



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

DOTTORATO DI RICERCA IN
SCIENZE E TECNOLOGIE DELLA SALUTE

Ciclo 37

Settore Concorsuale: 06/N1 - SCIENZE DELLE PROFESSIONI SANITARIE E DELLE
TECNOLOGIE MEDICHE APPLICATE

Settore Scientifico Disciplinare: MED/50 - SCIENZE TECNICHE MEDICHE APPLICATE

WHOLE BODY MRI (WB-MRI) IN PATIENTS AFFECTED BY MONOCLONAL
PLASMA CELL DISORDERS

Presentata da: Alice Rossi

Coordinatore Dottorato

Igor Diemberger

Supervisore

Alessandro Bevilacqua

Co-supervisore

Domenico Barone

Esame finale anno 2025

-INDEX-

ABSTRACT	3
INTRODUCTION	5
1.1 Diagnosis of Multiple Myeloma.....	5
1.2 Imaging in Myeloma.....	9
1.3 Radiomic study in oncology	19
PhD PROJECT	25
2.1 Study design.....	25
2.2 3T WB-MRI in myeloma: protocol set up	27
2.3 3T WB-MRI acceptancy.....	37
2.4 Inter reader agreement of WB-MRI in Myeloma.....	41
2.5 Prospective assessment of 3T WB-MRI and PET-CT in diagnosing multiple myeloma and its influence on patient care	53
2.6 Radiopsy: quantitative multiparametric WB-MRI for discrimination of Smoldering and Multiple Myeloma	63
REFERENCES	74

-ABSTRACT-

Purpose

This PhD research aimed to evaluate the impact of introduction of multiparametric Whole Body-Magnetic Resonance Imaging (WB-MRI) evaluation in patients affected by myeloma. It focused on diagnostic accuracy of WB-MRI with 18F-Fluorodeoxyglucose Positron Emission Tomography/Computed Tomography (PET-CT) and assessing its impact on clinical management. Additionally, it explored quantitative WB-MRI using advanced radiomics techniques.

Materials and Methods

The study was conducted within the AccuMRI IRST protocol, enrolling 177 patients from October 2020 to January 2024, with 134 patients undergoing both WB-MRI and PET-CT within one month. Clinical and laboratory data related to myeloma disease were collected. The first objective was to assess diagnostic accuracy of WB-MRI and PET-CT in the detection of bone marrow involvement (BMI) in patients with multiple myeloma (MM) and high-risk smoldering multiple myeloma (HR-SMM). Furthermore, inter reader agreement of 3 radiologists with different level of expertise was performed on reporting WB-MRI in 52 patients. Additionally, an evaluation of WB-MRI acceptability by 134 patients was performed. The second objective focused on explore quantitative WB-MRI, leveraging radiomics to extract imaging biomarkers, specifically on Apparent Diffusion Coefficient (ADC) and Fat Fraction (FF) maps, in a cohort of 84 patients (45 MM, 39 HR-SMM) to differentiate between HR-SMM and multiple myeloma MM. 6 Volumes of interest (VOI) were placed both in FF and in ADC maps on pelvic bone (4) and spine (2). RI and RS are placed around side of bone marrow trephine. Imaging data were analysed with first, second and higher order quantitative imaging biomarkers focusing on bone marrow biopsy sites. Logistic regression models were developed using LASSO feature selection and internally validated with cross-validation.

Results

WB-MRI demonstrated a sensitivity of 100% (66/66) and specificity of 97% (31/32) for BMI detection in MM, while PET-CT showed 89% sensitivity (59/66) and 97% specificity (31/32) ($p=0.02$ for sensitivity). Additionally, patients with BMI in at least one imaging modality compared to those without BMI in both WB-MRI and PET-CT,

demonstrated: in HR-SMM higher blood paraprotein levels ($p=0.01$), in MM higher blood paraprotein level ($p=0.007$) and lower-level haemoglobin ($p=0.002$). WB-MRI had a significant impact on clinical management, with treatment modifications consistent in 97% of cases with WB-MRI findings compared to 61% for PET-CT.

Inter-reader agreement for BMI evaluation using MY-RADS was excellent (Brennan and Prediger > 0.81) across most regions, with ICC values > 0.9 for FF and ADC evaluations except in the limbs (ICC=0.78 for ADC). Overall, WB-MRI was appreciated by the majority of patients (121/134, 90.3%), 68.7% of the patients preferred WB-MRI over other imaging modalities. The preference for imaging modalities was age-dependent (ANOVA $p = 0.011$), while (Chi-squared $p > 0.05$) was independent of sex and a primary cancer site.

Our preliminary radiomics models demonstrated robust performance in distinguishing between HR-SMM and MM, with a median AUC of 0.80 in the training phase and 0.70 in the test phase. The RS VOI near the biopsy site showed the highest predictive accuracy (AUC = 0.76 in test). The predictive power remained consistent across distant VOIs, with an AUC ranging from 0.69 to 0.78. Furthermore, the mean intensity of both FF and ADC sequences correlated moderately with the bone marrow plasma cell percentage (PC%), demonstrating the potential of radiomics to provide complementary information to standard biopsy-based biomarkers. The radiopsy model based on WB-MRI outperformed mean rFF alone but was inferior to plasma cell infiltration percentage for diagnostic accuracy.

Conclusion

This study demonstrates that WB-MRI provides superior diagnostic accuracy compared to PET-CT for detecting BMI in myeloma patients and significantly impacts clinical management. Quantitative WB-MRI, combined with radiomics, offers a promising non-invasive approach for improving risk stratification and guiding treatment decisions. Additionally, MY-RADS criteria for WB-MRI interpretations are reproducible, with excellent inter-reader agreement. WB-MRI is a clinically valuable, patient-friendly imaging modality for managing multiple myeloma.

-INTRODUCTION-

1.1 DIAGNOSIS OF MULTIPLE MYELOMA

Multiple myeloma (MM) is the second most common hematologic cancer. It accounts for approximately 1% of cancers and 10% of all hematologic malignancies. MM originates in plasma cells, which are a vital component of the immune system responsible for producing antibodies. In MM, these plasma cells become malignant, proliferating uncontrollably and leading to various health complications including bone lesions, anaemia, increased infections and kidney damage or failure [1] (Figure 1.1).

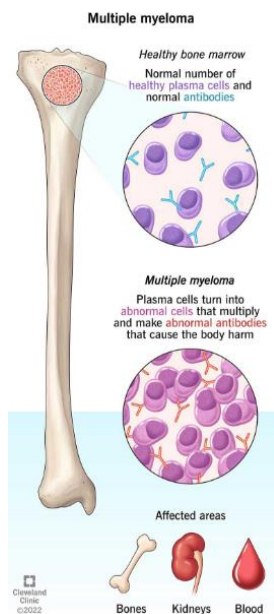


Fig.1.1
General characteristics of myeloma
<https://my.clevelandclinic.org/health/diseases/6178-multiple-myeloma>

Diagnosis typically involves a combination of blood tests to detect abnormal proteins (M proteins), urine tests for Bence Jones proteins, and bone marrow biopsy to identify malignant plasma cells [2,3]. The prognosis of MM varies significantly based on several factors, including the stage at diagnosis and individual response to treatment. The five-year survival rate is approximately 54% in the United States, but this can range from 40% to 82% depending on specific circumstances [4]. Early diagnosis and treatment are crucial for improving outcomes.

MM evolves from a clinically silent premalignant stage termed monoclonal gammopathy of undetermined significance (MGUS) with a risk of progression to MM or related malignancy at a rate of approximately 1% per year. A small subset of patients

has an intermediate clinical phenotype between MGUS and MM, and they are referred to as having smouldering multiple myeloma (SMM) that has a risk of progression of approximately 10% per year. In 2014 International Myeloma Working Group (IMWG) updated criteria for the diagnosis and treatment of these pathologies (Table 1.1) [3,5].

Disorder	Diagnostic Criteria
Monoclonal Gammopathy of Undetermined Significance (MGUS)	<ul style="list-style-type: none"> - M-protein: Serum monoclonal protein <3 g/dL. - Bone Marrow Involvement: Clonal plasma cells <10% in the bone marrow. - Absence of End-Organ Damage: No evidence of hypercalcemia, renal insufficiency, anaemia, or bone lesions (collectively known as CRAB features).
Smoldering Multiple Myeloma (SMM)	<p>Both criteria must be met:</p> <ul style="list-style-type: none"> • Serum monoclonal protein (IgG or IgA) ≥ 3 gm/dL, or urinary monoclonal protein ≥ 500 mg per 24 h and/or clonal bone marrow plasma cells 10–60% • Absence of myeloma defining events or amyloidosis
Multiple Myeloma (MM)	<p>Both criteria must be met:</p> <ul style="list-style-type: none"> • Clonal bone marrow plasma cells $\geq 10\%$ or biopsy-proven bony or extramedullary plasmacytoma • Any one or more of the myeloma defining events

Tab.1.1: IMWG diagnostic criteria for monoclonal plasma cells disorders.

Myeloma defining events are represented by the SLiM CRAB criteria, an updated framework for diagnosing active multiple myeloma established by IMWG. SLiM introduces specific biomarkers that can indicate multiple myeloma even in the absence of the classic end-organ damage CRAB symptoms, allowing for earlier diagnosis and intervention.

SLiM CRAB Criteria:

CRAB stands for the classic end-organ damage indicators associated with multiple myeloma:

- Calcium elevation: Serum calcium >11.5 mg/dL.

- Renal insufficiency: Serum creatinine >2 mg/dL or creatinine clearance <40 mL/min.
- Anaemia: Haemoglobin <10 g/dL or a decrease of >2 g/dL from baseline.
- Bone lesions: Presence of one or more osteolytic lesions seen on imaging (e.g., X-ray, MRI, or PET/CT).

SLiM Biomarkers

- S: Clonal plasma cells $\geq 60\%$ in the bone marrow.
- Li: Serum involved/uninvolved free light chain ratio ≥ 100 , with the involved free light chain level being ≥ 100 mg/L.
- M: More than one focal lesion on MRI, each ≥ 5 mm in size.

The incorporation of SLiM biomarkers allows for earlier diagnosis and treatment of multiple myeloma, potentially before significant organ damage occurs. This is crucial as patients identified under these criteria may have a high risk of progression to symptomatic disease. The presence of at least one SLiM biomarker or any CRAB feature is sufficient for diagnosing active multiple myeloma. Studies have shown that patients meeting SLiM criteria have an approximately 80% risk of developing myeloma-related organ damage within two years.

High-risk smoldering multiple myeloma (HR-SMM) represents a critical phase in the continuum from monoclonal gammopathy of undetermined significance (MGUS) to active multiple myeloma. HR-SMM is characterized by specific biomarkers that indicate a significant likelihood of progression to MM. Studies indicate that: the 2-year risk of progression for patients with $\geq 60\%$ clonal plasma cells can be as high as 95% in some cohorts; for those with an FLC ratio ≥ 100 , the 2-year progression risk is approximately 72%, with median time to progression (TTP) around 15 months. A systematic review and meta-analysis demonstrated that patients meeting high-risk criteria have longer TTP and lower progression risks than previously reported [6]. They may benefit from early therapy with lenalidomide as a single agent prolonging time to symptomatic MM with end-organ damage [7,8]. The SLiM CRAB criteria provide a structured framework for identifying these patients, enabling better risk stratification and management strategies aimed at improving outcomes. Regular monitoring and assessment are key components in managing this patient population effectively.

The SLiM CRAB criteria represent a departure from traditional management approaches that delayed treatment until the onset of symptoms. By recognizing

biomarker-defined myeloma as a legitimate disease state, clinicians can now intervene earlier to manage high-risk patients, with the potential to improve both outcomes and quality of life. As a matter of fact, end-organ damage in MM includes osteolytic bone lesions and renal failure, at times, they are not reversible, and they can cause significant morbidity to patients. Bone disease affects quality of life and represent the major cause of morbidity and mortality; as a matter of fact, 70% of MM patients has bone involvement at staging and 80-90% develop bone lesion during the disease [9]. As research continues to evolve, these criteria will likely play a pivotal role in shaping future therapeutic approaches in multiple myeloma care.

1.2 IMAGING IN MYELOMA

Bone disease is the most frequent feature of MM and impairs patients' quality of life representing a major cause of morbidity and mortality. Because 80% to 90% of all patients with MM develop bone marrow (BM) involvement, imaging plays a very important role in the management of MM. Imaging is pivotal in the accurate differentiation between solitary plasmacytoma and MM, it is necessary for detection of focal lesions that represent one of the myeloma defining events (MDE) as well as to predict the risk of progression from HR-SMM to active MM. Additionally, it could identify extramedullary disease (EMD) which is recognized as unfavourable prognostic feature [9,10].

A conventional radiographic skeletal survey was historically used for the assessment of bone lesions in patients with MM. Many studies have shown higher sensitivity for the detection of focal lesions with whole-body low-dose computed tomography (WB-LD-CT), ¹⁸F-fluorodeoxyglucose (FDG)-positron emission tomography-computed tomography (PET-CT) and whole-body magnetic resonance imaging (WB-MRI) compared to conventional radiographic skeletal surveys [9, 11]. Even if the diagnosis and response assessment primarily depend on the clinical and laboratory criteria [12], total-body imaging techniques are an indispensable element in the staging and post-treatment re-evaluation due to an accurate assessment of the degree of bone marrow involvement, replacing the conventional skeletal survey. This led to the introduction of these imaging methods into international guidelines. IMWG suggest a range of possible imaging investigations including CT, PET-CT, and MRI [13]. In the IMWG recommendations, WB-LD-CT is central in initial patient assessment for the detection of osteolytic lesions, but its interpretation is not so simple since focal lytic MM lesions on CT are difficult to distinguish from osteopenia and degenerative phenomena that are often present. An additional interpretative challenge on CT is the differential diagnosis of soft tissue attenuation in the bone marrow cavity that could be due to both replacement of red marrow by yellow marrow and pathologic tissue. PET-CT is considered as an alternative method at staging in MM patients. They are used as a triage method for patients eligible for WB-MRI in instances of negative or inconclusive results. [10,14]. IMWG recommends WB-MRI for all patients with a suspected

diagnosis of SMM or MM and in case of suspected relapse because of its superior sensitivity and utility in identifying myeloma-defining events. Furthermore, PET-CT is recommended as the modality of choice for assessing treatment response. WB-MRI is the first line imaging method for bone solitary plasmacytoma, whilst PET-CT in case of extramedullary plasmacytoma [13]. In the United Kingdom, WB-MRI is recommended as first line imaging for all patients with a suspected new diagnosis of myeloma [15]. The British Society for Haematology additionally recommended the use of WB-MRI for monitoring response of non-secretory myeloma, oligo-secretory myeloma, and for those patients with extramedullary disease [16].

In patients with MM, more than one Focal Lesion (FL) on MRI that is at least 5mm or greater in size is include in SLiM CRAB criteria as MDE for establish need of systemic treatment; this criterion allows the identification of bone lesions before the osteolytic damage visible on CT. Osteolytic lesions are visible on CT when bone destruction has already occurred and affected at least 30% of the bone trabeculae. Focal lesions represent foci of plasma cells accumulation within bone marrow and can be detected before destruction of mineralized bone [17].

Moreover, restrictions in the choice of modality have to be addressed. First, local availability and costs favour CT [15]. Second, patient characteristics may influence the choice. Although the number MRI-compatible metallic implants and medical devices is increasing, the technique is still contraindicated in patients with several older or incompatible models, due to the risk of device failure and heating. Claustrophobia is an additional serious limitation to MRI. The more time-consuming WB-MRI examination may be challenging for patients with active MM who suffer from bone pain, as they are asked to lie as motionless as possible during image acquisition. Otherwise, motion artifacts may impede the interpretation of MR images. For some patients, MRI is the preferred imaging modality due to the need to differentiate benign from malignant vertebral fractures or to assess painful complications and spinal cord compression [18,19,20].

¹⁸F-FDG PET-CT is a functional whole body imaging modality to assess disease presence, response or progression by revealing changes in metabolic activity of cells within neoplastic tissue even without the evidence of morphological changes. The recently published Italian criteria for myeloma for PET use (IMPeTUs) provided a standardization of post therapy PET-CT interpretation, in particular the refinement of the definition of complete metabolic response and its integration into MM staging

systems. These criteria are a comprehensive evaluation including various semiquantitative parameters (SUV max, PET-FLs, CT lytic lesions, fracture lesions, PMD/ EMD disease) and are based on the 5 points Deauville scale (D5-PS), just as for lymphoma [21].

As for PET-CT, MRI evaluation should include the entire body because up to 10% of patients have only appendicular skeleton involvement, and up to 50% of the lesions would be missed performing only the MRI spine [22,23].

WB-MRI is a radiation-free and contrast administration-free imaging method for detecting bone and soft tissue pathology. It combines high quality morphological images with “functional” information. Diffusion-weighted imaging (DWI) is based upon measuring the random Brownian motion of water molecules mostly in extracellular space within a tissue [24]. It is emerging as a core sequence of WB-MRI protocols for disease assessment because of its sensitiveness to tissue cellularity and cell viability offering excellent lesion-to-background contrast and quantification of the degree of water motion by calculation of the apparent diffusion coefficient (ADC) map; changes in ADC can reflect variations in cellularity [25, 26]. DWI and ADC are being used clinically as qualitative (DWI) and quantitative (ADC) indicators of disease presence, progression or response to treatment. The biological model of bone-marrow tumour response proposed by Padhani et al [26, 27] noted that the impact on quantitative MRI measurements is influenced by the mechanism of action of the drug in question (mode of tumour cell death), the intrinsic susceptibility of the tumour to treatment, and early and late stromal reactions of the bone marrow to tumour cell death. In this scheme (Figure 1.2), re-emergence of fat is part of the bone marrow repair process. An abundance of yellow bone marrow fat causes low signal intensity on high b-value DW images. Increasing bone marrow cellularity of red bone marrow and water content increases signal intensity and paradoxically increases ADC values. Successful therapy results in increases in ADC values and decreases in signal intensity consistent with decreasing cellularity but T2-shine through effects following therapy would increase both signal intensity and ADC values.

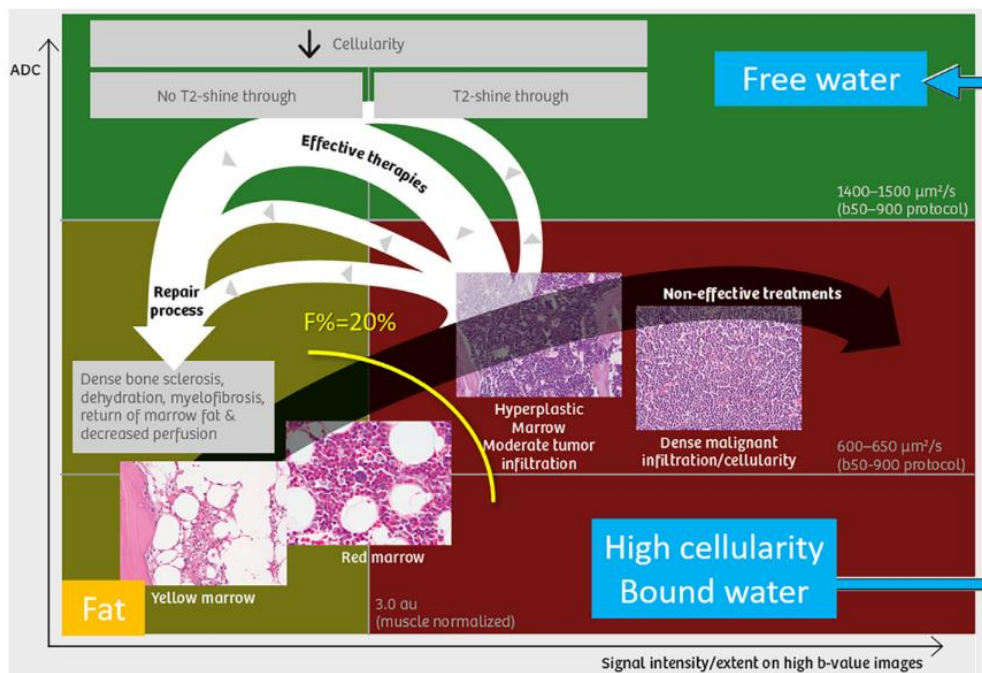


Fig.1.2. Changes in ADC values and high b-value signal intensity in response to therapy [27].

Relationships between bone marrow cell type and ADC; the green box represents tumour response with increased free tissue water, the burgundy box represents cellular tumour tissue with bound water, and the green-brown box represents tissue with absent viable tumour containing a variety of tissue elements. F% is fat fraction percentage.

Dixon-type pulse sequences T1 and the relative FF map, derived from fat and water images, allow quantitative measurement of fat in bone marrow and could be used to evaluate disease assessment [28,29,30]. The combination of anatomical sequences and at least two “functional” quantitative sequences made WB-MRI as multiparametric [31]. WB-MRI protocol should include DWI improve diagnostic sensibility and ADC and FF images in conjunction with all other sequences to maximise specificity in report. In WB-MRI, the quantitative assessment of bone lesions relies on ADC maps and FF maps, that allows both better baseline staging and response assessment in myeloma patients [10,32,33] (Figure 1.3).

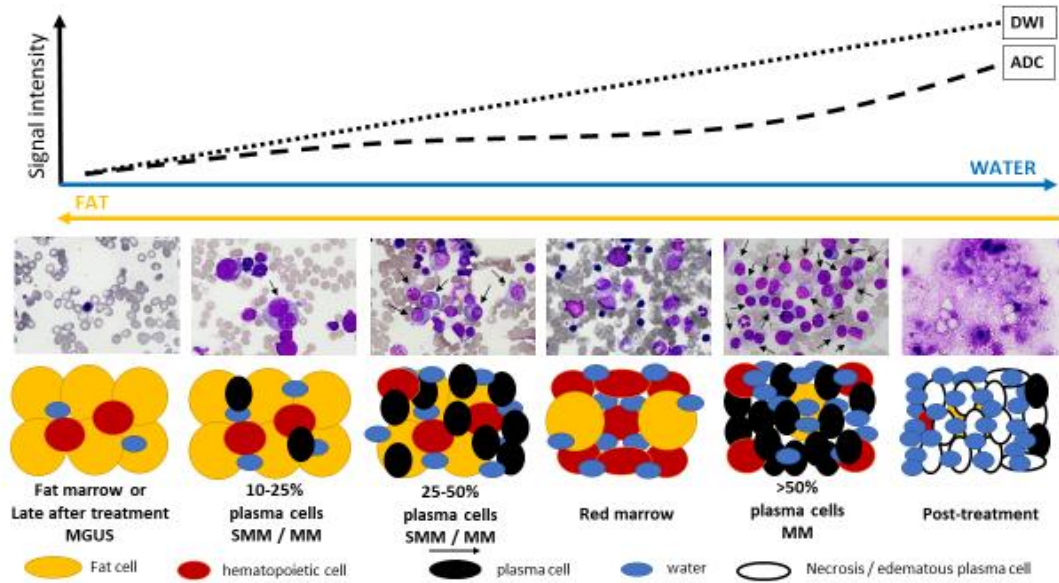


Fig.1.3: Changing in fat, water and plasma cell content during myeloma disease [33].

Interpretation of imaging studies of myeloma patients is often challenging. In the field of plasma cell disorders, application of the Guidelines for Acquisition, Interpretation, and Reporting of Whole-Body MRI in Myeloma: Myeloma Response Assessment and Diagnosis System (MY-RADS) is fundamental for standardization [23]. MR images show patterns of myeloma disease: normal, focal (focal lesions $\geq 5\text{mm}$), diffuse, focal on diffuse, micronodular (focal lesion $< 5\text{mm}$) and paramedullary/extramedullary tissue. Additionally, MY-RADS provides a mean for assessing response to treatment using Response Assessment Criteria (RAC), a Likert scale with five points. The ADC evaluation is included in RAC along with correlation with morphological information. For soft tissue disease, either paramedullary or extramedullary, are used as size criteria of Response Evaluation Criteria in Solid Tumors (RECIST) 1.1. Specifically, RAC 1 and 2 assess therapeutic response, RAC 3 stability, and RAC 4 and 5 progression of disease. It is proposed a structured report for WB-MRI with evaluation of a maximum of 5 “dominant” lesions and 2 RAC (primary for predominant pattern and secondary for other pattern in case of mix response) in each single skeletal region (skull, cervical spine, dorsal spine, lumbar spine, pelvis, extremity and ribs) and for paramedullary and extramedullary disease.

PET-CT and WB-MRI provide different and complementary information on the tissue under investigation giving information about BM in case of not only bone destruction,

but also assessing tumour burden and disease activity and they can identify bone disease before bone damage. Table 1.2 summarizes actual evidence about these methods [11].

	FDG PET-CT	WB-MRI
Scanning Time	15–20 min (including radiopharmaceutical injection- 60min)	40–60min
Contrast medium injection	yes	None
Radiation Exposure	6–15mSv	None
Standardisation and reproducibility	IMPeTUs criteria	MY-RADS guidelines
Focal bone lesions involvement	Higher Specificity	Higher sensitivity
Diffus bone marrow involvement		Highest sensitivity and specificity
Extramedullary and paramedullary tissues	Favored technique to assess EMD	Gold standard for skeletal-related events and spinal cord compression
Evaluation response to treatment	modality of choice Early changes in FLs (7 days)	Emerging data Early changes in ADC e rFF (21 days)
Impact on clinical decision	Lack of data (need further investigations). No statistically significant difference in the clinical decision to treat. Patient acceptance (WB-MRI>PET-CT)	
Prognostic Value (no direct comparisons between the two techniques)	FLs number, SUV max value, presence/absence of EMD	FLs number Diffuse pattern ADC and FF

Tab 1.2: pro and cons of WB-MRI and PET-CT in myeloma. Abbreviations: ADC (apparent diffusion coefficient); FDG PET-CT (fluoro deoxyglucose positron emission tomography computed tomography); FL (focal lesion); WB-MRI (multiparametric whole-body magnetic resonance imaging); SUVmax (maximum standardized uptake value); EMD (extramedullary disease).

A multitude of studies has examined performance of PET-CT relative to various MRI protocols.

Currently the high sensitivity of WB-MRI is explicitly acknowledged and WB-MRI with DWI (the protocol recommended by the IMWG for evaluation of monoclonal plasma cell disorders) is the most sensitive method to detect bone marrow infiltration for both bone focal lesions (FLs) and diffuse pattern and to evaluate extramedullary

tissues [10]. Moreover marrow biopsy results were more strongly associated with WB-MRI findings at the iliac crest [34]. PET-CT has higher specificity than WB-MRI and it is currently the modality of choice for assessing treatment response and for determining minimal residual disease-negative status [12,35,36].

Comparative studies have suggested that WB-MRI is more sensitive than PET-CT for the detection of BM infiltration at diagnosis of MM thanks to its high spatial resolution and sensitivity to diffuse pattern and to the lack of FDG uptake in a significant (10-15%) proportion of MM patients which has low-no uptake of FDG myeloma likely due to a low hexokinase-2 expression that is involved in the glycolytic pathway. WB-MRI lower specificity can be attributed to the lack of standardization of protocol and interpretation [34,37,38,39].

Findings on MRI also have prognostic value. The presence of more than 1 FL on baseline MRI is a poor prognostic indicator in terms of progression free survival (PFS); the presence of more than 7 lesions is associated with a significantly worse 5-year survival rate [10]. Although not currently considered a myeloma-defining event, a diffuse pattern of bone marrow infiltration is considered an indicator of poor prognosis and correlates with measures of tumour burden and markers or risk [39,40]. Recent studies indicate that the presence of more than three large focal lesions ($>5\text{ cm}^2$) on WB-MRI is an independent prognostic factor for both PFS and overall survival (OS). PET-CT also has a prognostic value; parameters such as the number of FLs, maximum standardized uptake value (SUVmax) and the presence of EMD correlating with shorter OS and PFS [11].

Response assessment in MM primarily depends on measurement of monoclonal protein secreted by plasma cells, in the serum or urine [3]. However, precise imaging disease assessment and evaluation response to anticancer therapy is fundamental in the optimal management of patients. WB-MRI and PET-CT are the modalities of choice also to evaluate minimal residual disease (MRD). Enhanced with DWI and ADC mapping, WB-MRI allows for the early identification of bone lesions and better differentiation between active and inactive disease; post-treatment changes in ADC values following chemotherapy can reflect reductions in tumour cellularity and these quantitative measurements can help identify residual disease even in the absence of morphological changes. Studies have shown that an increase in ADC values is associated with a favourable treatment response, while stable or decreasing ADC values may indicate persistent or progressive disease [11]. In recent original article, Belotti et al.

investigated the prognostic role of RAC criteria in 64 newly diagnosed MM patients after autologous stem cell transplantation (ASCT), and they combined the results of MY-RADS with those of minimal residual disease (MRD) assessment by multiparametric flow cytometry (MFC). They demonstrate that RAC criteria were able to independently stratify patients and to better predict their prognosis and the combined use of WB- MRI with MFC allowed a more precise evaluation of MRD. Patients with complete imaging response (RAC1) had better overall survival (OS) and post-ASCT progression free survival (PFS) than patients with imaging residual disease (RAC \geq 2) [41].

Given the heterogeneous data in this topic and the heterogeneous characteristics of MM, recent original articles, meta-analysis and state of the art suggest that it remains unclear which modality may be the best for response assessment and, maybe, a complementary role of WB-MRI and PET-CT could be considered for an optimal work-up [38,42, 43,44].

Quantitative imaging biomarkers (QIB) are emerging research fields aiming to translate quantitative imaging research into clinical practice. A biomarker is a measure: a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacological responses to a therapeutic intervention. An imaging biomarker is a biological feature, or biomarker detectable in an image relevant to a patient's diagnosis [45,46].

The data of WB-MRI (DWI/ADC and FF) may be used as novel possible QIB to evaluate disease at staging and response to therapy; in particular, MY RADS guidelines allow response assessment including DWI/ADC [23]. DWI and ADC are being used clinically as qualitative indicator of disease presence, progression or response to treatment [47,48].

Moreover, interesting recent publications underline how FF evaluation can aid in the differential diagnosis of malignant vertebral collapse and malignant focal lesion. Malignant lesions present high signal intensity on high b-value DWI, pathologic ADC values due to increased tumour cellularity and decreased fat content. In the study of Castagnoli et al [49] malignant bone lesions showed significantly lower median FF (13.87%) compared with normal bone marrow (89.76%) without significant difference in the median FF of malignant lesions from breast (14.46%), myeloma (13.12%) and prostate cancer (13.67%) ($p > 0.017$). Donners et al [50] found that combination high

DWI signal, mean ADC < 1100 $\mu\text{m}^2/\text{s}$ and mean FF < 20% can identify active bone metastases.

Schemeel et al [51] demonstrated that the ADC and proton density fat fraction (PDFF) values of malignant lesions were significantly lower compared to benign lesions (mean ADC $861 \times 10^{-6} \text{ mm}^2/\text{s}$ vs. $1323 \times 10^{-6} \text{ mm}^2/\text{s}$, $p < 0.001$; mean PDFF 3.1% vs. 28.2%, $p < 0.001$).

In this scenario, the data published so far on whole body quantitative imaging are particularly promising also in evaluation of response to therapy. Early changes in FF have been shown to predict response and it seems likely that ADC and FF will be complementary metrics [52,53,54]. Particularly noticeable is that increase in bone marrow FF and ADC values are most discriminant features and outperform decrease in size of FLs (IMWG criteria) to distinguish responders from non-responders' patients [38].

Although increasingly used and recommendation by international guidelines, WB-MRI usage has been confined mainly to expert centres. Important steps in the process of ensuring uniformity in the acquisition, interpretation, and reporting of WB-MRI are needed.

A recent Italian survey on WB-MRI in oncology involving 48 members demonstrated that WB-MRI was mainly performed at 1.5T MRI, with lymphoma, myeloma and prostate cancer having been the most common indications. The extreme variability in the choice of imaging protocols and use of contrast agents demonstrates the need of a standardization of WB-MRI application in clinical practice [55].

Recent publications support the standardization of WB-MRI through implementation of MY-RADS protocol involving scanners with a range of manufacturers, models and field strengths. Quantitative measures of image quality were developed and shown to be significantly correlated with radiological assessment. The OPTIMUM trial supports the standardization of ADC as a quantitative biomarker for assessing treatment response. Their rigorous protocol harmonization across multiple centres, which included DWI and T1-w Dixon sequences, confirms that ADC can reliably reflect minimal residual disease and changes in disease activity. By aligning protocol with similar recommendations, we can ensure that ADC measurements are both consistent and clinically meaningful [56,57]. Recommendations of UK quantitative WB-DWI technical workgroup and of Quantitative Imaging Biomarkers Alliance (QIBA) both for FF and for DWI support the standardization and efficacy [58,59,60].

It is important also training of radiologists in interpretation; a proposed learning curve of WB-MRI in MM demonstrated that 80 reports lead to a high level of inter-observer concordance [61].

1.3 RADIOMIC STUDY IN ONCOLOGY: Key Methodological Steps and Considerations

Radiomics is a promising field that enables the extraction of high-dimensional quantitative features from standard medical imaging modalities, such as CT, MRI, and PET. This approach is based on the idea that medical images contain intrinsic information regarding the biological behaviour, structure, and microenvironment of tumors, which are not readily discernible through visual inspection [62,63,64].

By mining these engineered or handcrafted features, which relate to image intensity, shape, and texture, derived from predefined mathematical transformations, radiomics aims to provide additional diagnostic, prognostic and predictive insights, potentially revolutionizing cancer care. However, to harness the full power of radiomics, studies must be methodologically rigorous, addressing several critical stages towards its multistep process (Figure 1.4).

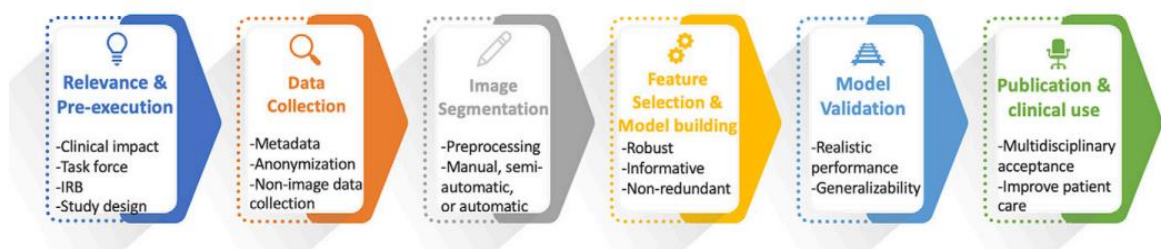


Fig. 1.4 Ideal radiomics workflow from conceptualization to clinical implementation [64].

1. Clinical Relevance and Study Design

The first step in any radiomics study is the formulation of a clear and clinically relevant hypothesis. This requires the involvement of a multidisciplinary team, including radiologists, oncologists, data scientists, bioinformatics... These ensure a comprehensive understanding of clinical need and real-world applications [64]. The primary goal should be the development and validation of a radiomics model or signature that addresses an unmet clinical need, such as improving risk stratification, the prediction of treatment response or survival outcomes in oncology [65].

Defining this clinical endpoint is crucial for guiding all subsequent steps in the study.

Once the clinical relevance is established, the following planning stages should be considered: adhere to specific radiomics checklist as Radiomics Quality Score (RQS) or METHodological radiomICs Score (METRICS) [63,65], consider multicentric collaboration and sample size (Figure 1.5). For radiomics studies, particularly those building machine learning models, ensuring sufficient sample size is crucial to avoid overfitting.

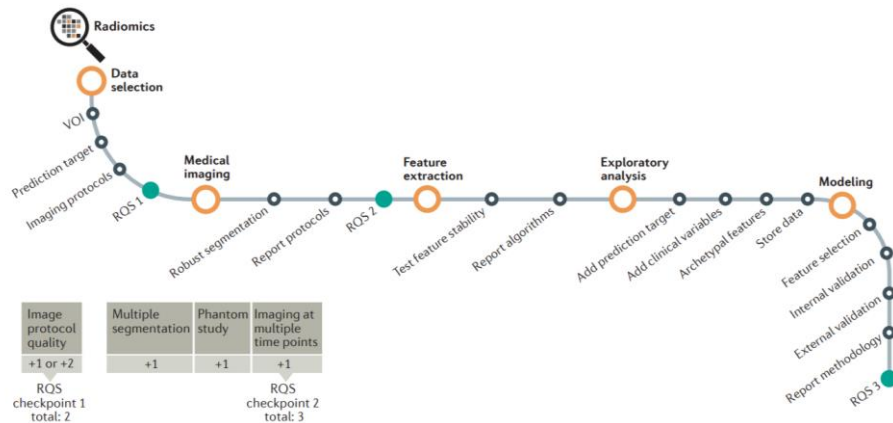


Figure 1.5: Flowchart depicting the workflow of radiomics and the application of the RQS [63]

2. Data collection

Accurate and reproducible data acquisition is the cornerstone of a successful radiomics study. High-quality imaging is necessary to capture subtle differences in tumor characteristics, and consistent imaging protocols across patients and timepoints are essential for minimizing variability [64,65]. Ideally, studies should adhere to well-documented imaging protocols are crucial to augment reproducibility and to eliminate confounding variability. Phantom studies can be used to assess the inter-scanner and inter-vendor variability of imaging features. Additionally, repeat scans (test-retest studies) should be conducted where possible to measure the stability of radiomic features over time [63].

Finally, engaging all stakeholders in data collection is integral to the non-image data collection process.

3. Image Segmentation and Preprocessing

After data acquisition, the next crucial step is image segmentation, which involves delineating the region of interest (ROI) or volume of interest (VOI) from which radiomic features will be extracted. This step can be performed manually, semi-

automatically, or fully automatically, but the method chosen significantly impacts the robustness of the study [64]. Accurate segmentation ensures that only the relevant areas of the image (such as the tumour and its peritumoral region) are analysed. However, variability in segmentation between different observers (inter-reader variability) and even the same observer over time (intra-reader variability) can introduce biases (METRICS).

Pre-processing of the images is another key consideration. Harmonization techniques like the ComBat algorithm have been proposed to mitigate these differences, but the lack of universal imaging standards is a major roadblock [65].

4. Feature Extraction, feature selection and model building

Once the ROI or VOI is segmented, quantitative features can be extracted. Feature extraction is a high-throughput process, meaning that hundreds to thousands of features can be derived from a single image. These features are then stored in a database and linked to the clinical or genomic data [63]. Importantly, not all extracted features will be useful; features must be carefully selected, and pre-processing steps (e.g., image normalization) should be clearly reported to avoid overfitting. The use of standardized software for feature extraction enhances reproducibility.

It is crucial that this process is carried out with rigor to ensure feature robustness and reliability. The following considerations are essential in this phase [63,65]:

- **Transparent Feature Extraction:** The use of standardized feature extraction software and clear documentation of the parameters involved in feature extraction is critical. Studies should explicitly mention the software and methods used for feature extraction, and, when possible, employ open-source tools to ensure reproducibility.
- **Dimensionality Reduction:** Given the high-dimensional nature of radiomics data, reducing the number of features is necessary to prevent overfitting and improve model generalizability. Feature selection techniques, such as principal component analysis (PCA) or clustering, should be employed to identify robust features that are most predictive of clinical outcomes. The METRICS and RQS guidelines both highlight the importance of appropriate dimensionality reduction techniques to avoid overfitting, which can jeopardize the performance of models in external datasets.

- **Elimination of Redundant and Non-Robust Features:** In addition to dimensionality reduction, the removal of redundant and non-robust features—those that are highly correlated with one another or sensitive to variations in image acquisition—is essential. This step ensures that only features that add distinct value to the model are retained.
- **Feature Selection:** Feature selection must be handled with care, particularly when the number of extracted features exceeds the available sample size. Dimensionality reduction techniques, such as principal component analysis (PCA), can help by reducing the number of features without losing essential information. Additionally, radiomics studies should integrate other clinical data (e.g., genomic, histopathologic) to develop more holistic predictive models.
- **Model Building:** a variety of machine learning algorithms, such as Least Absolute Shrinkage and selection (LASSO) regression, can be used to build predictive models from radiomic features before applying ANOVA to analyse variance.

5. Model Validation

Model validation is essential for ensuring the reproducibility, generalizability and reliability of the radiomics model to predict the desired end point. Internal validation can be achieved using techniques such as cross-validation or bootstrapping, but external validation—testing the model on independent datasets from different institutions or populations—is the gold standard to confirm that the model performs well in real-world settings, including different institutions, imaging devices, and patient populations [63,64,65]. The use of external datasets is particularly important in radiomics, where differences in scanners, imaging protocols, and patient populations can greatly affect model performance. Calibration and discrimination metrics, such as the area under the receiver operating characteristic curve (AUC), should be reported to evaluate the model’s predictive accuracy [63,65].

Ensuring sufficient sample size is crucial to avoid overfitting particularly those building machine learning models. The one-third rule is commonly applied [64,65]:

- **Training and validation sets:** For model development, the sample size must be large enough that approximately one-third of the data is used for validation. For

instance, in a radiomics model with 10 features, a minimum of 133 samples may be required, with 100 for training and 33 for validation

- External validation: External validation requires additional datasets from different populations, making multi-center studies important for ensuring generalizability. This suggests that larger sample sizes, potentially 200 or more patients, may be needed for comprehensive model validation
- Transparency is also key. Researchers are encouraged to publish their data, code, and models in open-access repositories, allowing for independent verification and replication of their work

Radiomics research must move beyond retrospective studies. Prospective clinical trials, incorporating radiomics models as decision support tools, are necessary to validate their utility and efficacy in real-world oncology settings.

Figure 1.6 shows the summary of common mistakes within radiomics study and suggested strategy to overcome them [64].







	Common Mistakes	Strategies to overcome the mistakes
 Relevance & pre-execution	Lack of integration among stakeholders, suboptimal sample size, overlook of potential bias	Early multidisciplinary engagement, critical study feasibility review, institutional and societal support for multicentric studies
 Data collection	Absence of streamlined workflow, low quality input data, heterogeneous data	Improved workflow, standardized protocols and reports, protected time to specialists involved on data collection
 Image segmentation and preprocessing	Segmentation errors, inter- and intra-reader variability, and lack of integration with PACS	Pre-segmentation training, automatic segmentation tools, quality control
 Feature extraction, selection and model building	Lack of adherence to standardization, feature extraction errors, incorrect feature selection (over- or underfitting)	Prefer standardized and robust feature extraction, careful feature selection, nested cross-validation from small data, consider explainability
 Model validation	Lack of transparency, data leakage to the validation set	Rigorous, systematic, and transparent approaches, careful data handling, external validation
 Publication and clinical use	Bias favoring positive results, focus on model accuracy rather than workflow implementation and reproducibility	Promote standardized guidelines, resources dedicated to improved workflow

Fig. 1.6 common mistakes and possible strategies to overcome them in oncological setting [64].

Conducting a radiomics study involves several critical steps, each requiring careful attention to methodological detail to ensure robustness and clinical applicability. From the initial study design and data acquisition through feature extraction, model building, and validation, each phase presents unique challenges and opportunities. Addressing these challenges through standardized methodologies, rigorous validation, and open science practices will be key to fully realizing the potential of radiomics in personalized oncology.

With sustained multidisciplinary collaboration and a focus on real-world validation, radiomics has the potential to revolutionize cancer management and improve patient outcomes.

Moreover, as highlighted in the assessment of diagnostic imaging [66], demonstrating the clinical value of any imaging technique—radiomics included—requires a combination of high-quality research data and real-world evidence. To ensure that radiomics is not only technically sound but also clinically effective, it must undergo rigorous clinical testing similar to other diagnostic tools. This underscores the necessity of integrating both robust clinical trials and real-world data to validate its utility in patient care.

-PhD PROJECT-

2.1 STUDY DESIGN

The aim of this study is to evaluate the impact of introduction of WB-MRI, performed on 3 T scanner according to MY-RADS guidelines [23], in our institution in terms of diagnostic accuracy, in qualitative and quantitative terms, compared to clinical and laboratory data (IMWG criteria) and to PET-CT, in 3 cohorts of patients:

COHORT 1

Newly diagnosed patients affected by HR-SMM.

COHORT 2

Patients affected by a first diagnosis of MM

COHORT 3

Patients affected by relapsed or refractory MM (RR-MM)

This is a single centre, observational, prospective, study. The study was approved by C.E.R.O.M comitato etico della Romagna (AccuMRI trial IRST code 100.15). All investigations were performed by routine clinical practice. All participants signed an informed consent to be enrolled in the study, in addition to the standard consent required by law for MRI scans. From October 2020 to January 2024, we prospectively enrol 177 myeloma patients, 25 patients affected by metastatic prostate cancer, 54 patients affected by advanced breast cancer and other 51 patients affected by other malignancy (19 melanoma, 8 neuroendocrine tumours, 7 lymphoma, 5 ovarian cancer, 4 renal cancer, 2 gastric cancer, 2 thyroid cancer, 1 gastrointestinal stromal tumour, 2 paraganglioma) and 3 patients affected by Ly Fraumeni syndrome.

Primary objectives:

- To optimize multiparametric WB-MRI along MY-RADS
- Investigating patients' acceptance of WB-MRI
- To compare the diagnostic accuracy of WB-MRI with PET-CT in conjunction with clinical and laboratory data following IMWG criteria [3] in evaluation of bone marrow infiltration in patients with myeloma

- Quantitative imaging WB-MRI evaluation of bone marrow infiltration compared to clinical-laboratory data and PET-CT in staging and follow-up patients affected by plasma cell disorders
- To evaluate the impact of WB-MRI on decision making and therapeutic management of patients
- To evaluate inter reader agreement of WB-MRI

Secondary objective:

Explorative analysis of the potential value of radiomics features for risk stratification (pilot study)

2.2 3T WB-MRI IN MYELOMA: PROTOCOL SET-UP

General consideration about MR Imaging at 3 T versus 1.5 T

The transition from 1.5 Tesla (T) to 3T MRI scanners presents several significant differences in terms of image quality, safety, cost and accessibility. Understanding these factors is crucial for healthcare providers and patients alike.

Image Quality

Magnetic Field Strength and Signal-to-Noise Ratio: The primary difference is the magnetic field strength; a 3T MRI has twice the strength of a 1.5T MRI. This increased strength enhances the signal-to-noise ratio (SNR), leading to clearer and more detailed images, which is particularly beneficial for detecting small lesions and intricate structures within the body [67]

Time scan: The generally faster scan times associated with 3T machines can enhance patient comfort and reduce anxiety associated by reducing the duration of the procedure. This efficiency is particularly advantageous in high-volume medical settings [67]

Resolution: The enhanced magnetic field of a 3T scanner allows for the acquisition of images with greater clarity and detail. This is especially beneficial in applications for specialized imaging needs, such as neurological assessments, vascular studies, and examinations requiring high spatial resolution, while 1.5T MRIs are often adequate for general imaging purposes [67,68]

Artifacts [67]: While 3T scanners provide superior image quality, they may also introduce more artifacts - unwanted features in the images - due to the stronger magnetic field's sensitivity to movement of blood and fluid within the body. These artifacts can sometimes complicate the interpretation of scans, particularly in patients with certain implants or conditions

Common Artifacts in 3T MRI Images and their minimization:

- **Susceptibility Artifacts:** These artifacts arise from differences in magnetic susceptibility between adjacent tissues, particularly near metallic implants or air-tissue interfaces. They appear as distortions or signal voids in the images. They can be reduced using spin echo sequences instead of gradient echo sequences, which are less sensitive to susceptibility effects. Additionally, employing techniques like iterative decomposition of water and fat (IDEAL) can help

mitigate these artifacts by improving fat suppression and separating fat and water signals effectively

- **Motion Artifacts:** Caused by patient movement, breathing, or pulsation of blood and cerebrospinal fluid, these artifacts can manifest as ghosting or blurring in the images. Strategies for minimization include patient immobilization, using cardiac or respiratory gating techniques, and applying saturation bands to reduce the effects of motion. Fast imaging techniques such as echo-planar imaging (EPI) can also help reduce motion-related artifacts
- **Aliasing Artifacts (Wrap-Around):** It occurs when anatomical structures outside the field of view are incorrectly mapped onto the image, appearing at the opposite edge. It can be mitigated by increasing the field of view (FOV) during image acquisition to ensure that all relevant anatomy is captured within the imaging area
- **Truncation Artifacts:** Result from under-sampling data during image acquisition, leading to sharp edges between areas of high contrast being misrepresented. It can be minimized by increasing the matrix size (more encoding steps) and reducing the FOV can help alleviate truncation artifacts
- **Cross-Excitation Artifacts:** Caused by overlapping excitation of adjacent slices during sequential acquisitions, leading to decreased signal intensity in overlapping regions. Minimization is obtained through increasing slice gaps and using interleaved slice acquisition can help reduce these artifacts
- **Zipper Artifacts:** Result from interference by external radiofrequency (RF) sources, appearing as bright or dark lines on the image. There is the need to ensure that all external RF sources are removed from the scanning environment and inspect for any breaches in RF shielding before scans

Safety Considerations

Implants and Foreign Bodies: One of the significant challenges with 3T MRI scanners is their interaction with metallic implants. Many devices that are safe in a 1.5T environment may not be safe at 3T due to the stronger magnetic field. This necessitates more rigorous screening protocols for patients with implants [69,70,71]

Specific Absorption Rate (SAR): The SAR, which measures how much energy the body absorbs during an MRI scan, is higher in 3T scanners. This can lead to increased heating of tissues, making it less suitable for certain patients [67]

Facilities equipped with 1.5T MRI scanners are more common, making them easier to access for patients needing imaging services. In contrast, while the number of 3T scanners is increasing, they may not be as readily available in all regions. 3T MRI machines are generally more expensive to purchase and maintain, often costing up to 30-40% more than their 1.5T counterparts.

WB-MRI sequences and protocol

All the WB-MRIs are performed in a 3T MRI scanner (Ingenia Philips, Eindhoven, Netherlands). A head/neck helmet-like coil and a flexible surface body coil are applied to patients for WB-MRI scans. No intravenous gadolinium contrast was administered.

The typical duration of our WB- MRI examination is 45 min.

The WB-MRI basic protocols at our institution are based on MY-RADS guidelines [23] and consist of:

- sagittal T1-weighted of the whole spine
- sagittal STIR T2-weighted sequences of the whole spine
- axial diffusion-weighted whole-body imaging with background body signal suppression (DWIBS) technique at two b-Value (b50 e b800) from vertex to knees
- axial TSE T2 from skull base to knees
- axial T1 Gradient Echo (GRE) mDixon sequences from vertex to knees (in phase, out of phase, fat only, water only) from vertex to knees

Additional segments and/or sequences can be added according to clinical needs.

Post-processing [71]:

- From the DWI and the Dixon-type images, the Apparent Diffusion Coefficient (ADC) map and the relative Fat Fraction (rFF) maps are respectively reconstructed. ADC is calculated with a monoexponential fitting to the signal intensities in the different b-value DWI images.
- The rFF maps were calculated by $(\text{fat- only image})/(\text{fat- only image}+\text{water- only image}) \times 100$ “ $\text{rFF}\% = (F/(F + W)) \times 100$ ”.

- To facilitate image reading and reporting, we unify (compose) the different stations of each sequence into single-stack of images with WB coverage (Mobiview merge).
- A series of Maximum Intensity Projection (MIPs) reconstruction of the high b-value DW images, rotating around the cranio-caudal axis, allow an ‘at-a-glance’ overview for disease assessment and lesion detection. Typically, these are displayed in inverse grey scale, thus providing an appearance similar to that of PET.

Diffusion Weighted sequences. Diffusion Weighted Imaging evaluates the random motion of water molecules in tissues and the degree of its restriction. DWI sequences are acquired with different “b values”. The b value represents the strength or magnitude of the diffusion-weighting gradients applied during the imaging process. It quantifies how sensitive the DWI sequence is to the movement of water molecules within tissues. Higher b values indicate stronger diffusion weighting and are used to probe tissue characteristics like cellularity and microstructural changes. Consequently, a high signal in high b value sequences expresses a significantly restricted diffusion, that may be expression of high cellularity (such as in the spleen). Conversely, low b value sequences are more similar to T2 weighted sequences and are sensible to the water content of the tissue. Therefore, an area of high signal in low b value sequences that shows a signal drop in high b value sequences, likely represents a structure characterized by very low cellularity, such as the bladder or liver cysts.

Two b-values are needed for generating apparent diffusion coefficient (ADC) maps for lesion characterization and response assessment, the lowest no lower than 50 s/mm² in order to reduce perfusion-related signals and the highest between 800 and 1000 s/mm² in order to avoid kurtosis signal [24].

To simplify the interpretation of DWI images and quantitatively assess the degree of diffusion, ADC maps can be calculated [26,72]. Once these maps have been produced, ADC can be measured (in $\mu\text{m}^2/\text{s}$) with a 2D Region of interest (ROI) on lesions > 1 cm [23,60].

Free breathing Body Suppressed Diffusion Weighted Imaging (DWIBS) sequences in WB-MRI can efficiently identify focal and diffuse patterns of malignant disease in bone marrow; they are also able to detect lymph nodes (that can be then measured in morphological sequences), lesions within the viscera (such as in the liver), involvement of the peritoneal leaflets and of the retroperitoneum and, eventually, relatively large

areas of local recurrence [73]. Additionally, even though the measurement of ADC values is not easy to standardize between different scanners and they may be influenced by the chosen b values, however, general thresholds for the assessment of the bone marrow can be suggested. Specifically, MY-RADS guidelines state that:

- Normal bone marrow normally has ADC values below 600 - 700 $\mu\text{m}^2/\text{s}$
- Viable tumour lies between 700 - 1400 $\mu\text{m}^2/\text{s}$.
- Necrotic and treated disease normally has ADC values > 1400 $\mu\text{m}^2/\text{s}$.

In MY-RADS protocol, at least two b values DWI sequences (one with low b value ~ 50, one with high b value ~ 800-1000) are acquired, with the possibility of adding an intermediate b value ~ 600. Images are acquired in the axial plane, with 5 mm slices, without gaps, from the vertex to the knees. Coronal, high b value, inverted MIP images can be reconstructed, obtaining the so-called “PET-like” images, that can be used to grossly assess the disease burden, to preliminarily evaluate the areas of disease and for pictorial purposes.

T1-Dixon sequences. T1-Dixon sequences in a single acquisition, in-phase and out-of-phase images are acquired and can be used to obtain “fat-only”, “water-only”, “relative fat fraction” (FF) and “relative water fraction” maps.

Dixon imaging proves to be a valuable method for assessing bone lesions, as it quantifies the presence of microscopic intracellular fat content. This technique is well-established for evaluating bone marrow and has shown high levels of accuracy, with reported sensitivity and specificity reaching up to 95%, enabling the differentiation between neoplastic and non-neoplastic lesions [27,74]

T1-Dixon in phase images and Water Only images can be used as morphologic images; due to their high resolution, are useful to measure bone lesions. T1-Dixon images can be acquired on an axial plane with 5 mm contiguous slices, from vertex to knees. The relative Fat Fraction (FF) maps are calculated by (fat- only image)/(fat-only image + water-only image) x 100%. FF maps allow for a quantitative measure of fat within lesions, representing a further diagnostic aid when characterizing lesions based on fat content, especially within the bone marrow but also in other organs [75, 76, 77].

Other morphological sequences. Axial T2 5mm sequences (from vertex to mid thighs) and sagittal 4-5 mm T1 and T2/STIR sequences of the whole spine are acquired for the morphological assessment of the whole body.

T2 sequences can be used to characterize visceral lesions and to measure nodes and body structures as well as for identifying pleural effusion and ascites. Whole sagittal spine sequences allow a better understanding of vertebral fracture and spinal cord compression.

The information derived from the three groups of sequences should always be integrated and should not be considered by themselves, so to avoid misinterpretation of findings.

In general lesions that show:

- high signal intensity (SI) on high b value DWI sequences
- intermediate to low values (700-1000 $\mu\text{m}^2/\text{s}$) in ADC maps
- low FF values ($\leq 20\text{-}25\%$)
- low signal intensity in T1w imaging

are consistent with highly cellular, active, bone Focal Lesion FL (figure 2.1). These characteristics of signal intensity (SI) are present also in diffuse pattern of bone marrow infiltration [23, 49].

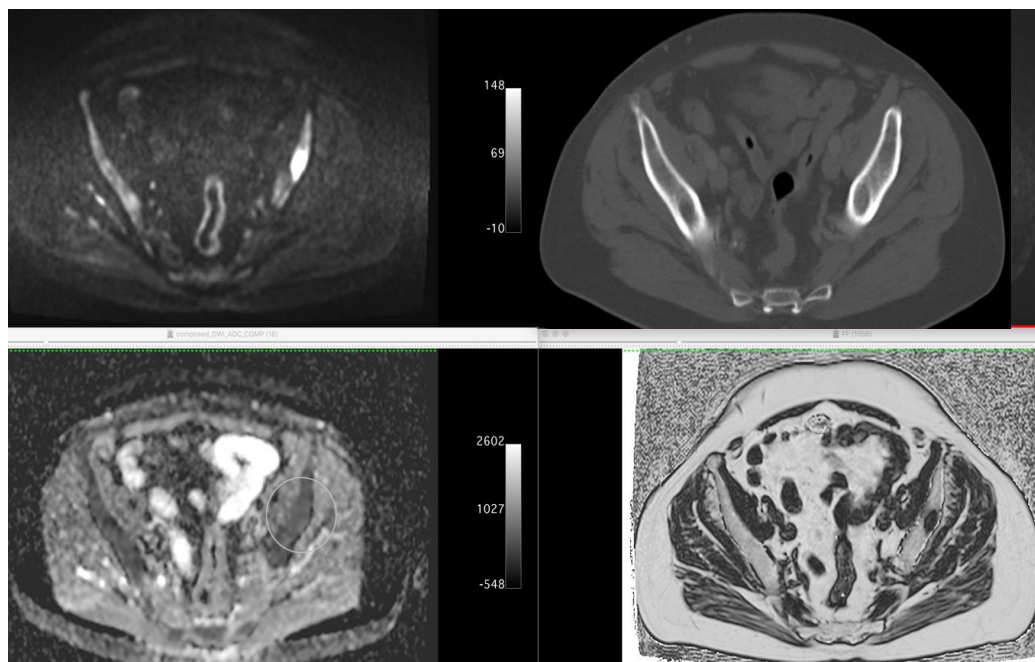


Fig. 2.1: left iliac bone lesion shows high DWI, low FF and low-intermediate ADC, in keeping with active disease. However, the lesion is poorly depicted on CT scan, due to its ill-defined ground glass appearance and a density similar to that of adjacent normal marrow. This image shows how CT scan is not reliable for identifying low density non osteolytic lesions.

The accurate examination of all sequences can lead to a proper differential diagnosis between benign and malignant lesions. For example, a benign haemangioma of the spine may sometimes appear as a high SI on DWI and intermediate SI on ADC, but the

fat content shown by FF images and the appearance in morphological sequences are reassuring. Conversely, cystic and/or degenerative lesions within the bone might appear as foci of extremely low FF, but the high ADC signal suggests a benign nature. Moreover, treated metastatic lesions may keep a high DWI signal, but the high ADC values are suggestive for response to therapy. This aspect is called “T2 shine-through” and refers to high signal on DWI images that is not due to restricted diffusion, but rather to high T2 signal which 'shines through' to the DWI image. T2 shine through occurs because of long T2 decay time in some normal tissue [77,78].

MY-RADS guidelines also introduced a score based on the Response Assessment Criteria (RAC) for assessing the likelihood of response to treatment. This 1 to 5 score is expressed after evaluating qualitative and quantitative characteristics of each lesion both in functional imaging and in morphological sequences (table 2.1). According to guidelines, in clinical reporting template a primary and a secondary RAC score could be expressed for each anatomical region and for paramedullary and extramedullary tissue for mixed response to treatment. However, in clinical practice, overall primary and secondary RAC scores can be expressed at the end of the report for simplicity.

A particular case where there can be diagnostic doubt between DWI and morphological sequences is the so called “T1-pseudoprogression” that refers to a worsening T1 appearance in WB-MRI follow-up of lesion visible only in DWI at baseline WB-MRI with increasing ADC values due to increased bone marrow oedema related to marked cell death. Increase in BM fat (“dot sign” within the lesion or “halo sign” in the peripheral part) is also a strong sign of response and may lower ADC values.

Figure 2.2 and table 2.2 show WB-MRI protocol based on MY-RADS.

Category	Description	Main characteristics	
RAC 1	Highly likely to be responding	Unequivocal decrease in number or size of FL Fat dot or halo sign in FL or return of normal SI	Increase of ADC from $<1400 \mu\text{m}^2/\text{sec}$ to >1400 or $>40\%$ For soft tissue disease, RECIST 1.1 for PR/CR
RAC 2	Likely to be responding	Slight decreases in number and size of FL but not to fulfil RAC1	Increase of ADC from $<1000 \mu\text{m}^2/\text{sec}$ to <1400 or 25-40% For soft tissue disease, RECIST 1.1 for PR
RAC 3	Stable disease	No change either for active or inactive disease	
RAC 4	Likely to be progressing	Equivocal appearance of new lesion(s) Reemerging of lesion	Increase SI on DWI high b-value and decrease of ADC <1400 For soft tissue disease, RECIST 1.1 not for PD
RAC 5	Highly likely to be progressing	Unequivocal new lesions >5 $>10\text{mm}$ New critical fractures or spinal cord compression Evolution from focal to diffuse pattern	New lesion with ADC between 600-1000 For soft tissue disease, RECIST 1.1 for PD

Tab. 2.1: MY-RADS Response Assessment Categories integrating MRI features parameters for a comprehensive assessment of treatment response.

Philips Ingenia 3T (bore diameter 70cm)	DWIBS	T1 (spine)	T2 STIR (spine)	T1 mDIXON -XD FFE (3D)	T2 Multivane + Trigger
Imaging plane	Axial	Sagittal	Sagittal	Axial	Axial
Field of view (cm)	450x400	200x350	200x350	430 x 430	430 x 430
Matrix size	128 x 112	224x346	224 x 346	268 x 217	288 x 288
Repetition time (ms)	4737 (shortest)	680-700	3400-5000	3.8 (shortest)	2014 (shortest)
Echo time (ms)	62 shortest)	10	60	1.31 (shortest)	126 (shortest)
Fast imaging / factor	EPI factor 47	TSE factor4	TSE factor17	none	TSE factor 50 / 1 perc. 240 %
Parallel Imaging factor	2.4	2	2	3	2.9
No. of signals averaged for high b-value images/NSA	1 (b=50), 6 (b=800)	1	1	1	1
Section thickness(mm)	5	4	4	3	5
Gap (mm)	1	0.4	0.4	0	0.5
Voxel (mm)	3.5x3.5x6	0.9x1x 4	0.9 x 1 x 4	1.6 x 1.97 x 3	1.5 x 1.5 x 5
Slices	30	18	18	83	48
Direction of motion probing gradients/ gradient mode	enhanced	default	default	maximum	maximum
Receiver bandwidth	3181.2 Hz	290.6 Hz	267.6 Hz	1504.4 Hz	438.4 Hz
Fat suppression	STIR (TI = 220 ms)	none	STIR (TI = 200 ms)	Dixon	none
b-values (s/mm ²)	50 and 800	0	0	0	0

Tab 2.2: Specific scan parameters of sequences.

Recently we added mDixon Quant sequence to the protocol in order to objectively identify bone marrow oedema and fat. This sequence yields fat-only and water-only images and proton density fat fraction (PDFF) parameter maps. The signal model used for this sequence includes a 7-peak fat model and an R2* correction factor that is common to both the water and fat signals. PDFF measurements have been shown to be accurate and precise in the liver and in the bone marrow. Their reproducibility across sites suggests potential for use in multicentre studies and for the development of thresholds which could be used to define and quantify disease [27,50,51,81].

ADC values stability and reproducibility were checked weekly through a dedicated phantom developed by the Italian association of Medical Physics, whereas FF values were checked with an in house built phantom following the guidelines developed by Bush et al [57, 81,82].



Fig 2.2 Example of images acquired according to WB-MRI MY-RADS protocols: A: T1, B: STIR; C and D: DWIBS b 50 and 800, E:ADC map; F,G,H and I: Dixon T1 in phase, opposite phase, fat only and water only; L: rFF map; M: T2w; N: MIP DWIBS b800

2.3 3T WHOLE BODY MRI ACCEPTANCY

As various whole body imaging modalities can be used for diagnosis and follow-up in different settings, understanding the patient experience may be beneficial in modern healthcare where it is crucial to recognize and meet patients' needs and preferences [83]. To guarantee patient-centered care and good quality scans, acceptance of WB-MRI is a pivotal topic and the studies available yielded somewhat divergent findings but none of them was performed exclusively on 3T MRI scanners. This study aims to evaluate cancer patients' experience of WB-MRI performed on a 3T scanner compared to other diagnostic total body examinations in particular PET-TC or Contrast Enhancement CT (CE-CT).

Materials and methods

From October 2020 to March 2022, all the patients who underwent a first-in-life WB-MRI at our institution for clinical reasons (cancer staging or follow-up) were asked to participate in this study for their feedback on the procedure. The questionnaire was divided into several parts: the first included demographic information and the examination date. The second part concerned the patient's experience: the physical and psychological reactions during the scan, such as dizziness, involuntary muscle contraction, tingling, tickles, increased temperature, sweating, fatigue, fear, headache, nausea, and tachycardia. All the data were evaluated using a four-point scale (0 not present; 1 present; 2 unpleasant, but tolerable; 3 intolerable). The third part evaluated the global level of satisfaction of WB-MRI considering exam duration, noise, the narrowness of the tube, and comfort of positioning (1 at ease; 2 low uneasiness; 3 moderate uneasiness; 4 strong uneasiness). Finally, in patients who had already performed other total body examinations, preference between WB-MRI and other diagnostic modalities was investigated. Statistical analyses were performed using the statistical and data management package MedCalc for Windows (Version 5.0.1.0 Ostend, Belgium). ANOVA or Pearson's Chi-squared test was performed to investigate the influence of age (ANOVA), sex (Chi-squared), and primary cancer (Chi-squared) on the acceptance of WB-MRI compared to other total body imaging modalities (CT or PET). Statistical significance was set at $P < 0.05$ for all tests.

Results

A total of 134 patients (73 males, 61 females) with a mean age of 61.3 ± 13.8 years were enrolled. Multiple myeloma (56/134, 41.8%) and prostate cancer (20/134, 14.9%) were the most common cancer types among patients. All patients had already performed other total body imaging techniques the previous year (CT and PET/CT); 83 patients (61.9%) had already performed MRI in different body districts with a similar scanner 1.5T or 3T, and only one patient had already undergone WB-MRI in another hospital on 1.5T scanner. In response to an open-ended question, none of these patients reported differences in symptoms or approval compared to previous MRIs. The most frequent symptom reported by the patients was a localized increase in temperature (51/134, 38.1%), which was deemed tolerable in all cases, whereas sweating was reported in only nine cases (9/134—6.7%) and always tolerable (figure 2.3).

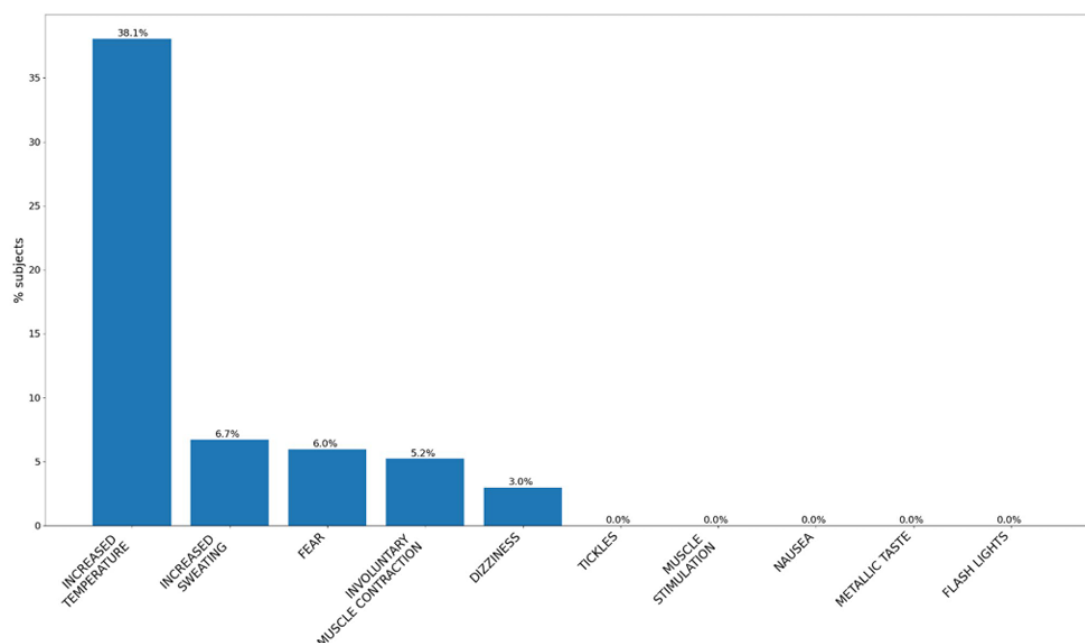


Fig. 2.3 Percentage of symptoms recorded during 3T WB-MRI exams.

Numbness and tingling of the limbs were both reported in 46 (46/134, 34.4%) cases and were deemed unpleasant or intolerable in 9 (9/46, 19.6%) and 3 (3/46, 6.5%) cases, respectively (Figure 2.3). Most patients were able to tolerate the length of the WB MRI as well as the noise and positioning. Only a few patients found the length of the examination, confined space, and positioning to be uncomfortable. 113 patients (113/134, 84.3%) perceived the WB-MRI duration as tolerable. Overall, WB-MRI was appreciated by the vast majority of patients (121/134—90.3%) who said they would

probably undergo the procedure again, whereas 13/134 (9.7%) said they would not repeat the examination in the future due to the narrowness of tube (3/13), lengthy examination time (8/13), and uncomfortable positioning (2/13). When asked to express a preference between the whole body imaging modalities, patients preferred the WB-MRI in 68.7% of cases (92/134), followed by Computed Tomography in 15.7% of cases (21/134) and Positron Emission Tomography in 7.4% (10/134), with a 8.2% (11/134) of patients who said they did not have any preference. The preference for imaging modalities was influenced by age ($p = 0.011$) with younger patients preferring WB-MRI, independently of sex and primary cancer ($p > 0.05$).

Discussion

Our study showed that most cancer patients of all types accept and highly appreciate WB-MRI compared to the other total body imaging modalities.

Low patient acceptability may lead to a lack of compliance during image acquisition, which can negatively affect the image quality [84] due to motion artifacts, especially if the examination requires a long acquisition time. Our study showed that most cancer patients of all types accept and highly appreciate WB-MRI compared to the other total body imaging modalities. The most common symptom reported by patients in our study was a localized increase in temperature, most likely due to the high overall Specific Energy Dose (SED) of WB-MRI, which implies an increase in body temperature, sometimes associated with sweating. Thus, maintaining a medium–high ventilation level inside the bore, allowing optimal heat dissipation, is crucial. It is thus critical to position the patient appropriately and comfortably, for example, with a knee cushion and armrest if available, to prevent patient movement. The examination's noise level was another issue raised frequently by patients. The loudness of several sequences—most notably DWI and T1W GRE mDixon—is quite high, and wearing both earplugs and earmuffs improve patient acceptability of the examination by reducing the perceived acoustic noise and enabling operator-patient communication, such as during breath-hold commands. Additionally, several patients suggested that listening to music could help reduce the discomfort induced by the loudness of the scan.

To our knowledge, only a few studies have investigated the patient experience of WB-MRI performed on 1.5 T scanner and compared it to other total body imaging modalities. Scan acceptability is greater in younger patients probably due to awareness

of the long-life expectancy and risks associated with the use of ionizing radiation also known in the general population.

In our experience, the majority (90.3%) of patients appreciated WB-MRI and said they would probably undergo the procedure again. These data are in line with the studies by Adams et al. [85] and Oliveri et al. [86] in cancer patients and with those of Busacchio et al. [87] in self-referring asymptomatic subjects screening. According to the study of Dyrberg et al. on the pleasure of WB-MRI in prostate cancer patients, WB-MRI had a satisfactory overall enjoyment level [88]. The Ryder A et al. study regarding patients affected by Myeloma demonstrated a high level of the overall satisfaction for both WB-MRI and other methods [89].

On the other side, the studies of Evans [90,91] in patients affected by colon rectal cancer and lung cancer founds that the overall satisfaction was lower for WB-MRI than PET/CT and, above all, compared to CE-CT.

Conclusions

Our study reveals that WB-MRI exams are well-accepted by adult patients suffering from various types of malignancies and bone disorders. This is a good starting point when thinking about how to apply the technology in clinical practice because WB-MRI seems to be just as pleasant as other total body imaging modalities (PET/CT, CE-CT). Evaluating a new imaging method through the eyes of the patients is an important factor in successfully developing and implementing WB-MRI in clinical practice. This analysis is essential to make our procedures more patient-friendly and offer a more patient-centered approach [92].

2.4 INTER READER AGREEMENT OF WB-MRI IN MYELOMA

Thanks to the collaboration between IRCCS Policlinico Sant'Orsola (Bologna) and IRST, in view of starting WB-MRI at Policlinico Sant'Orsola, in order to enhance the interpretation of WB-MRI images and to activate multicentric trial, we evaluate the inter-observer agreement of WB-MRI examination reports produced by readers of different expertise. The present study aims to evaluate the reproducibility of MY-RADS assessments among radiologists with different expertise in WB- MRI. Additionally, by evaluating the inter reader agreement of ADC and rFF values among radiologists, we seek to further investigate the methodology's reproducibility and reliability, which is crucial for early disease detection, patient stratification-risk and evaluation response to therapy, ultimately leading to improved survival outcomes.

Material And Methods

Study Population

All patients who have been referred to IRST from September 2020 to May 2023 for a WB-MRI have been retrospectively analysed. Only those who met the following inclusion criteria were enrolled: (1) Age > 18 years (2) New Diagnosis or suspected or confirmed relapse of Multiple Myeloma following IMVG criteria; (4) complete WB-MRI acquired along MY-RADS guidelines.

Exclusion criteria were: (1) Age < 18 years (2) Concomitant presence of a malignancy other than myeloma that might cause bone metastatic disease (3) Incomplete WB-MRI study or study affected by major artefacts.

Image Analysis

Three radiologists with different experience in WB-MRI were selected, including: 1 expert radiologist (AR) with more than 5 years in WB-MRI and 15 years' experience in MRI, 1 young radiologist (AC) with 3 years' experience in MRI and 6 months experience in WB-MRI, one senior resident doctor (EA) with 6 months experience in general MRI and 3 months experience in WB-MRI in myeloma patients, with MY-RADS teaching and a soft skill course on WB-MRI.

The three radiologists independently performed a clinical reporting template for each WB-MRI according to MY-RADS reporting system evaluating images on a Picture

Archiving and Communication System (PACS, AGFA Enterprise). Before evaluating WB-MRIs, all radiologists knew the patient's clinical and laboratory situation, which is essential for interpreting the investigation. Morphological images (fat- and fluid-sensitive sequences, low and high b-value DWI images) and quantitative ADC and FF maps are evaluated on PACS workstations by “linking” and browsing them in the different planes using multiplanar registration.

The analysis was conducted following MY-RADS guidelines in each 7 different anatomical regions (skull, cervical spine, thoracic spine, lumbo-sacral spine, pelvis, thorax and limbs including the femurs and the proximal humeri). Presence of paramedullary and extra-medullary disease was also assessed. Possible patterns of disease of bone marrow infiltration included: negative, focal, diffuse, focal on diffuse and micronodular.

Patients who received a follow-up WB-MRI scan were also evaluated for response assessment criteria (RAC) score following the indications expressed in MY-RADS [23]. MY-RADS guidelines suggest identifying 2 RAC for regions in order to document heterogeneous response (mixed/discordant responses): a primary RAC for the predominant pattern seen in most lesions and a secondary RAC for the other highest pattern seen in lesions within the region. Primary and secondary RAC scores from 1 to 5 were expressed for each anatomical segment as before.

All data were subsequently collected in a single database to assess the inter-reader agreement both for baseline evaluation and for the assessment of response to treatment.

Quantitative Analysis

To further evaluate reproducibility of ADC and rFF values among different radiologists, a quantitative analysis was conducted. For patients with follow-up, up to five lesions per patient, each with a minimum diameter of 1 cm [23,60], were contoured with a 2D circular Region of Interest (ROI) covering the central part of the lesion to assess the consistency of ADC and FF values among readers. All measurements were performed on PACS. To ensure that the same lesions were assessed by all readers, the first reader selected lesions anywhere in the acquisition volume, analyzed them, and marked them on PACS. The second and third readers then performed their analysis independently on each selected lesion. Data of size of ROI (cm²), ADC value (μm²/s) and fat content in rFF (%) were collected from the 3 readers. All data was subsequently collected in a single database to assess the inter-reader agreement both for ADC and FF values.

Study Design and Data Analysis

The imaging data were independently evaluated by the study's three raters, who classified the disease patterns. To quantify interrater agreement Cohen's Kappa (κ) and Brennan and Prediger's Coefficient (BPC) were calculated. Kappa, a widely used measure of agreement, adjusts for chance agreement but can be affected by imbalanced data distributions, where certain categories dominate the ratings. To address this limitation, BPC was also calculated as it is less sensitive to data imbalance and can complement kappa in such scenarios, including cases where the high-agreement but low-kappa paradox occurs. To detect cases where κ may not accurately reflect agreement beyond chance, the percentage of observed agreement was reported alongside these chance-corrected measures. As the Response Assessment Categories (RACs) are ordered (RAC 1 and 2 indicating response, RAC 3 stable disease, and RAC 4 and 5 disease progression), a weighted kappa was calculated, with linear weights applied to assign lower penalties to disagreements between adjacent categories compared to non-adjacent ones.

Both κ and BPC were estimated with 95% confidence intervals (CIs) calculated based on the variance due to the sampling process of subjects, assuming the set of raters was fixed. Agreement levels were interpreted using the following benchmark scale: Excellent agreement: 0.81-1.00; Substantial agreement: 0.61-0.80; Moderate agreement: 0.41-0.60; Fair agreement: 0.21-0.40; Slight agreement: 0.00-0.20; Poor agreement: <0.00.

To assess the reliability of mean ADC and FF values contoured by different radiologists, the Intra-Class Coefficient (ICC) was calculated. Reliability was classified based on previously established thresholds: poor reliability (< 0.5), moderate reliability (0.5–0.75), good reliability (0.75–0.90), and excellent reliability (> 0.90).

All statistical analyses were performed with STATA 15, MedCalc (v. 12.1.0 for Microsoft Windows 2000/XP/Vista/7; MedCalc Software) and IBM SPSS Statistics (IBM Corp.).

Results

Study population

Within the PACS of our institution, 70 patients were identified meeting the inclusion criteria 1 and 2; of these 15 had a WB-MRI negative for bone lesions and 2 had an

incomplete exam so they were excluded from the final analysis. Of these 53 patients 1 had more were excluded for presence of major artifacts due to obesity.

In the end 52 patients [23 Female, 29 Male; Mean age: 61 years (range 40-81 years); Isotype: IgA (12), IgG (32), Micromolecular (2), Non secretory (6)] were included in the study because of diagnosed with active multiple myeloma due to IMWG criteria (MM, 47) or presence of clinic lab data compatible with relapse-refractory MM (RR, 5). 25 (19 MM, 6 RR) patients were re-evaluated after therapy and included in the dedicated analysis. A total amount of 77 WB-MRIs were evaluated.

The 25 patients with paired WB-MRIs were further investigated for quantitative analysis MRI (qMRI) studies: 8 patients were excluded for presence of 2 micronodular pattern, 5 negative pattern, 1 focal lesion < 1 cm. 17 patients with focal or diffuse pattern of bone marrow infiltration met the criteria for inclusion in qMRI study and were finally evaluated. Across these 17 patients, 143 final lesions were quantified both in ADC and in rFF maps.

The analysis of inter observer agreement at staging

Agreement levels across anatomical regions were consistently high, with most regions demonstrating substantial to almost perfect agreement based on Cohen's Kappa, which ranged from 0.66 to 0.86. Brennan and Prediger's Coefficient (BPC) generally indicated higher agreement levels than kappa, with values ranging from 0.79 to 0.98, reflecting substantial agreement across most regions. The extramedullary region, however, displayed the lowest kappa value (0.23; 95% CI: -0.08–0.54), suggesting only slight agreement despite a high observed agreement of 96%. In contrast, BPC for this region was 0.92 (95% CI: 0.84–1), highlighting a case where kappa, sensitive to imbalanced data distributions, underestimated the true agreement. This discrepancy can be attributed to the fact that only two patients presented extramedullary disease, and one of the radiologists incorrectly identified extramedullary involvement in a patient (Table 2.3, Figure 2.4).

The analysis of the 2 radiologists with intermediate and initial experience in diagnosing focal lesions reveal that the radiologist with intermediate experience identified every patient with a focal pattern and never missed a diagnosis, whereas the reader with initial experience overlooked two focal on diffuse patterns that were interpreted as diffuse and the diffuse pattern with extramedullary disease.

These data indicate a good level of inter-reader agreement across most sections.

Region	% Agreement	Kappa	Kappa 95% CI	Brennan & Prediger	B&P 95% CI
Cranium	88%	0.66	0.49-0.82	0.85	0.75-0.94
Cervical	88%	0.82	0.71-0.93	0.86	0.77-0.94
Dorsal	88%	0.84	0.74-0.94	0.86	0.77-0.94
Lumbar	87%	0.81	0.70-0.92	0.83	0.73-0.93
Pelvis	90%	0.86	0.77-0.95	0.87	0.79-0.96
Thorax	88%	0.83	0.73-0.93	0.86	0.77-0.94
Limbs	83%	0.73	0.61-0.85	0.79	0.69-0.89
Overall Skeleton	93%	0.81	0.77-0.85	0.85	0.81-0.88
Paramedullary	94%	0.74	0.52-0.96	0.87	0.76-0.98
Extramedullary	96%	0.23	-0.08-0.54	0.92	0.84-1

Tab.2.3: different statistical evaluation of agreement in different anatomical regions

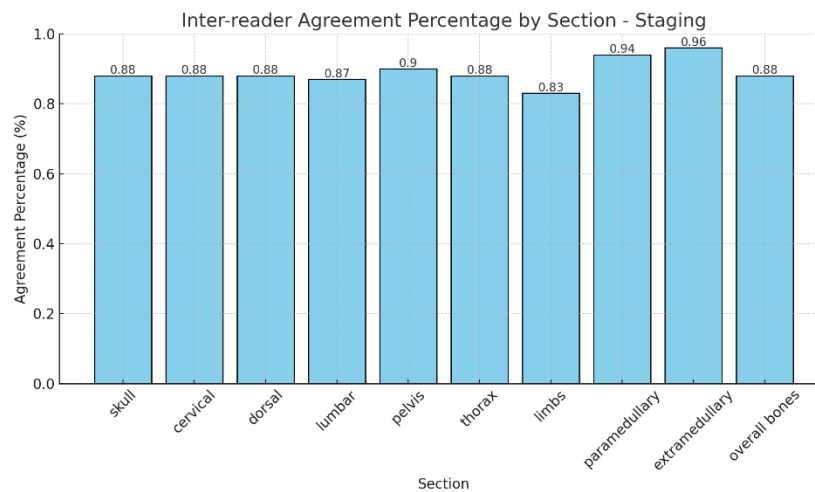


Fig. 2.4: Percentage agreement in the identification of different pattern of BMI across anatomical regions

The analysis of inter observer agreement at follow-up

Notably no mixed responses were observed by all readers in all regions, so we consider only one RAC for any anatomic segments. Cohen's Kappa values indicated substantial to excellent agreement, varying between 0.83 in the pelvis and 0.95 in the limbs. Similarly, Brennan and Prediger's agreement coefficients mirrored these results, with values ranging from 0.87 in the pelvis to 0.97 in the limbs. A summary of these data is shown in table 2.4 and figure 2.5.

Region	% Agreement	Kappa	Kappa 95% CI	Brennan & Prediger	B&P 95% CI
Cranium	95%	0.91	0.79-1	0.96	0.84-1
Cervical	92%	0.87	0.74-1	0.90	0.79-1
Dorsal	95%	0.92	0.80-1	0.93	0.84-1
Lumbar	92%	0.88	0.74-1	0.90	0.79-1
Pelvis	89%	0.83	0.67-0.99	0.87	0.74-1
Thorax	92%	0.87	0.72-1	0.90	0.79-1
Limbs	97%	0.95	0.86-1	0.97	0.90-1
Overall Skeleton	93%	0.89	0.84-0.94	0.91	0.88-0.95

Tab.2.4: different statistical evaluation of agreement in different anatomical regions between RAC

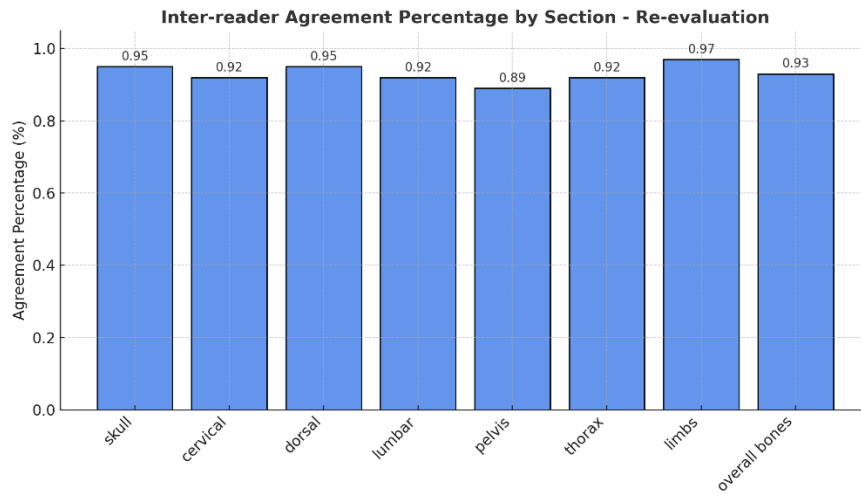


Fig. 2.5: Percentage agreement in the identification of RAC across anatomical regions

Cohen's weighted kappa further confirmed strong agreement, with values ranging from 0.82 in the pelvis to 0.96 in the arms; similarly, the Brennan and Prediger's weighted agreement coefficients were high, confirming strong inter-observer agreement across all regions (figure 2.6).

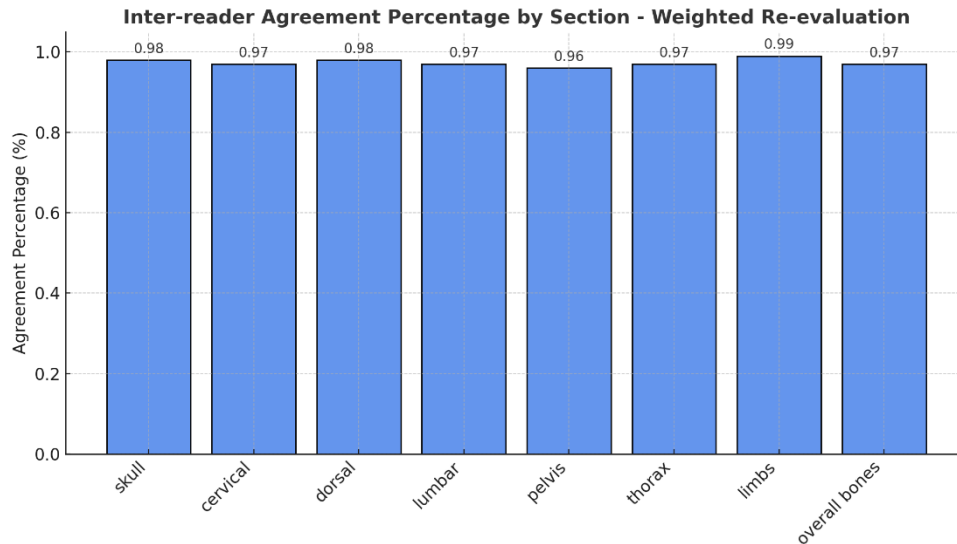


Fig. 2.6: Different percentage agreement for class of RAC

Quantitative analysis

We evaluated a total amount of 143 lesions respectively localized in: cervical spine (13), dorsal spine (20), lumbar spine (42), pelvis (28), limbs (17), thorax (21) and Skull (2). The only 2 lesions of the skull were excluded from the analysis for the statistical insignificance. The final analysis considers 141 lesions.

The results of ICC analysis on ADC and FF values and ROI area are graphically in figure 2.7A for all the anatomical region taken in consideration and with the detailed regions in fig.2.7B. We report excellent agreement among the three radiologists in both ROI area, ADC and FF values except for the limbs where we obtained a moderate agreement of 0.63 for ROI area, a good agreement for ADC values and an excellent agreement for FF values.

Considering the whole skeleton the ICC values for whole skeleton measurements indicate excellent inter-observer reliability across all metrics. The highest agreement was observed for Fat Fraction (FF) with an ICC of 0.98 (95% CI: 0.99-0.99), the Apparent Diffusion Coefficient (ADC) showed slightly lower but still strong agreement, with an ICC of 0.94 (95% CI: 0.95-0.97).

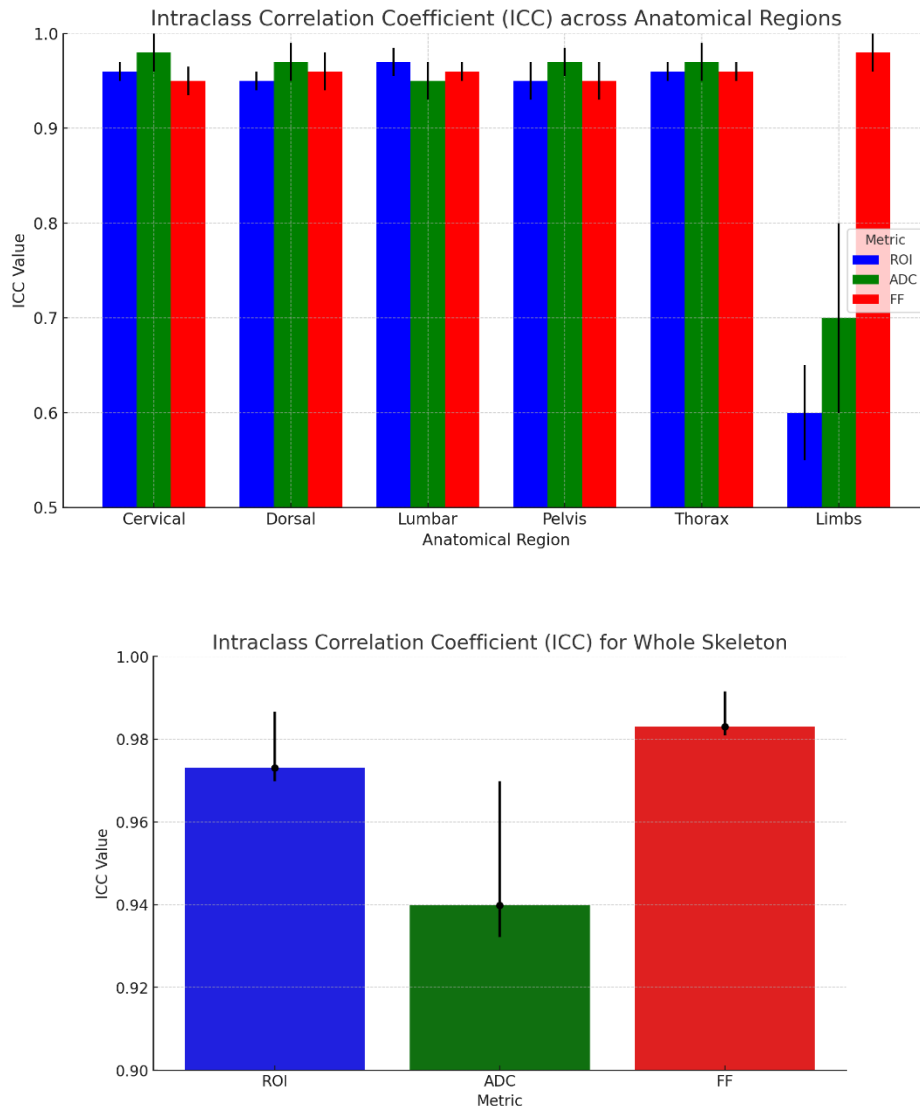


Fig 2.7: ICC for ADC, FF and ROI area for all anatomical regions (A) and for the whole skeleton (B)

Discussion

Our results indicate substantial to excellent agreement across all metrics, confirming the reproducibility of MY-RADS assessments among radiologists with varying experience levels. Importantly, high agreement was observed for focal lesions, which is critical for early detection and management of bone marrow infiltration in MM as well as for application of SLiM CRAB criteria [3].

At the time of writing, there is a limited number of studies comparing the reproducibility of RADS in bone disease and no one with quantitative assessment of ADC and rFF.

The inter-reader agreement observed in our study is consistent with findings by Croft et al. [93] who evaluated interobserver agreement between 3 expert radiologists (>8 years

of experience) in baseline WB-MRI in myeloma. They reported an overall skeleton ICC of 0.91, demonstrating excellent agreement in the assessment of myeloma-related bone disease but there ICC varied between skeletal regions with spine, pelvis and ribs showing good inter-observer agreement, whereas skull and long bones were moderate. In contrast to their study, we didn't observe variations in agreement across different anatomical regions, even if we observed that the senior resident missed some diagnosis particularly in the paramedullary and extramedullary areas highlighting the complexity of evaluating soft tissue involvement in MM and enhanced radiologist training.

Furthermore, our results align with those of Pricolo et al. [94], who studied inter-observer agreement using the MET-RADS-P system in metastatic prostate cancer between 3 radiologists of different level of expertise. Their findings of excellent agreement in bone assessments between radiologists supports a high reproducibility of RADS criteria. For the primary/dominant RAC pattern, the agreement between readers was excellent for the metastatic findings in cervical, dorsal, and lumbosacral spine, pelvis, limbs, lungs and other sites (K:0.81-1.0), substantial for thorax, retroperitoneal nodes, other nodes and liver (K:0.61-0.80), moderate for pelvic nodes (K:0.56), fair for primary soft tissue and not assessable for skull due to the absence of findings. For the secondary RAC pattern, agreement between readers was excellent for the metastatic findings in cervical spine (K:0.93) and retroperitoneal nodes (K:0.89), substantial for those in dorsal spine, pelvis, thorax, limbs and pelvic nodes (K:0.61-0.80), and moderate for lumbosacral spine (K:0.44). Pricolo and colleagues found lower agreement particular in node evaluation and in the assessment in secondary RAC, underscoring even in this study the importance of experience and continuous education in improving the accuracy of WB-MRI assessments particularly in these areas.

Lai et al. [95] highlighted the superior consistency of WB-MRI over whole body low dose CT for experienced observers; as a matter of fact interobserver agreement for WB-MRI was superior to WB-CT overall and for each region, without overlap in whole-skeleton confidence intervals (ICC 0.98 versus 0.77, 95%CI 0.96-0.99 versus 0.45-0.91). However, for inexperienced observers, although there is a trend for a better interobserver score for the whole skeleton on WB-MRI (ICC 0.95, 95%CI 0.72-0.98) than on WB-CT (ICC 0.72, 95%CI 0.34-0.88), the confidence intervals overlap

This study confirms the need of expertise and that the overall reliability of WB-MRI in assessing myeloma burden is high.

Moreover, Lecouvet et al [96] demonstrated that MRI has a better reproducibility rather than PET-CT; an intra- and inter-reader agreement very good for MRI (64 axial MRI and 20 WB-MRI) ($k = 0.90$ [0.81; 1.00] and 0.88 [0.78; 0.98]), whilst intra- and inter-reader agreement was good for PET/CT ($k = 0.80$ [0.69; 0.91] and 0.71 [0.56; 0.86]). MRIs were evaluated by two radiologists with respectively > 25 and >3 years of expertise.

In all articles, the expertise of the radiologists played a significant role. Experienced readers showed higher consistency compared to junior readers, particularly in complex regions. This suggests that the variability in regions like the limbs or skull and soft tissue could be improved with more experience as suggest by Berardo and colleagues that demonstrated that 80 WB-MRI reported in myeloma leads to a high level of inter-observer concordance [61].

Our study's second section examines readers' agreement while utilizing 2D ROIs to assess ADC and FF values, which are becoming more and more recognized as potential imaging biomarkers.

At present there are clear data supporting the use of proton density fat fraction (PDFF) for evaluation of liver steatosis that could represent a marker of metabolic associated fatty liver disease (MAFLD) and a consensus profile has recently been published [59]. PDFF thresholds for mild, moderate, and severe steatosis have been proposed at 5%, 15%, and 25%, respectively [97]. Furthermore, Schmeel et al [80] demonstrated that PDFF showed high linearity ($r^2 = 0.972$ - 0.978) and small mean bias (0.6-1.5%) with 95% limits of agreement within $\pm 3.4\%$ across field strengths, imaging platforms, and readers. Repeatability and reproducibility of PDFF were high, with the mean overall coefficient of variation being 0.86% and 2.77%, respectively. The overall intraclass correlation coefficient was 0.986 as a measure for an excellent interreader agreement. So they conclude that MRI-based quantification of vertebral bone marrow PDFF is highly accurate, repeatable, and reproducible among readers, field strengths, and MRI platforms, indicating its robustness as a quantitative imaging biomarker for multicentric studies. Moreover, ADC is emerging as possible biomarkers for clinical practise and the Diffusion-Weighted Imaging Biomarker Committee of the Quantitative Imaging Biomarkers Alliance (QIBA) has recently reviewed the ADC Profile and published a report to monitor lesions in the brain, liver, prostate, and breast. Specifically, changes in mean ADC exceeding 8% for brain lesions, 27% for liver lesions, 27% for prostate lesions, and 15% for breast lesions are claimed to represent true changes with 95%

confidence. These guidelines should enable successful clinical application of ADC as a quantitative imaging biomarker and guarantee repeatable ADC measurements that provide a reliable means of assessing treatment response [60].

Although bone lesions assessment is not included in these QIBA guidelines, we proceeded with their quantitative assessment in accordance with the MY-RADS and recently published studies about the use of ADC and rFF evaluation of bone lesions [30, 47, 49, 50, 51].

We observed an excellent interobserver agreement both in evaluation of ADC values and in rFF percentage across all anatomic regions. Our data are substantially in agreement with some articles published about the reproducibility of these values.

Castagnoli et al. [49] found excellent inter-reader agreement (ICC = 0.95) reported by on rFF evaluation of normal marrow (89.76%) and malignant bone lesions from breast (14.46%), myeloma (13.12%) and prostate cancer (13.67%) ($p > 0.017$). The excellent inter-reader agreement reported by aligns with our findings, supporting the reliability of FF measurements across different anatomical regions. Similar results on high reproducibility were found on ADC values by Michoux et al [48] that highlights the repeatability (RC) and reproducibility (RDC) limits for ADC measurements in an oncologic multicentre setting using WB-MRI. Based on the upper limit of the 95% confident interval on RC and RDC (from both readers), a change in ADC in an individual patient must be superior to 12% (cerebrum white matter), 16% (paraspinal muscle), 22% (renal cortex), 26% (central and peripheral zones of the prostate), 29% (renal medulla), 35% (liver), 45% (spleen), 50% (posterior iliac crest), 66% (L5 vertebra), 68% (femur), and 94% (acetabulum) to be significant. According to their work, for an ADC change to be considered clinically relevant, it must exceed specific thresholds, which vary depending on the organ but the coefficient of variation of ADC was not influenced by other factors (centre, reader). Small variations in ADC measurements, especially in regions such as bone marrow and complex tissues, may not necessarily reflect real pathological changes, but could be attributed to intrinsic reproducibility limits. Moreover, previous published data by Giles et al [98] showed the potential of WB- MRI imaging as a biomarker of treatment response in myeloma with good sensitivity. ADC analysis was found to be highly repeatable (mean coefficient of variation of ADC was 3.8% in healthy volunteers and 2.8% in myeloma patients) and also they found that mean ADC increased in 95% (19 of 20) of responding patients and decreased in all non-responders ($p = 0.002$). A 3.3% increase in ADC helped identify

response with 90% sensitivity and 100% specificity; an 8% increase (greater than repeatability of cohort 1b) resulted in 70% sensitivity and 100% specificity. There was a significant negative correlation between change in ADC and change in laboratory markers of response ($r = -0.614$; $p = 0.001$).

Another interesting test re-test study about reproducibility of functional parameters was performed on WB-MRI in advanced prostate cancer and it states that bone metastases' mean ADC and FF measurements of single lesions and global disease volumes are repeatable, supporting their potential role as quantitative biomarkers in metastatic bone disease [99]. Moreover, the excellent reproducibility of ROIs area supports the validity of reproducibility of ADC and FF. The good results of reproducibility of ADC and FF found in our study could enhance the Ingenia 3 T scanner's ability to perform multiparametric WB-MRI even though this type of scanner is not present in the OPTIMUM/MUKnine trial, a multicentric trial performed in England to improve WB-MRI in myeloma patients in different scanners [56, 57].

Possible limits of our study are the small sample size of patients especially in follow up group, the vast majority of bone lesions rather than soft tissue (only 2 patients with extra medullary disease and 9 with para medullary tissue) and the absence of mixed response (and so secondary RAC different from primary RAC) that is a tricky evaluation.

Conclusion

Our findings demonstrate that WB-MRI using the MY-RADS criteria provides a robust framework for the consistent evaluation of multiple myeloma. Their reproducibility supports its integration into clinical practice for the staging and management of patients. Furthermore, the reproducibility of quantitative ADC and FF measurements in our study supports their use as reliable biomarkers for disease assessment. The findings underscore the importance of protocol standardization and optimal training of radiologists to improve diagnostic consistency of multiparametric WB-MRI through structured training programs, with an emphasis on the MY-RADS criteria and “functional” imaging sequences like ADC and FF.

2.5 PROSPECTIVE ASSESSMENT OF 3T WHOLE-BODY MRI AND PET-CT IN DIAGNOSING MULTIPLE MYELOMA AND ITS INFLUENCE ON PATIENT CARE

The aim of this study is to compare the diagnostic accuracy of multiparametric WB-MRI (performed on a 3 Tesla scanner) with PET-CT in conjunction with clinical and laboratory data (International Myeloma Working Group criteria) in patients with myeloma and to evaluate the clinical relevance of the two imaging methods in patient management.

Materials and methods:

From October 2020 to January 2024, we prospectively enrolled consecutive myeloma patients who underwent to a 3 Tesla WB-MRI (performed according to MY-RADS guidelines); clinical and laboratory data were also collected for all the patients. Of these, we included in this part of the study only patients who underwent WB-MRI and FDG PET-CT within 1 month at staging or follow-up. Independently and without knowledge of the outcome of the other imaging modality, one radiologist and one nuclear medicine physician assessed each patient's level of bone involvement by analysing various patterns of bone marrow infiltration (BMI) for focal lesions (FL) as well as diffuse/micronodular infiltration (per-patient and per-pattern of BMI analysis) in each imaging modality.

Diagnosis and relapse were defined as per International Myeloma Working Group (IMWG) criteria using laboratory parameters including bone trephine biopsy and assessment of end-organ damage. Baseline clinical data of presentation at trial entry (new diagnosis or suspected of confirmed relapse), laboratory myeloma markers (in particular: haemoglobin, calcium, glomerular filtrate rate (GFR), serum paraprotein; bone marrow plasma cells percentage at trephine biopsies) were collected into an anonymized database. GFR was obtained as the well-established clinical practise through CKD-EPI equations taking into account age, sex, race and creatine level [100]. Bone marrow biopsies were performed at the right posterior iliac crest without image guidance for all the patients, as for clinical practise.

Two expert specialist haematologists, during a multidisciplinary weekly meeting, reviewed all clinical, biological, histopathological and imaging data available and, in

consensus, reached an agreement on management plans [3,5,12,17]. This was used as a reference to evaluate WB-MRI and PET-CT diagnostic performance and impact on management (either treat as myeloma with or without additional radiotherapy or surgery, or treat for non-myeloma related event or active surveillance). The multidisciplinary meeting is composed by at least a panel of experts of 1 radiologist, 1 nuclear medicine physician, 1 radiotherapist, 1 pathologist and 2 haematologists. Statical tests used for the analysis were: Mc-Nemar Test to compare sensitivity and specificity of WB-MRI and PET-CT and to assess which imaging technique has the strongest impact on the change of management, T-Test or Wilcoxon rank-sum test to assess the correlation between lab data and imaging findings.

Results:

Study population

This study evaluated 137 patients, with 54% male (n=73) and 46% female (n=61), median age 66 years (IQR: 57-74). Table 2.5 summarized patients' characteristics.

Overall (N=137)	
Sex – n (%)	
male	73 (54%)
female	64 (46%)
Age at MRI	
median, IQ-IIIQ	66, 57-74
ISS – n (%)	
I	39 (48%)
II	23 (28%)
III	23 (24%)
ND	55
Disease stage – n (%)	
smouldering	39 (28%)
NDMM	37 (27%)
post-ASCT	37 (27%)
RDMM	24 (18%)
Distance MRI-PET	
median, IQ-IIIQ	7, 2-20

Tab 2.5: patients' characteristics of this study.

The stages of the disease were distributed as follows:

- Group 1: High Risk Smouldering Multiple Myeloma: 28% (n=39)
- Group 2: Multiple Myeloma 72% (n=98): Newly Diagnosed Multiple Myeloma (ND-MM): 27% (n=37); Post-Autologous Stem Cell Transplant (Post-ASCT): 27% (n=37); Relapsed/Refractory Multiple Myeloma (RRMM): 18% (n=24)

The median time between WB-MRI and PET-CT scans was 7 days.

Comparison of WB-MRI and PET Findings

WB-MRI Imaging Patterns were: Negative: 60 patients (44%); Diffuse: 12 patients (9%); Micronodular: 3 patients (2%); Focal: 41 patients (30%); Focal on Diffuse: 21 patients (15%); PET-CT Imaging Patterns were: Negative: 73 patients (53%); Diffuse: 6 patients (4%); Micronodular: 0 patients (0%); Focal: 38 patients (28%); Focal on Diffuse: 20 patients (15%) (Table 2.6).

		PET-CT				Total
		Negative	Diffuse	Focal	Focal on diffuse	
WB-MRI	Negative	58	1	0	1	60
	Diffuse	8	4	0	0	12
	Micronodular	3	0	0	0	3
	Focal	3	1	34	3	41
	Focal on diffuse	1	0	4	16	21
Total		73	6	38	20	137

Tab 2.6: Overall WB-MRI and PET-CT patterns of bone marrow involvement.

In the HR-SMM cohort (group 1), WB-MRI showed presence of BMI in 10/39 patients (26%) while PET-CT in 4/39 patients (10%), determining a sensitivity for WB-MRI and PET-CT of 91% and 36% respectively ($p=0.07$), whereas the specificity was of 100% for both ($p=1$; Table 2.7). Additionally, in the same cohort, patients with BMI, demonstrated in at least one imaging modality, demonstrated higher blood paraprotein levels compared to patients without BMI in both WB-MRI and PET-CT ($p=0.01$; table 2.9).

	Reference				Reference			
WB-MRI	Negative	Positive	Total	PET-CT	Negative	Positive	Total	
Negative	28	1	29	Negative	28	7	35	
Positive	0	10	10	Positive	0	4	4	
Total	28	11	39	Total	28	11	39	
Sensitivity %	91 (59-100)			Sensitivity %	36 (11-69)			<i>p=0.07</i>
Specificity %	100 (88-100)			Specificity %	100 (88-100)			<i>p=1</i>

Tab 2.7: Sensibility and specificity of WB-MRI and PET-CT in HR-SMM group.

In the MM cohort (group 2), WB-MRI demonstrated BMI in 67/98 patients (68%) while PET-CT in 60/98 (61 %) cases, thus leading to a sensitivity of 100% and 89% respectively for WB-MRI and PET-CT ($p=0.02$). and a specificity of 97% for both. These results agree with the literature data about the ability of WB-MRI to depict diffuse and micronodular pattern of BMI (15 cases vs 6 cases). Even in this cohort of patients WB-MRI and PET-CT shared the same specificity (97%; $p=1$. Table 2.8). In the MM cohort, there is also a correlation between positive imaging findings, obtained in at least one modality between WB-MRI and PET-CT, and blood paraprotein level ($p=0.007$) and haemoglobin ($p=0.002$). Specific Laboratory Data Analysis are show in Tables 2.9 (HR-SMM) and 2.10 (MM).

	Reference				Reference			
WB-MRI	Negative	Positive	Total	PET-CT	Negative	Positive	Total	
Negative	31	0	31	Negative	31	7	38	
Positive	1	66	67	Positive	1	59	60	
Total	32	66	98	Total	32	66	98	
Sensitivity %	100 (95-100)			Sensitivity %	89 (79-96)			<i>p=0.02</i>
Specificity %	97 (84-100)			Specificity %	97 (84-100)			<i>p=1</i>

Tab 2.8. Sensibility and specificity of WB-MRI and PET-CT in MM group

	Negative WBMRI & PET	Positive WBMRI OR PET	<i>p</i>
Plasma cells % (PC%)			
median, IQ-IIIQ	20, 13-40	33.5, 20-45	
mean (SD)	27.0 (16.9)	33.8 (13.6)	0.24
Monoclonal para protein (g/L)			
median, IQ-IIIQ	19.4, 14.7-28.2	26.5, 23.2-35.4	
mean (SD)	19.8 (10.0)	29.8 (10.0)	0.01
haemoglobin (g/dl)			
median, IQ-IIIQ	12.6, 11.6-13.5	11.7, 10.9-12.3	
mean (SD)	12.4 (1.6)	11.7 (1.3)	0.11
Glomerular filtrate rate			
median, IQ-IIIQ	85.5, 67-102	93, 82-103	
mean (SD)	82.4 (20.8)	90.5 (14.8)	0.37
calcium (mg/dL)			
median, IQ-IIIQ	9.5, 9.2-9.8	9.3, 9.2-9.8	
mean (SD)	9.5 (0.4)	9.4 (0.5)	0.51

Tab. 2.9: Laboratory Data Analysis (High Risk -Smoldering Multiple Myeloma)

	Negative WBMRI & PET	Positive WBMRI OR PET	<i>p</i>
Plasma Cells % (PC%)			
median, IQ-IIIQ	20, 13-75	51, 20-80	
mean (SD)	38.5 (35.0)	49.7 (31.8)	0.14
Monoclonal para protein (g/L)			
median, IQ-IIIQ	3.25, 0.6-14.4	15.9, 2.7-32.4	
mean (SD)	7.4 (8.0)	19.3 (18.8)	0.007
haemoglobin (g/dL)			
median, IQ-IIIQ	12.5, 11.8-13.2	10.9, 9.5-13	
mean (SD)	12.5 (1.4)	11.2 (2.0)	0.002
Glomerular filtrate rate			
median, IQ-IIIQ	79, 63-95	78, 51-90	
mean (SD)	77.9 (20.6)	71.2 (25.9)	0.22
calcium (mg/dL)			
median, IQ-IIIQ	9.2, 8.8-9.7	9.4, 8.9-9.9	
mean (SD)	9.3 (0.6)	9.4 (0.7)	0.29

Tab.2.10: Laboratory Data Analysis (Multiple Myeloma)

Analysis of the ability to identify focal pattern reveals:

- Group 1 (HR-SMM): WB- MRI detected focal lesions in 10% (4/39), while PET-CT detected focal lesions in 8% (3/39). The percent agreement was 92% (95% CI: 84-100%), with a Cohen's kappa of 0.53 (95% CI: -0.04-1.00).
- Group 2 (MM): WB- MRI detected focal lesions in 59% (58/98), and PET-CT detected focal lesions in 56% (55/98). The percentage agreement was 97% (95% CI: 93-100%), with a Cohen's kappa of 0.94 (95% CI: 0.87-1.00). Sensitivity of 100 % and specificity of 98 % for WB-MRI and 97% and 100% respectively for PET-CT (p=0.5 for sensitivity).

Impact on management of patients

Overall change of management was 74/137 (54%), no change 63/137 (46%).

If we consider positive findings, there are more positive WB-MRI (66/74, 89%) than PET-CT (56/74, 76%) in patients whose management changes, therefore WB-MRI positive findings have a greater impact on management change with $p = 0.002$ (Table 2.11 and 2.12).

MRI	Change of management			PET	Change of management		
	No change	Change	Total		No change	Change	Total
Negative	52	8	60	Negative	55	18	73
Positive	11	66	77	Positive	8	56	64
Total	63	74	137	Total	63	74	137
$p < 0.001$				$p < 0.001$			

Tab 2.11: positive and negative WB-MRI and PET-CT.

MRI	PET		
	Positive	Negative	Total
Positive	56	10	66
Negative	0	8	8
Total	56	18	74
$p = 0.002$			

Tab 2.12: positive WB-MRI and PET-CT.

Both imaging techniques are associated with management change. This change was consistent with WB-MRI in 72/74 (97%) compared to PET-CT where we found consistency in 45/74 (61%) with a significant difference (p -value < 0.001) and so a stronger impact on management change of WB-MRI (Table 2.13).

	PET-CT		
WB-MRI	<i>change</i>	<i>no change</i>	Total
<i>change</i>	43	29	72
<i>no change</i>	2	0	2
Total	45	29	74
$p < 0.001$			

Tab. 2.13: Overall change of management.

Furthermore, WB-MRI led to a diagnosis of active multiple myeloma in 4 patients previously diagnosed HR-SMM thank to findings of M criteria of SLiM CRAB criteria. WB-MRI led to a change of management, also, for incidental findings in 7 patients for non-myeloma-related event. In 4 patients it detected other malignancies (1 right colon cancer, 2 pancreatic adenocarcinomas and 1 cholangiocarcinoma) that required further oncological and surgical intervention and 3 spinal cord compression requiring immediately neurosurgeon evaluation and radiotherapy.

Discussion

WB-MRI and PET-CT provide different perspectives on myeloma patients helping in the identification of BMI before osteolytic end organ damage occurs. Utilization of functional imaging modalities is increasing since their superiority in sensibility and specificity rather than WB LD CT and skeleton X-Ray and they are insert into international guidelines. Because PET-CT discriminate between diseases that are metabolically active and those that are not, it is useful in identifying lesions with high metabolic activity, such as focal lesions (FLs) and extramedullary disease (EMD) and it is particularly effective in evaluation after treatment. On the other hand, WB-MRI examines the water and fat content of tissues and following MY-RADS guidelines, is particularly sensitive in identifying early bone marrow infiltration both in case of focal lesions and in case of diffuse or micronodular pattern of BMI. The inclusion of DWI

sequences with ADC and FF maps leading to a multiparametric WB-MRI and has increased its diagnostic power, allowing for the detection of subtle changes in bone marrow cellularity and fat content, which are indicative of disease burden also in the assessment after treatment even in RR-MM patients, owing to the proven excellent capability of WB-MRI to depict bone marrow pathology, particularly in the context of the extremely heterogeneous nature of myeloma and its spatial variability [42, 79; 101,102]. The selection of imaging modalities to use now mostly depends on patient's needs, local availability and national recommendation, but also on costs [15].

Comparative studies have suggested that multiparametric WB-MRI is more sensitive than PET-CT for the detection of BMI at diagnosis of MM thanks to its high spatial resolution and sensitivity to diffuse pattern [10, 29]

At the time of writing, there is a limited number of studies prospectively comparing diagnostic performance of WB-MRI and PET-CT in myeloma patients and evaluating management change.

Our study first compared the diagnostic accuracy of the two imaging methods for the detection of BMI and it was high for both. An overall sensitivity of 100% and 89% respectively for WB-MRI and PET-CT ($p=0.02$) and a specificity of 97% for both was found in MM group. Moreover, WB-MRI detected in MM group diffuse pattern alone or combining with focal pattern in 27 patients, while PET-CT in 23 cases, in HR -SMM patients diffuse pattern was found in 6 WB-MRI and in 2 PET-CT; WB-MRI depicted also 2 micronodular patterns. Even if they are SLiM CRAB criteria considers only FL, we consider diffuse pattern and micronodular pattern sign of BMI in patients affected by HR-SMM and factor to access to some clinical trials for treatment. In this setting, the value of WB-MRI was extremely high showing a sensitivity of 91% and specificity of 100%, whilst for PET-CT we found a sensitivity of 36% and specificity of 100%.

At present only another prospective trial has been published by Messiou C et al. about prospective comparison of multiparametric WB-MRI and FGD PET-CT and our result agrees with the results of iTIMM trial [39]. The iTIMM trial highlights the superior sensitivity of WB-MRI over PET-CT in detecting multiple myeloma: WB-MRI identified focal lesions in 83% of participants (50 out of 60), while PET-CT detected them in only 60% (36 out of 60); WB MRI detected diffuse disease in 82% of participants (49 out of 60), compared to just 17% (10 out of 60) for PET-CT. WB-MRI showed significantly higher detection for focal lesions at all anatomic sites (except ribs, scapulae, and clavicles) and for diffuse pattern at all sites; moreover, participants with

diffuse disease at MRI had higher plasma cell infiltration ($p = 0.03$) and paraprotein levels ($p = 0.02$) compared with those without diffuse disease. All genetically high-risk tumors showed diffuse infiltration at WB-MRI. In concordance to our study, there was a correlation between paraprotein level and positive findings on WB-MRI or PET-CT both in HR-SMM and MM patients, however there wasn't a correlation with PC but we found a correlation with haemoglobin and MM positive imaging findings. The inferior diagnostic performance of PET-CT in this study compared to our results could be explained using IMPeTUs criteria that enhance the diagnostic power of PET-CT.

Westerland et al retrospectively [103] compared the two modalities and the impact on clinical management and found that WB-MRI had a higher per patient sensitivity for bone disease because it detected a higher number of lesions per patient with concordance between PET-CT and WB-MRI in terms of disease positivity and lesion number in only 59% of patients.

Moreover, Lecouvet et al [96] demonstrated that MRI is significantly more sensitive and reproducible than PET-CT to detect BMI in MM. The sensitivity of MRI to detect BMI (97% [90%; 100%]) was significantly superior to that of PET-CT (76% [64%; 85%]) ($p < 0.001$). The specificity of MRI (86% [57%; 98%]) was lower than that of PET-CT (93% [66%; 100%]), without reaching statistical significance ($p = 0.32$). A significantly higher plasma cells percentage was observed in positive MRI or PET-CT findings.

However, PET-CT showed high performance, including for evaluation of diffuse infiltration, in another prospective study on 30 new diagnosed MM [104]; although WB-DWI detected more FLs than did PET-CT, there was no difference in the detection of bone disease on a per-patient basis.

Regarding the impact on change in management, imaging with PET-CT and WB-MRI resulted in 54% (74/137) of patients with WB-MRI consistent with the change in 97% (72/74) and PET-CT in 61% (45/74). Moreover, positive WB-MRI has a greater impact on management change. Westerland et al [103] found a change in management thanks to the use of functional imaging in up to 23.9% (11/46) of patients, compared to review of clinical data alone. WB-MRI resulted in a decision to treat in an additional 6.5% (3/46) of patients compared with PET-CT. According to our data, also in this study negative PET-CT scan had a slight negative impact on management (i.e. surveillance versus treatment), mitigated by positive clinical data. However, treatment decisions

were not statistically different and either modality would be appropriate in initial staging, depending on local availability and expertise.

In contrast to these data, in the research of Lecouvet there was a strong correlation between decisions regarding patient management and PET-CT findings ($p < 0.001$). However, as underlined by authors, a limit of that study could be that there is a mix of 64 axial MRI and only 20 WB-MRI [96].

Our study has some limitations: first, we did not focalize on number of focal lesions, but we assessed a per pattern evaluation; second, we did not evaluate the performance of WB-MRI and PET-CT in the post treatment assessment and minimal residual disease; third, there is a lack of prognostic value.

Conclusion

Our prospective trial supports the utmost role of WB-MRI and FGD PET-CT in the assessment of patients affected by SMM and MM at both diagnosis and relapses with a superior sensibility of WB-MRI in the evaluation of BMI and higher impact on therapeutic approach. While PET-CT provides a quantitative measure through SUV, WB-MRI's ADC and FF measurements offer a non-ionizing alternative for assessing disease burden without radiation exposure. The future of myeloma management will likely see greater integration of both techniques, with PET-CT providing metabolic assessments and WB-MRI offering detailed structural evaluations. This combined approach could improve diagnostic accuracy and treatment monitoring, enhancing patient outcomes.

2.6 RADIOPSY: QUANTITATIVE MULTIPARAMETRIC WB-MRI FOR DISCRIMINATION OF SMOLDERING AND MULTIPLE MYELOMA

Multiple Myeloma (MM) evolves from precursor stages such as smoldering multiple myeloma (SMM), with distinct risks of progression. Whole-body MRI (WB-MRI) is a sensitive imaging tool for assessing bone marrow infiltration (BMI) in MM and SMM. This study aims to develop predictive models using WB-MRI quantitative imaging biomarkers (QIBs) extracted from Apparent Diffusion Coefficient (ADC) and relative Fat Fraction (rFF) maps, to distinguish between SMM and MM.

Materials and Methods

Population demographics clinical and biological data

Inclusion criteria of this part of the study were execution of 3 T WB-MRI (performed according to MY-RADS guidelines), suspected new diagnosis of HR-SMM or MM defined as per IMWG criteria using laboratory parameters including bone trephine biopsy and assessment of end-organ damage. Exclusion criteria were patients with MRI-unsafe prostheses and devices patients with other known malignancies, patients whose tests are of suboptimal quality, or whose test has been suspended, or is incomplete.

Baseline clinical data of presentation at trial entry (new diagnosis or relapse), laboratory markers (b2-microglobulin, albumin, haemoglobin, calcium, lactate dehydrogenase, glomerular filtrate rate, serum free light chains and ratio, serum paraprotein, bone marrow cellularity and infiltration at trephine biopsies (plasma cells percentage, BMPC%) were collated into an anonymized database.

Bone marrow biopsies were performed at the right posterior iliac crest without image guidance for all the patients, as for clinical practise. Glomerular Filtration Rate (GFR) was obtained as the well-established clinical practise through CKD-EPI Equations taking into account age, sex, race and creatine level [100].

Patients' characteristics are summarized in Table 2.14. Mann Whitney U was employed to evaluate differences in demographical, clinical and biological indicators.

Variable	n (%) or Median (Range)
Sex	
Male	46 (55%)
Female	38 (45%)
Monoclonal Protein Isotype	
IgG	50 (60%)
IgA	20 (24%)
Light Chain	10 (12%)
Non Secretory	4 (4%)
ISS Stage	
Stage I	45 (53%)
Stage II	24 (28%)
Stage III	15 (18%)
Disease Stage	
Smouldering	39 (46%)
MM	45 (54%)
Bence Jones Protein	
Negative	28 (33%)
Positive - Not Evaluable	22 (26%)
Median (g)	0.105 (0.0 – 1.86)
Serum M-protein (g/L)	19.6 (0.0 - 85.5)
Bone marrow plasma cells (BMPC %)	39.0 (6.0 - 100.0)
Hemoglobin (g/dL)	11.9 (7.0 - 16.5)
Platelet Count (10⁹/L)	214.5 (58.0 - 692.0)
Calcium (mg/dL)	9.4 (8.3 - 11.1)
Albumin (g/L)	41.0 (27.0 - 49.0)
LDH (U/L)	167.5 (87.0 - 352.0)
Beta2 microglobulin (mg/L)	2.60 (1.83 - 3.85)
Glomerular Filtration Rate (ml/min)	80.00 (64.50 - 95.00)
Alkaline Phosphatase (U/L)	66.5 (32.0 - 446.0)

Tab. 2.14. Clinical Characteristics of the Study Population (n=84)

Image segmentation and Analysis

An experienced radiologist and a medical physicist placed four identical cylindrical volumes of interest (VOIs), labelled Left inferior (LI), left superior (LS), right inferior (RI), and right superior (RS), each with a volume of 2.6 cm³, on the pelvic bone and two

more VOIs were positioned on the dorsal and lumbar vertebrae (the vertebrae numbers may vary due to disease invasion), both on ADC and rFF sequences as shown in figure 2.8. Specifically, the RS and RI VOIs were placed near the biopsy site, where plasma cell percentage (BMPC %) is assessed. Particular attention was given to avoiding any focal lesions during VOI placement, especially in patients with MM, as well as to potential focal spots near the biopsy site.

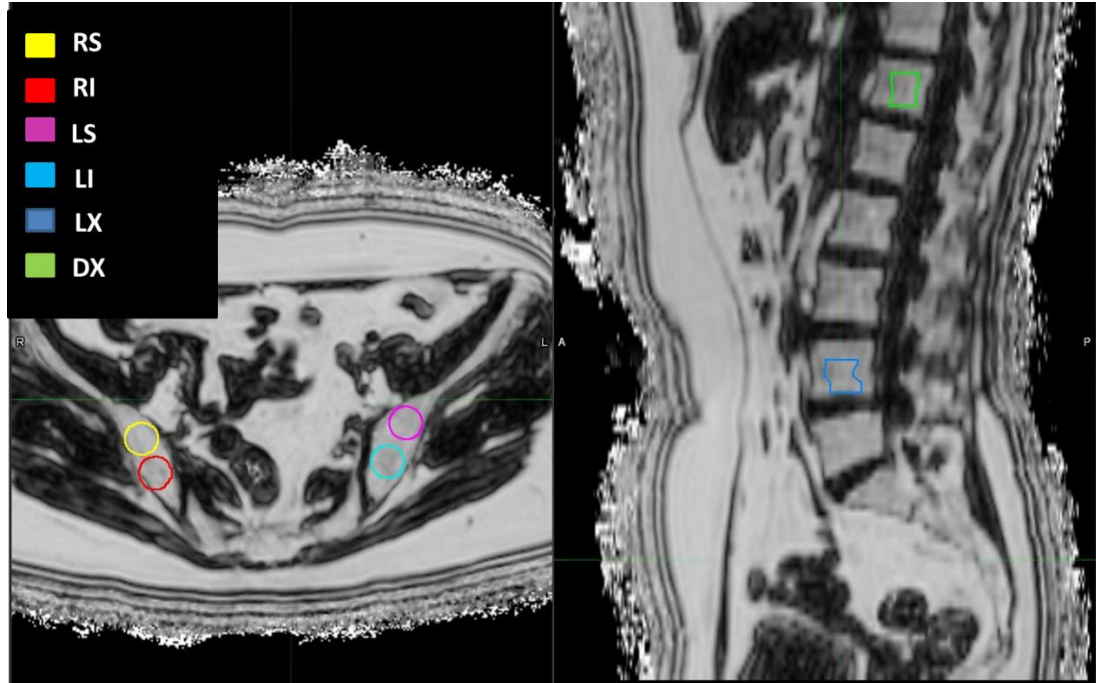


Fig. 2.8: Example of patient contouring with cylindrical VOIs at different sites of pelvic bone and vertebrae labelled as right superior (RS), right inferior (RI), left superior (LS), left inferior (LI), lumbar vertebra (LX) dorsal vertebra (DX).

Statistical Analyses

Mean ADC and rFF values extracted from the RS VOI, which are most representative of disease invasion due to their proximity to the biopsy site, were correlated using Pearson's ρ with relevant laboratory biomarkers (statistically different among SM and MM at Mann Whitney test) and previously reported in literature as predictors for SM and MM [7;105].

Histogram, second and higher order QIBs were extracted from all VOIs using S-IBEX, an IBSI-compliant software [106]. QIBs with an Intra-Class Correlation Coefficient (ICC) < 0.5 , evaluated for RS and RI, were discarded. The Pearson correlation coefficient (ρ) was employed to eliminate redundancy among QIBs employing a threshold of 0.85.

The Least Absolute Shrinkage and Selection Operator (LASSO) algorithm was used to select the most predictive QIBs for both ADC and rFF sequences across all VOIs. The QIBs most frequently selected by LASSO across the different VOIs were further analysed and employed to build predictive models.

Single and Multi-feature models, internally validated using 3-fold cross-validation, were developed with two-thirds of the dataset and tested on the remaining one-third. This procedure was repeated 30 times to assess the median performance of the models on the various VOIs, including RS, RI (the VOIs near the biopsy site), LS, LI, L5, and D11. De Long Test will be employed to score the significance of model's performance on different VOIs.

In the end "radiopsy" model based on RS VOI will be further analysed through precision and recall calculation and compared with prediction performance of Plasma cell invasion percentage and with mean values of rFF and ADC sequences. Calibration and decision curves will be used for model performance and benefit comparison. The best logistic regression model developed from RS VOI will be tested on the distant VOIs RI, LS, LI, L5 and D11 to demonstrate that "radiopsy" can be used to evaluate clinical status from sites distant from biopsy.

Radiomics score:

We evaluated the quality of radiomics research employing METRICS (Excellent 88%) and Radiomics Quality Score [63,65].

Results

Patients' characteristics and histopathology

102 patients were included in this study. 18 patients have been discarded for bad image quality or presence of artifact in pelvis or incomplete exam, resulting in 84 patients included for the quantitative analysis (46 male, 38 female) as shown in Figure 2. 45 patients were affected by Multiple Myeloma and 39 affected by Smouldering Multiple Myeloma.

Patients' demographic, clinical and biological characteristics and their evaluation with Mann Whitney U are shown in Table 2.15 where BMPC%, HB, serum albumin and Beta2 microglobulin are significantly different with those in the SMM group.

	SM mean	SM std	MM mean	MM std	p-value
Age	58.91	10.38	64.93	9.47	0.025
Sex(M/F)	18/20		28/20		
M-Protein (g/L)	19.45	10.10	28.19	23.73	0.211
BMPC %	25.30	15.45	63.83	25.97	0.000
HB (g/dL)	12.44	1.71	10.69	2.08	0.001
PLT (10 ⁹ /L)	231.87	60.60	205.17	79.34	0.202
Serum calcium(mg/dL)	9.48	0.35	9.53	0.74	0.932
Serum Albumin (g/L)	42.67	2.97	38.60	5.64	0.001
LDH (U/L)	160.53	30.22	167.91	53.03	0.969
GFR (ml/min)	81.65	20.15	73.06	22.97	0.110
Beta2 microglobulin (mg/L)	2.28	0.83	5.25	4.29	0.000
Alkaline phosphatases (U/L)	63.50	19.10	91.69	75.25	0.104

Tab. 2.15: Patient demographical and laboratory data

Statistical Analyses

Correlation of relevant lab biomarker with mean ADC and rFF values extracted from RS VOI are shown in Figure 2.9 where SMMM represent the patient clinical status of SM or MM.

It is noticeable that mean intensity both in ADC and FF is moderately associated with BMPC% with a $\rho = 0.45$ and 0.43 respectively ($p < 0.01$). Scatterplot of ADC and rFF against BMPC% is shown in figure 2.9b and 2.9c respectively.

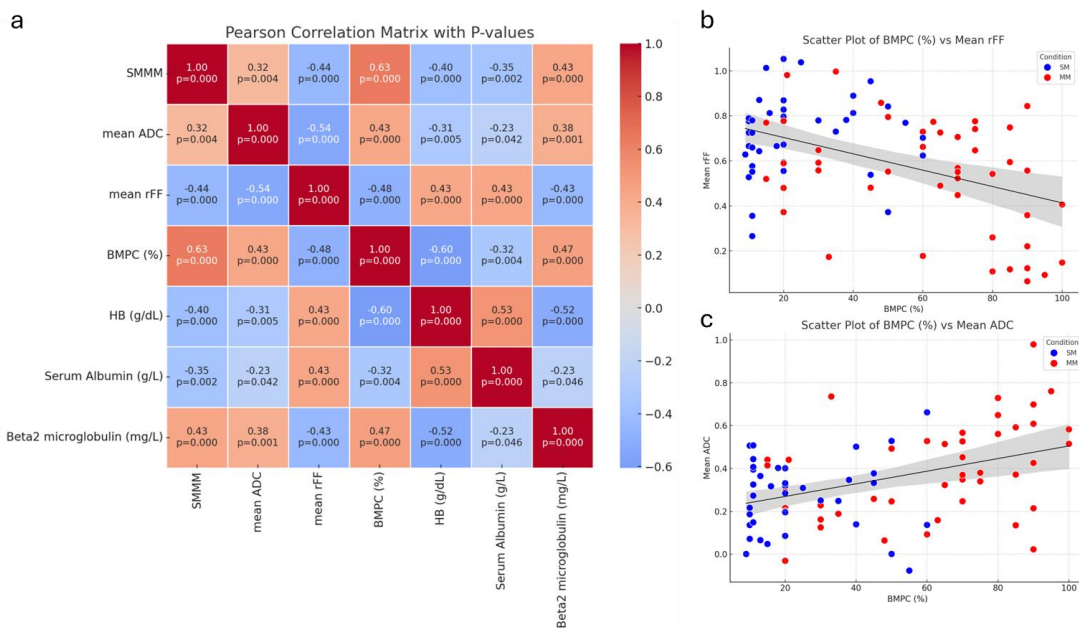


Fig.2.9: Correlation between mean intensity of ADC and rFF and relevant lab biomarkers (a). Scatterplots of correlation between BMPC and rFF (b) and ADC (c).

144 histogram, second and higher order QIBs were extracted from the 6 VOIs contoured on WB-MRI ADC and rFF sequences for each patient. After applying ICC on RS and RI, 23 stable QIBs were obtained for ADC and 43 for rFF confirming the general higher reproducibility of rFF sequences compared to ADC. Pearson ρ correlation coefficient was applied with a threshold of 0.85 to the previously selected QIBs to remove redundancy.

We ended up with 15 independent QIBs both for ADC and rFF sequences to be tested as predictors of MM. LASSO was applied on all the VOIs contoured to assess the mostly selected QIBs related to the clinical diagnosis of SM or MM both for ADC and rFF. Four most frequent QIBs for each imaging modality were selected for further analysis to maximize prediction and avoid overfitting. Mean intensity, complexity, zone distance non uniformity, and the coefficient of variation were selected for rFF sequences whereas mean intensity, zone size non uniformity, 90th intensity percentile and zone distance non uniformity were selected for ADC. These QIBs were finally employed to build 3 -fold cross validated logistic regression-based models on the different VOIs as shown in Table 2.16 where we find an average Area Under the ROC (Receiver operating characteristic) Curve (AUC) of 0.80 in training and 0.70 in test with a slight non statistically significant variability ($p>0.05$) of the VOI taken into consideration at De Long test.

Model						
(AUC)	Median Train	95% CI low	95% CI up	Median Test	95% CI low	95% CI up
RS	0.80	0.79	0.89	0.76	0.61	0.87
RI	0.80	0.77	0.87	0.71	0.53	0.82
LS	0.80	0.77	0.86	0.73	0.56	0.84
LI	0.78	0.73	0.85	0.66	0.52	0.84
LXX	0.78	0.69	0.82	0.66	0.54	0.82
DXX	0.79	0.61	0.84	0.69	0.53	0.81
Mean AUC	0.80			0.70		

Tab. 2.16: Logistic regression based models performances in training and test with 95% confidence interval (CI)

In figure 2.10 we show the performance of the logistic regression models based on Radiopsy QIBs evaluated on RS compared to a logistic regression model trained only with plasma cell infiltration parameter or only with mean intensity of rFF sequences.

The median AUC for the radiopsy model was 0.80 (0.74-0.87) in training and 0.76 (0.59 – 0.885) in test whereas for mean intensity rFF scored an AUC of 0.74 (0.68 - 0.82) in training and 0.74 (0.55 – 0.83) in test. The AUC of the radiopsy model resulted to be statistically significant different from the mean intensity rFF one in training but not in test.

The BMPC based model scored an AUC of 0.89 (0.87-0.91) and 0.90 (0.87-0.91) in training and test respectively and was statistically significantly different from both the previous models ($p < 0.01$) (Figure 2.10).

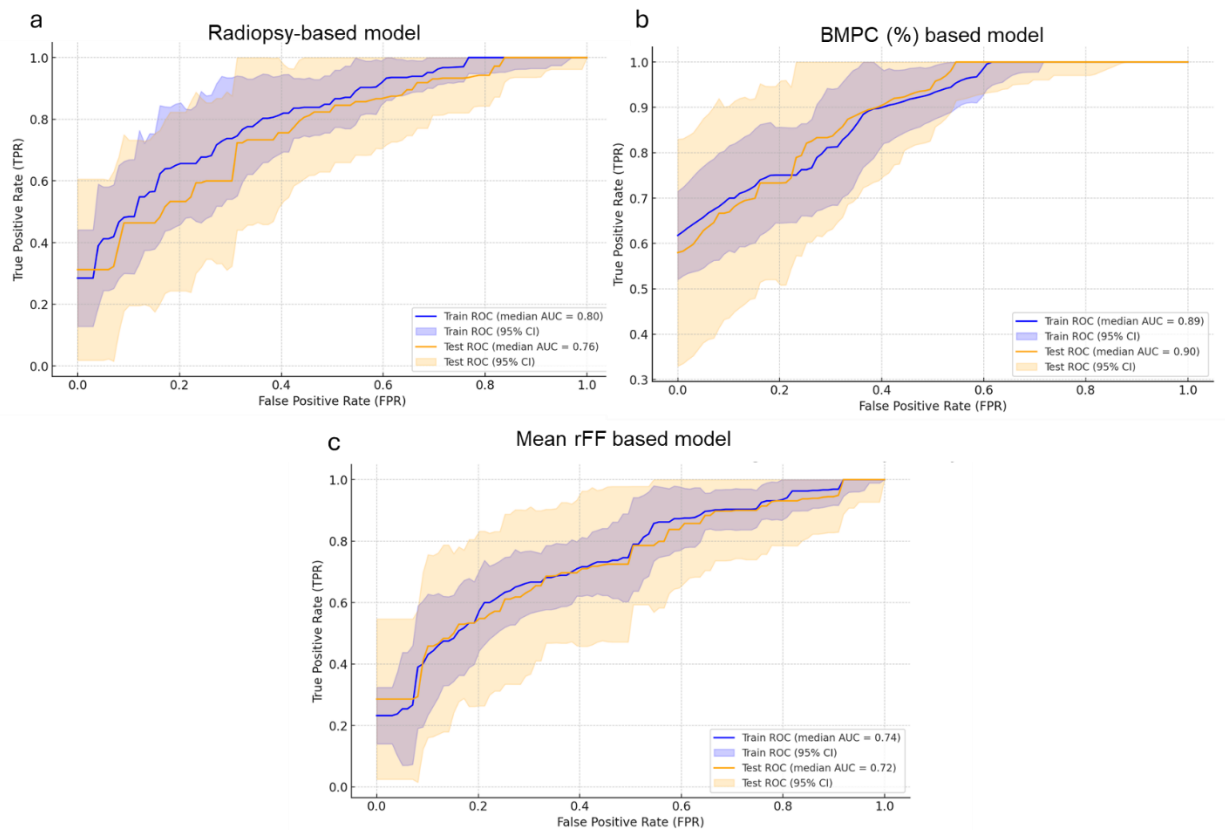


Fig. 2.10: ROC curve for training and test of logistic regression models based on (a) Radiopsy QIBs, (b) BMPC (%) and (c) mean rFF

In the end we select the best radiopsy logistic regression model that had 0.90 test AUC composed by Mean intensity, Complexity, Zone distance non-uniformity and Intensity-based coefficient of variation (from rFF sequences) and it was employed to be tested on the distant VOIs giving the performances summarized in Table 2.16 and the ROC curves given in Figure 2.11. The difference in performances resulted to be not statistically significant.

VOI	Test AUC	Test Precision	Test Recall
RI	0.74	0.71	0.70
LI	0.77	0.72	0.72
LS	0.78	0.75	0.65
LXX	0.69	0.62	0.83
DXX	0.69	0.59	0.83

Tab. 2.16: Numerical performances of Radiopsy model developed on RS and tested on distant VOIs in term of AUC, precision and recall.

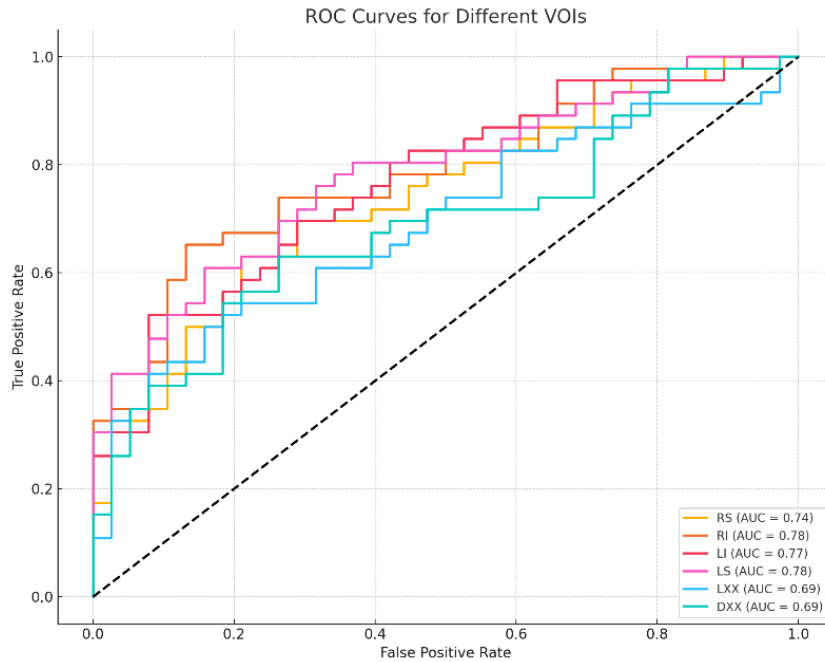


Fig. 2.11: ROC curves showing the performances of Radiopsy model on distant VOIs. AUCs are presented in the legend.

Discussion

The findings from this study underscore the potential of imaging biomarkers derived from WB-MRI sequences, particularly ADC and rFF, in distinguishing between MM and SMM patients. Notably, radiopsy models demonstrated the ability to predict the clinical status non-invasively across various VOIs, maintaining stable predictive performance even when assessments were conducted at locations distant from biopsy sites. The median AUC obtained using the most relevant feature from ADC and rFF

sequences was 0.80 in the training phase and 0.70 in the test phase across all the VOIs taken into consideration.

Among the VOIs evaluated, the RS VOI showed the highest predictive accuracy, achieving an AUC of 0.80 (95% CI: 0.79-0.89) in training and 0.76 (95% CI: 0.61-0.87) in testing. This superior performance suggests the importance of QIBs extracted from regions close to biopsy sites, where disease activity is more closely monitored. The consistency of AUC values across different VOIs, including the LS VOI (median test AUC of 0.73, 95% CI: 0.56-0.84) and the RI VOI (median test AUC of 0.71, 95% CI: 0.53-0.82), highlights the robustness of these models and their potential application across different anatomical locations. We further compared the performances of radiopsy model at RS site with BMPC which is considered one of the most powerful indicators of disease presence.

ROC curves in Figure 2.10 compare the performance of the radiopsy model (AUC = 0.76 in the test set) with BMPC (AUC = 0.90 in the test set), still revealing a statistically significant superiority of BMPC, as expected from the gold standard biomarker used for the myeloma diagnosis. Endeeping the analysis of the radiopsy model, we highlight that there is a notable and statistically significant difference between the radiopsy model's performance in the training set (AUC = 0.80) and the performance when the most important metric of the model alone is mean rFF with an AUC of 0.74, suggesting that higher-order QIBs enhance predictive performance. With an AUC of 0.76 for the Radiopsy model and 0.72 for mean intensity rFF in the test set, the difference between the two models is not statistically significant, but the Radiopsy model nevertheless demonstrates better prediction abilities. BMPC% evaluation provides a superior benefit for high-threshold probability patients, indicating that the BMPC model offers a better balance between sensitivity and specificity in this range, reducing false positives while still capturing key true positives. Radiopsy model can offers complementary information to PC% when a biopsy is not feasible or for non-invasively follow-up.

Given the known heterogeneous diffuse spatial nature of multiple myeloma, we also evaluated the Radiopsy model's ability to predict patient status in distant VOIs. The model achieved encouraging results, with a median AUC of 0.75 when evaluating VOIs on the pelvic bone and 0.69 when assessing vertebrae. These findings suggest that the Radiopsy model can effectively evaluate distant invasion in the pelvic bone, though we observed a performance drop when assessing vertebrae.

The importance of imaging biomarkers in the radiomic model is underline by the mean intensity of Fat Fraction and ADC emerging as strong predictors of disease, showing moderate correlation with BMPC% at 0.45 and 0.43, respectively. These findings are consistent with those of Sun et al. [105], who reported correlations of 0.60 and 0.49, respectively. Wennmann et al. also investigated radiomic models in a multicenter study of 512 patients from eight centers, reporting a strong correlation (0.71) between radiomic QIBs and BMPC values [107] . However, direct comparisons with our study are difficult, they observed that patients with high BMPC had lower feature values for metrics such as the "first-order numeric mode value," which aligns with our findings on mean ADC.

Moreover, Latifoltojar et al. [108] and Koutoulidis et al. [53] studied mean values of rFF and ADC, along with radiomic QIBs, to differentiate between patients with very good partial response and those with partial response. Latifoltojar found that mean rFF had an AUC of 0.95 (95% confidence interval 0.87–1.00), whereas other MRI biomarkers such mean ADC were less reliable in predicting patient outcomes. Similarly, Koutoulidis demonstrated that mean and median rFF intensity, along with the 90th percentile, distinguished between these patient groups with AUC values of 0.82, 0.84, and 0.85, respectively. These same QIBs are also present in our analysis, which seeks to distinguish between SM and MM. This study will be considered in our further investigation, particularly in evaluating the progression of SMM to MM during follow-up.

Reproducibility remains a challenge, particularly for texture QIBs derived from ADC sequences. Wennmann et al. [107] highlighted the variability of radiomic QIBs across different MRI sequences and scanners, with only a small subset achieving acceptable reproducibility, even after applying normalization techniques. Their study found that only 4% of texture QIBs exhibited an $ICC \geq 0.8$ across all experimental conditions, emphasizing the influence of scanning parameters and hardware on feature stability. In line with these findings, our study observed similar variability in ADC-derived QIBs, necessitating a lower threshold for ICC acceptability (0.5). In contrast, FF-derived QIBs demonstrated greater stability. The variability observed in our results supports the conclusions of Wennmann et al., indicating that while textural QIBs from ADC sequences hold diagnostic promise, their reliability is often undermined by variations in acquisition parameters.

Despite these challenges, our findings reinforce the potential of radiomic QIBs from WB-MRI, particularly ADC and rFF, in assessing disease burden in MM and SMM. While QIBs may offer complementary insights into disease heterogeneity and clonal diversity, they should be used in conjunction with traditional biopsy-derived metrics, rather than as standalone diagnostic tools, to enhance patient stratification and treatment planning in clinical practise.

Our study has some main limitation: first the small sample size and lack of prognostic value; second this analysis is based only on images from one MRI scanner in patients enrolled in a single centre study, however it has already been activated a multicenter data collection from other Italian research centers that use multiparametric WB-MRI to external validate our results. Current validation relies on applying the Radiopsy model to distant locations and doesn't account for scanner variability.

Conclusion

In conclusion, the integration of QIBs with existing clinical biomarkers holds promise for improving precision diagnostic and medicine of multiple myeloma. Future studies with large cohorts and external validation are warranted to confirm our findings and further refine its clinical utility.

REFERENCES

- 1- Malard F, Neri P, Bahlis NJ, Terpos E, Moukalled N, Hungria VTM, Manier S, Mohty M. Multiple myeloma. *Nat Rev Dis Primers*. 2024 Jun 27;10(1):45. doi: 10.1038/s41572-024-00529-7. PMID: 38937492.
- 2- <https://www.mayoclinic.org/diseases-conditions/multiple-myeloma/symptoms-causes/syc-20353378>
- 3- Rajkumar SV, Dimopoulos MA, Palumbo A, Blade J, Merlini G, Mateos MV, Kumar S, Hillengass J, Kastritis E, Richardson P, Landgren O, Paiva B, Dispenzieri A, Weiss B, LeLeu X, Zweegman S, Lonial S, Rosinol L, Zamagni E, Jagannath S, Sezer O, Kristinsson SY, Caers J, Usmani SZ, Lahuerta JJ, Johnsen HE, Beksac M, Cavo M, Goldschmidt H, Terpos E, Kyle RA, Anderson KC, Durie BG, Miguel JF. International Myeloma Working Group updated criteria for the diagnosis of multiple myeloma. *Lancet Oncol*. 2014 Nov;15(12):e538-48. doi: 10.1016/S1470-2045(14)70442-5. Epub 2014 Oct 26. PMID: 25439696
- 4- Zhu DT, Park A, Lai A, Zhang L, Attar H, Rebbeck TR. Multiple myeloma incidence and mortality trends in the United States, 1999-2020. *Sci Rep*. 2024 Jun 24;14(1):14564. doi: 10.1038/s41598-024-65590-4. PMID: 38914692; PMCID: PMC11196710
- 5- Cavo M, Rajkumar SV, Palumbo A, Moreau P, Orlowski R, Bladé J, Sezer O, Ludwig H, Dimopoulos MA, Attal M, Sonneveld P, Boccadoro M, Anderson KC, Richardson PG, Bensinger W, Johnsen HE, Kroeger N, Gahrton G, Bergsagel PL, Vesole DH, Einsele H, Jagannath S, Niesvizky R, Durie BG, San Miguel J, Lonial S; International Myeloma Working Group. International Myeloma Working Group consensus approach to the treatment of multiple myeloma patients who are candidates for autologous stem cell transplantation. *Blood*. 2011 Jun 9;117(23):6063-73. doi: 10.1182/blood-2011-02-297325. Epub 2011 Mar 29. PMID: 21447828; PMCID: PMC3293742
- 6- Ludwig H, Kainz S, Schreder M, Zojer N, Hinke A. SLiM CRAB criteria revisited: temporal trends in prognosis of patients with smoldering multiple myeloma who meet the definition of 'biomarker-defined early multiple myeloma'-a systematic review with meta-analysis. *EClinicalMedicine*. 2023 Mar

- 20;58:101910. doi: 10.1016/j.eclinm.2023.101910. PMID: 36969337; PMCID: PMC10033724
- 7- Mateos MV, Kumar S, Dimopoulos MA, González-Calle V, Kastritis E, Hajek R, De Larrea CF, Morgan GJ, Merlini G, Goldschmidt H, Gheraldes C, Gozzetti A, Kyriakou C, Garderet L, Hansson M, Zamagni E, Fantl D, Leleu X, Kim BS, Esteves G, Ludwig H, Usmani S, Min CK, Qi M, Ukropec J, Weiss BM, Rajkumar SV, Durie BGM, San-Miguel J. International Myeloma Working Group risk stratification model for smoldering multiple myeloma (SMM). *Blood Cancer J.* 2020 Oct 16;10(10):102. doi: 10.1038/s41408-020-00366-3. PMID: 33067414; PMCID: PMC7567803.
 - 8- Rajkumar SV, Kumar S, Lonial S, Mateos MV. Smoldering multiple myeloma current treatment algorithms. *Blood Cancer J.* 2022 Sep 5;12(9):129. doi: 10.1038/s41408-022-00719-0. PMID: 36064707; PMCID: PMC9445066.
 - 9- Zamagni E, Tacchetti P, Cavo M. Imaging in multiple myeloma: How? When? *Blood.* 2019 Feb 14;133(7):644-651. doi: 10.1182/blood-2018-08-825356. Epub 2018 Dec 26. PMID: 30587527
 - 10- Wu F, Bernard S, Fayad LM, Ilaslan H, Messiou C, Moulopoulos LA, Mulligan ME. Updates and Ongoing Challenges in Imaging of Multiple Myeloma: *AJR Expert Panel Narrative Review.* *AJR Am J Roentgenol.* 2021 Oct;217(4):775-785. doi: 10.2214/AJR.21.25878. Epub 2021 May 12. PMID: 33978464
 - 11- Rossi A, Cattabriga A, Bezzi D. Symptomatic Myeloma: PET, Whole-Body MR Imaging with Diffusion-Weighted Imaging or Both. *PET Clin.* 2024 Oct;19(4):525-534. doi: 10.1016/j.cpet.2024.05.004. Epub 2024 Jul 4. PMID: 38969566
 - 12- Kumar S, Paiva B, Anderson KC, et al. International myeloma working group consensus criteria for response and minimal residual disease assessment in multiple myeloma. *Lancet Oncol* 2016;17:e328–46.
 - 13- Hillengass J, Usmani S, Rajkumar SV, Durie BGM, Mateos MV, Lonial S, Joao C, Anderson KC, García-Sanz R, Riva E, Du J, van de Donk N, Berdeja JG, Terpos E, Zamagni E, Kyle RA, San Miguel J, Goldschmidt H, Giralt S, Kumar S, Raje N, Ludwig H, Ocio E, Schots R, Einsele H, Schjesvold F, Chen WM, Abildgaard N, Lipe BC, Dytfeld D, Wirk BM, Drake M, Cavo M, Lahuerta JJ, Lentzsch S. International myeloma working group consensus recommendations

- on imaging in monoclonal plasma cell disorders. *Lancet Oncol.* 2019 Jun;20(6):e302-e312. doi: 10.1016/S1470-2045(19)30309-2. Erratum in: *Lancet Oncol.* 2019 Jul;20(7):e346. doi: 10.1016/S1470-2045(19)30423-1. PMID: 31162104
- 14-Moulopoulos LA, Koutoulidis V, Hillengass J, Zamagni E, Aquerreta JD, Roche CL, Lentzsch S, Moreau P, Cavo M, Miguel JS, Dimopoulos MA, Rajkumar SV, Durie BGM, Terpos E, Delorme S. 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 Oct 4;8(10):95. doi: 10.1038/s41408-018-0124-1. PMID: 30287814; PMCID: PMC6172202
- 15-Myeloma Diagnosis and Management.
<https://www.nice.org.uk/guidance/ng35/evidence/appendices-af-pdf-2306487278>
- 16-Chantry A, Kazmi M, Barrington S, Goh V, Mulholland N, Streetly M, Lai M, Pratt G; British Society for Haematology Guidelines. Guidelines for the use of imaging in the management of patients with myeloma. *Br J Haematol.* 2017 Aug;178(3):380-393. doi: 10.1111/bjh.14827. Epub 2017 Jul 5. PMID: 28677897.
- 17-Zamagni E, Cavo M, Fakhri B, Vij R, Roodman D. Bones in Multiple Myeloma: Imaging and Therapy. *Am Soc Clin Oncol Educ Book.* 2018 May 23;38:638-646. doi: 10.1200/EDBK_205583. PMID: 30231385.
- 18-Dimopoulos MA, Hillengass J, Usmani S, Zamagni E, Lentzsch S, Davies FE, Raje N, Sezer O, Zweegman S, Shah J, Badros A, Shimizu K, Moreau P, Chim CS, Lahuerta JJ, Hou J, Jurczyszyn A, Goldschmidt H, Sonneveld P, Palumbo A, Ludwig H, Cavo M, Barlogie B, Anderson K, Roodman GD, Rajkumar SV, Durie BG, Terpos E. Role of magnetic resonance imaging in the management of patients with multiple myeloma: a consensus statement. *J Clin Oncol.* 2015 Feb 20;33(6):657-64. doi: 10.1200/JCO.2014.57.9961. Epub 2015 Jan 20. PMID: 25605835.
- 19-Treitl KM, Ricke J, Baur-Melnyk A. Whole-body magnetic resonance imaging (WBMRI) versus whole-body computed tomography (WBCT) for myeloma imaging and staging. *Skeletal Radiol.* 2022 Jan;51(1):43-58. doi:

- 10.1007/s00256-021-03799-4. Epub 2021 May 24. PMID: 34031705; PMCID: PMC8626374.
- 20- Nanni C. PET-FDG: Impetus. *Cancers* (Basel). 2020 Apr 22;12(4):1030. doi: 10.3390/cancers12041030. PMID: 32331374; PMCID: PMC7226158.
- 21- Cruz IAN, Fayad LM, Ahlawat S, Lederman HM, Nico MAC, Ormond Filho AG, Guimarães JB. Whole-Body MRI in Musculoskeletal Oncology: A Comprehensive Review with Recommendations. *Radiol Imaging Cancer*. 2023 May;5(3):e220107. doi: 10.1148/rycan.220107. PMID: 37144975; PMCID: PMC10240252.
- 22- Messiou C, Hillengass J, Delorme S, Lecouvet FE, Moulopoulos LA, Collins DJ, Blackledge MD, Abildgaard N, Østergaard B, Schlemmer HP, Landgren O, Asmussen JT, Kaiser MF, Padhani A. Guidelines for Acquisition, Interpretation, and Reporting of Whole-Body MRI in Myeloma: Myeloma Response Assessment and Diagnosis System (MY-RADS). *Radiology*. 2019 Apr;291(1):5-13. doi: 10.1148/radiol.2019181949. Epub 2019 Feb 26. PMID: 30806604.
- 23- Koh DM, Blackledge M, Padhani AR, Takahara T, Kwee TC, Leach MO, Collins DJ. Whole-body diffusion-weighted MRI: tips, tricks, and pitfalls. *AJR Am J Roentgenol*. 2012 Aug;199(2):252-62. doi: 10.2214/AJR.11.7866. PMID: 22826385
- 24- Tunariu N, Blackledge M, Messiou C, Petralia G, Padhani A, Curcean S, Curcean A, Koh DM. What's New for Clinical Whole-body MRI (WB-MRI) in the 21st Century. *Br J Radiol*. 2020 Nov 1;93(1115):20200562. doi: 10.1259/bjr.20200562. Epub 2020 Sep 9. PMID: 32822545; PMCID: PMC8519652.
- 25- Padhani AR, Makris A, Gall P, Collins DJ, Tunariu N, de Bono JS. Therapy monitoring of skeletal metastases with whole-body diffusion MRI. *J Magn Reson Imaging*. 2014 May;39(5):1049-78. doi: 10.1002/jmri.24548. Epub 2014 Feb 10. PMID: 24510426.
- 26- Padhani AR, Lecouvet F, Petralia G, Koh DM. Re: Alonso Garcia-Ruiz, Carlos Macarro, Francesca Zacchi, et al. Whole-body Magnetic Resonance Imaging as a Treatment Response Biomarker in Castration-resistant Prostate Cancer with Bone Metastases: The iPROMET Clinical Trial. *Eur Urol*. In press. <https://doi.org/10.1016/j.eururo.2024.02.016>. *Eur Urol*. 2024 Sep;86(3):e67-e68. doi: 10.1016/j.eururo.2024.04.035. Epub 2024 May 17. PMID: 38760293.

- 27-van Vucht N, Santiago R, Lottmann B, Pressney I, Harder D, Sheikh A, Saifuddin A. The Dixon technique for MRI of the bone marrow. *Skeletal Radiol*. 2019 Dec;48(12):1861-1874. doi: 10.1007/s00256-019-03271-4. Epub 2019 Jul 15. PMID: 31309243.
- 28-Lecouvet FE, Chabot C, Taihi L, Kirchgesner T, Triqueneaux P, Malghem J. Present and future of whole-body MRI in metastatic disease and myeloma: how and why you will do it. *Skeletal Radiol*. 2024 Sep;53(9):1815-1831. doi: 10.1007/s00256-024-04723-2. Epub 2024 Jul 15. PMID: 39007948; PMCID: PMC11303436
- 29-Padhani AR, Lecouvet FE, Tunariu N, Koh DM, De Keyzer F, Collins DJ, Sala E, Schlemmer HP, Petralia G, Vargas HA, Fanti S, Tombal HB, de Bono J. METastasis Reporting and Data System for Prostate Cancer: Practical Guidelines for Acquisition, Interpretation, and Reporting of Whole-body Magnetic Resonance Imaging-based Evaluations of Multiorgan Involvement in Advanced Prostate Cancer. *Eur Urol*. 2017 Jan;71(1):81-92. doi: 10.1016/j.eururo.2016.05.033. Epub 2016 Jun 14. PMID: 27317091; PMCID: PMC5176005.
- 30-Donners R, Blackledge M, Tunariu N, Messiou C, Merkle EM, Koh DM. Quantitative Whole-Body Diffusion-Weighted MR Imaging. *Magn Reson Imaging Clin N Am*. 2018 Nov;26(4):479-494. doi: 10.1016/j.mric.2018.06.002. Epub 2018 Sep 15. PMID: 30316462.
- 31-Hameed M, Sandhu A, Soneji N, Amiras D, Rockall A, Messiou C, Wallitt K, Barwick TD. Pictorial review of whole body MRI in myeloma: emphasis on diffusion-weighted imaging. *Br J Radiol*. 2020 Nov 1;93(1115):20200312. doi: 10.1259/bjr.20200312. Epub 2020 Aug 26. PMID: 32667830; PMCID: PMC8519646
- 32-Ormond Filho AG, Carneiro BC, Pastore D, Silva IP, Yamashita SR, Consolo FD, Hungria VTM, Sandes AF, Rizzatti EG, Nico MAC. Whole-Body Imaging of Multiple Myeloma: Diagnostic Criteria. *Radiographics*. 2019 Jul-Aug;39(4):1077-1097. doi: 10.1148/rg.2019180096. PMID: 31283452..
- 33-Pawlyn C, Fowkes L, Otero S, Jones JR, Boyd KD, Davies FE, Morgan GJ, Collins DJ, Sharma B, Riddell A, Kaiser MF, Messiou C. Whole-body diffusion-weighted MRI: a new gold standard for assessing disease burden in patients with multiple myeloma? *Leukemia*. 2016 Jun;30(6):1446-8. doi:

- 10.1038/leu.2015.338. Epub 2015 Dec 9. PMID: 26648535; PMCID: PMC4895156.
- 34- Sachpekidis C, Mosebach J, Freitag MT, Wilhelm T, Mai EK, Goldschmidt H, Haberkorn U, Schlemmer HP, Delorme S, Dimitrakopoulou-Strauss A. Application of (18)F-FDG PET and diffusion weighted imaging (DWI) in multiple myeloma: comparison of functional imaging modalities. *Am J Nucl Med Mol Imaging*. 2015 Oct 12;5(5):479-92. PMID: 26550539; PMCID: PMC4620175.
 - 35- Matteucci F, Paganelli G, Martinelli G, Cerchione C. PET/CT in Multiple Myeloma: Beyond FDG. *Front Oncol*. 2021 Jan 25;10:622501. doi: 10.3389/fonc.2020.622501. PMID: 33569348; PMCID: PMC7868556.
 - 36- Rasche L, Angtuaco E, McDonald JE, Buros A, Stein C, Pawlyn C, Thanendrarajan S, Schinke C, Samant R, Yaccoby S, Walker BA, Epstein J, Zangari M, van Rhee F, Meissner T, Goldschmidt H, Hemminki K, Houlston R, Barlogie B, Davies FE, Morgan GJ, Weinhold N. Low expression of hexokinase-2 is associated with false-negative FDG-positron emission tomography in multiple myeloma. *Blood*. 2017 Jul 6;130(1):30-34. doi: 10.1182/blood-2017-03-774422. Epub 2017 Apr 21. PMID: 28432222; PMCID: PMC5501152;
 - 37- Lecouvet FE, Vekemans MC, Van Den Berghe T, Verstraete K, Kirchgesner T, Acid S, Malghem J, Wuts J, Hillengass J, Vandecaveye V, Jamar F, Gheysens O, Vande Berg BC. Imaging of treatment response and minimal residual disease in multiple myeloma: state of the art WB-MRI and PET/CT. *Skeletal Radiol*. 2022 Jan;51(1):59-80. doi: 10.1007/s00256-021-03841-5. Epub 2021 Aug 7. PMID: 34363522; PMCID: PMC8626399
 - 38- Messiou C, Porta N, Sharma B, Levine D, Koh DM, Boyd K, Pawlyn C, Riddell A, Downey K, Croft J, Morgan V, Stern S, Cheung B, Kyriakou C, Kaczmarek P, Winfield J, Blackledge M, Oyen WJG, Kaiser MF. Prospective Evaluation of Whole-Body MRI versus FDG PET/CT for Lesion Detection in Participants with Myeloma. *Radiol Imaging Cancer*. 2021 Sep;3(5):e210048. doi: 10.1148/rycan.2021210048. PMID: 34559006; PMCID: PMC8489453
 - 39- Mouloupoulos LA, Dimopoulos MA, Kastiris E, Christoulas D, Gkatzamanidou M, Roussou M, Koureas A, Migkou M, Gavriatopoulou M, Eleutherakis-Papaiaikovou E, Gika D, Koutoulidis V, Terpos E. Diffuse pattern of bone

- marrow involvement on magnetic resonance imaging is associated with high risk cytogenetics and poor outcome in newly diagnosed, symptomatic patients with multiple myeloma: a single center experience on 228 patients. *Am J Hematol*. 2012 Sep;87(9):861-4. doi: 10.1002/ajh.23258. Epub 2012 May 28. PMID: 22641455.
- 40-Belotti A, Ribolla R, Cancelli V, Villanacci A, Angelini V, Chiarini M, Giustini V, Facchetti GV, Roccaro AM, Ferrari S, Peli A, Bottelli C, Cattaneo C, Crippa C, Micilotta M, Frittoli B, Grazioli L, Rossi G, Tucci A. Predictive role of diffusion-weighted whole-body MRI (DW-MRI) imaging response according to MY-RADS criteria after autologous stem cell transplantation in patients with multiple myeloma and combined evaluation with MRD assessment by flow cytometry. *Cancer Med*. 2021 Sep;10(17):5859-5865. doi: 10.1002/cam4.4136. Epub 2021 Jul 15. PMID: 34263564; PMCID: PMC8419770.
- 41-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 Jan;75:60-64. doi: 10.1016/j.mri.2020.10.005. Epub 2020 Oct 16. PMID: 33075451;
- 42-Rasche L, Alapat D, Kumar M, Gershner G, McDonald J, Wardell CP, Samant R, Van Hemert R, Epstein J, Williams AF, Thanendrarajan S, Schinke C, Bauer M, Ashby C, Tytarenko RG, van Rhee F, Walker BA, Zangari M, Barlogie B, Davies FE, Morgan GJ, Weinhold N. Combination of flow cytometry and functional imaging for monitoring of residual disease in myeloma. *Leukemia*. 2019 Jul;33(7):1713-1722. doi: 10.1038/s41375-018-0329-0. Epub 2018 Dec 20. PMID: 30573775; PMCID: PMC6586541
- 43-Rama S, Suh CH, Kim KW, Durieux JC, Ramaiya NH, Tirumani SH. Comparative Performance of Whole-Body MRI and FDG PET/CT in Evaluation of Multiple Myeloma Treatment Response: Systematic Review and Meta-Analysis. *AJR Am J Roentgenol*. 2022 Apr;218(4):602-613. doi: 10.2214/AJR.21.26381. Epub 2021 Oct 27. PMID: 34704461
- 44-<https://www.myesr.org/research/biomarkers-inventory>
- 45-Califf RM. Biomarker definitions and their applications. *Exp Biol Med* (Maywood). 2018 Feb;243(3):213-221. doi: 10.1177/1535370217750088. PMID: 29405771; PMCID: PMC5813875

- 46- Suh CH, Yun SJ, Jin W, Lee SH, Park SY, Ryu CW. ADC as a useful diagnostic tool for differentiating benign and malignant vertebral bone marrow lesions and compression fractures: a systematic review and meta-analysis. *Eur Radiol.* 2018 Jul;28(7):2890-2902. doi: 10.1007/s00330-018-5330-5. Epub 2018 Feb 15. PMID: 29450718
- 47- Michoux NF, Ceranka JW, Vandemeulebroucke J, Peeters F, Lu P, Absil J, Triqueneaux P, Liu Y, Collette L, Willekens I, Brussaard C, Debeir O, Hahn S, Raeymaekers H, de Mey J, Metens T, Lecouvet FE. Repeatability and reproducibility of ADC measurements: a prospective multicenter whole-body-MRI study. *Eur Radiol.* 2021 Jul;31(7):4514-4527. doi: 10.1007/s00330-020-07522-0. Epub 2021 Jan 6. PMID: 33409773
- 48- Castagnoli F, Donners R, Tunariu N, Messiou C, Koh DM. Relative fat fraction of malignant bone lesions from breast cancer, prostate cancer and myeloma are significantly lower than normal bone marrow and shows excellent interobserver agreement. *Br J Radiol.* 2023 Dec;96(1152):20230240. doi: 10.1259/bjr.20230240. Epub 2023 Oct 31. PMID: 37750943; PMCID: PMC10646620.
- 49- Donners R, Figueiredo I, Tunariu N, Blackledge M, Koh DM, de la Maza MLDF, Chandran K, de Bono JS, Fotiadis N. Multiparametric bone MRI can improve CT-guided bone biopsy target selection in cancer patients and increase diagnostic yield and feasibility of next-generation tumour sequencing. *Eur Radiol.* 2022 Jul;32(7):4647-4656. doi: 10.1007/s00330-022-08536-6. Epub 2022 Jan 29. PMID: 35092476; PMCID: PMC9213271.
- 50- Schmeel FC, Enkirch SJ, Luetkens JA, Faron A, Lehnen N, Sprinkart AM, Schmeel LC, Radbruch A, Attenberger U, Kukuk GM, Mürtz P. Diagnostic Accuracy of Quantitative Imaging Biomarkers in the Differentiation of Benign and Malignant Vertebral Lesions : Combination of Diffusion-Weighted and Proton Density Fat Fraction Spine MRI. *Clin Neuroradiol.* 2021 Dec;31(4):1059-1070. doi: 10.1007/s00062-021-01009-1. Epub 2021 Mar 31. PMID: 33787957; PMCID: PMC8648653
- 51- Winfield JM, Blackledge MD, Tunariu N, Koh DM, Messiou C. Whole-body MRI: a practical guide for imaging patients with malignant bone disease. *Clin Radiol.* 2021 Oct;76(10):715-727. doi: 10.1016/j.crad.2021.04.001. Epub 2021 Apr 30. PMID: 33934876.

- 52-Koutoulidis V, Terpos E, Papanikolaou N, Fontara S, Seimenis I, Gavriatopoulou M, Ntanasis-Stathopoulos I, Bourgioti C, Santinha J, Moreira JM, Kastiris E, Dimopoulos MA, Moulopoulos LA. Comparison of MRI Features of Fat Fraction and ADC for Early Treatment Response Assessment in Participants with Multiple Myeloma. *Radiology*. 2022 Jul;304(1):137-144. doi: 10.1148/radiol.211388. Epub 2022 Apr 5. PMID: 35380497.
- 53-Berardo S, Sukhovei L, Andorno S, Carriero A, Stecco A. Quantitative bone marrow magnetic resonance imaging through apparent diffusion coefficient and fat fraction in multiple myeloma patients. *Radiol Med*. 2021 Mar;126(3):445-452. doi: 10.1007/s11547-020-01258-z. Epub 2020 Aug 18. PMID: 32812173.
- 54-Albano D, Stecco A, Micci G, Sconfienza LM, Colagrande S, Reginelli A, Grassi R, Carriero A, Midiri M, Lagalla R, Galia M. Whole-body magnetic resonance imaging (WB-MRI) in oncology: an Italian survey. *Radiol Med*. 2021 Feb;126(2):299-305. doi: 10.1007/s11547-020-01242-7. Epub 2020 Jun 22. PMID: 32572763.
- 55-Rata M, Blackledge M, Scurr E, Winfield J, Koh DM, Dragan A, Candito A, King A, Rennie W, Gaba S, Suresh P, Malcolm P, Davis A, Nilak A, Shah A, Gandhi S, Albrizio M, Drury A, Roberts S, Jenner M, Brown S, Kaiser M, Messiou C. Implementation of Whole-Body MRI (MY-RADS) within the OPTIMUM/MUKnine multi-centre clinical trial for patients with myeloma. *Insights Imaging*. 2022 Jul 28;13(1):123. doi: 10.1186/s13244-022-01253-0. PMID: 35900614; PMCID: PMC9334517.
- 56-Keaveney S, Dragan A, Rata M, Blackledge M, Scurr E, Winfield JM, Shur J, Koh DM, Porta N, Candito A, King A, Rennie W, Gaba S, Suresh P, Malcolm P, Davis A, Nilak A, Shah A, Gandhi S, Albrizio M, Drury A, Pratt G, Cook G, Roberts S, Jenner M, Brown S, Kaiser M, Messiou C. Image quality in whole-body MRI using the MY-RADS protocol in a prospective multi-centre multiple myeloma study. *Insights Imaging*. 2023 Oct 15;14(1):170. doi: 10.1186/s13244-023-01498-3. PMID: 37840055; PMCID: PMC10577121.
- 57-Barnes A, Alonzi R, Blackledge M, Charles-Edwards G, Collins DJ, Cook G, Coutts G, Goh V, Graves M, Kelly C, Koh DM, McCallum H, Miquel ME, O'Connor J, Padhani A, Pearson R, Priest A, Rockall A, Stirling J, Taylor S, Tunariu N, van der Meulen J, Walls D, Winfield J, Punwani S. UK quantitative WB-DWI technical workgroup: consensus meeting recommendations on

- optimisation, quality control, processing and analysis of quantitative whole-body diffusion-weighted imaging for cancer. *Br J Radiol.* 2018 Jan;91(1081):20170577. doi: 10.1259/bjr.20170577. Epub 2017 Dec 7. PMID: 29076749; PMCID: PMC5966219.
- 58- https://qibawiki.rsna.org/images/6/63/MRI-Based_PDF_of_the_Liver_QIBA_Profile_2024_06-19_CONSENSUS-maintenance.pdf
- 59- https://qibawiki.rsna.org/images/b/b0/QIBA_DWIProfile_Stage3_15Dec2022_v3.pdf
- 60- Negroni D, Cassarà A, Trisoglio A, Soligo E, Berardo S, Carriero A, Stecco A. Learning curves in radiological reporting of whole-body MRI in plasma cell disease: a retrospective study. *Radiol Med.* 2021 Nov;126(11):1451-1459. doi: 10.1007/s11547-021-01391-3. Epub 2021 Jul 26. PMID: 34309766; PMCID: PMC8558285.
- 61- Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology.* 2016 Feb;278(2):563-77. doi: 10.1148/radiol.2015151169. Epub 2015 Nov 18. PMID: 26579733; PMCID: PMC4734157;
- 62- Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, Sanduleanu S, Larue RTHM, Even AJG, Jochems A, van Wijk Y, Woodruff H, van Soest J, Lustberg T, Roelofs E, van Elmpt W, Dekker A, Mottaghy FM, Wildberger JE, Walsh S. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol.* 2017 Dec;14(12):749-762. doi: 10.1038/nrclinonc.2017.141. Epub 2017 Oct 4. PMID: 28975929;
- 63- Horvat N, Papanikolaou N, Koh DM. Radiomics Beyond the Hype: A Critical Evaluation Toward Oncologic Clinical Use. *Radiol Artif Intell.* 2024 Jul;6(4):e230437. doi: 10.1148/ryai.230437. PMID: 38717290; PMCID: PMC11294952.).
- 64- Kocak B, Akinci D'Antonoli T, Mercaldo N, Alberich-Bayarri A, Baessler B, Ambrosini I, Andreychenko AE, Bakas S, Beets-Tan RGH, Bressen K, Buvat I, Cannella R, Cappellini LA, Cavallo AU, Chepelev LL, Chu LCH, Demircioglu A, deSouza NM, Dietzel M, Fanni SC, Fedorov A, Fournier LS, Giannini V, Girometti R, Groot Lipman KBW, Kalarakis G, Kelly BS, Klontzas ME, Koh DM, Kotter E, Lee HY, Maas M, Marti-Bonmati L, Müller H, Obuchowski N,

- Orlhac F, Papanikolaou N, Petrash E, Pfaehler E, Pinto Dos Santos D, Ponsiglione A, Sabater S, Sardanelli F, Seeböck P, Sijtsma NM, Stanzione A, Traverso A, Ugga L, Vallières M, van Dijk LV, van Griethuysen JJM, van Hamersvelt RW, van Ooijen P, Vernuccio F, Wang A, Williams S, Witowski J, Zhang Z, Zwanenburg A, Cuocolo R. METHodological RadiomICs Score (METRICS): a quality scoring tool for radiomics research endorsed by EuSoMII. *Insights Imaging*. 2024 Jan 17;15(1):8. doi: 10.1186/s13244-023-01572-w. PMID: 38228979; PMCID: PMC10792137.
- 65-Stewart C, Davenport MS, Miglioretti DL, Smith-Bindman R. Types of Evidence Needed to Assess the Clinical Value of Diagnostic Imaging. *NEJM Evid*. 2024 Jul;3(7):EVIDra2300252. doi: 10.1056/EVIDra2300252. Epub 2024 Jun 25. PMID: 38916414
- 66-Rajiah P, Bolen MA. Cardiovascular MR imaging at 3 T: opportunities, challenges, and solutions. *Radiographics*. 2014 Oct;34(6):1612-35. doi: 10.1148/rg.346140048. PMID: 25310420.
- 67-Schick F, Pieper CC, Kupczyk P, Almansour H, Keller G, Springer F, Mürtz P, Endler C, Sprinkart AM, Kaufmann S, Herrmann J, Attenberger UI. 1.5 vs 3 Tesla Magnetic Resonance Imaging: A Review of Favorite Clinical Applications for Both Field Strengths-Part 1. *Invest Radiol*. 2021 Nov 1;56(11):680-691. doi: 10.1097/RLI.0000000000000812. PMID: 34324464.
- 68-Documento SIRM- sicurezza in risonanza magnetica 2021 - Sinossi per il radiologo
- 69- Nordin LE, Åberg K, Kihlberg J, Owman T, Hansson B, Björkman-Burtscher IM, Petersen C, Lundberg P. ESR Essentials: basic physics of MR safety-practice recommendations by the European Society for Magnetic Resonance in Medicine and Biology. *Eur Radiol*. 2024 Aug 13. doi: 10.1007/s00330-024-10999-8. Epub ahead of print. PMID: 39136707.
- 70-Decreto Ministeriale 14 gennaio 2021 - Determinazione degli standard di sicurezza e impiego per le apparecchiature a risonanza magnetica e individuazione di altre tipologie di apparecchiature a risonanza magnetica settoriali non soggette ad autorizzazione. - Gazzetta Ufficiale Serie Generale n. 65 del 16-03-2021.
- 71-Summers P, Saia G, Colombo A, Pricolo P, Zugni F, Alessi S, Marvaso G, Jereczek-Fossa BA, Bellomi M, Petralia G. Whole-body magnetic resonance

- imaging: technique, guidelines and key applications. *Ecancermedicalscience*. 2021 Jan 7;15:1164. doi: 10.3332/ecancer.2021.1164. PMID: 33680078; PMCID: PMC7929776.
- 72- Padhani AR, Gogbashian A. Bony metastases: assessing response to therapy with whole-body diffusion MRI. *Cancer Imaging*. 2011 Oct 3;11 Spec No A(1A):S129-45. doi: 10.1102/1470-7330.2011.9034. PMID: 22185786; PMCID: PMC3266569.
- 73- Takahara T, Imai Y, Yamashita T, Yasuda S, Nasu S, Van Cauteren M. Diffusion weighted whole body imaging with background body signal suppression (DWIBS): technical improvement using free breathing, STIR and high resolution 3D display. *Radiat Med*. 2004 Jul-Aug;22(4):275-82. PMID: 15468951.
- 74- Ji X, Huang W, Dong H, Shen Z, Zheng M, Zou D, Shen W, Xia S. Evaluation of bone marrow infiltration in multiple myeloma using whole-body diffusion-weighted imaging and T1-weighted water-fat separation Dixon. *Quant Imaging Med Surg*. 2021 Feb;11(2):641-651. doi: 10.21037/qims-20-289. PMID: 33532264; PMCID: PMC7779908.
- 75- Disler, D.G.; McCauley, T.R.; Ratner, L.M.; Kesack, C.D.; Cooper, J.A. In-phase and out-of-phase MR imaging of bone marrow: Prediction of neoplasia based on the detection of coexistent fat and water. *AJR Am. J. Roentgenol*. 1997, 169, 1439–1447;
- 76- Douis, H.; Davies, A.M.; Jeys, L.; Sian, P. Chemical shift MRI can aid in the diagnosis of indeterminate skeletal lesions of the spine. *Eur. Radiol*. 2016, 26, 932–940
- 77- Koh DM, Collins DJ. Diffusion-weighted MRI in the body: applications and challenges in oncology. *AJR Am J Roentgenol*. 2007 Jun;188(6):1622-35. doi: 10.2214/AJR.06.1403. PMID: 17515386.
- 78- Lecouvet FE, Larbi A, Pasoglou V, Omoumi P, Tombal B, Michoux N, Malghem J, Lhommel R, Vande Berg BC. MRI for response assessment in metastatic bone disease. *Eur Radiol*. 2013 Jul;23(7):1986-97. doi: 10.1007/s00330-013-2792-3. Epub 2013 Mar 1. PMID: 23455764.
- 79- Schmeel FC, Vomweg T, Träber F, Gerhards A, Enkirch SJ, Faron A, Sprinkart AM, Schmeel LC, Luetkens JA, Thomas D, Kukuk GM. Proton density fat fraction MRI of vertebral bone marrow: Accuracy, repeatability, and

reproducibility among readers, field strengths, and imaging platforms. *J Magn Reson Imaging*. 2019 Dec;50(6):1762-1772. doi: 10.1002/jmri.26748. Epub 2019 Apr 13. PMID: 30980694.

- 80-Fedeli L, Benelli M, Busoni S, Belli G, Ciccarone A, Coniglio A, Esposito M, Nocetti L, Sghedoni R, Tarducci R, Altabella L, Belligotti E, Bettarini S, Betti M, Caivano R, Carnì M, Chiappiniello A, Cimolai S, Cretti F, Fulcheri C, Gasperi C, Giacometti M, Levrero F, Lizio D, Maieron M, Marzi S, Mascaro L, Mazzocchi S, Meliadò G, Morzenti S, Niespolo A, Noferini L, Oberhofer N, Orsingher L, Quattrocchi M, Ricci A, Savini A, Taddeucci A, Testa C, Tortoli P, Gobbi G, Gori C, Bernardi L, Giannelli M, Mazzoni LN; Italian Association of Physics in Medicine (AIFM) Working Group on MR Quantification, Intercomparison and Quality Assurance. On the dependence of quantitative diffusion-weighted imaging on scanner system characteristics and acquisition parameters: A large multicenter and multiparametric phantom study with unsupervised clustering analysis. *Phys Med*. 2021 May;85:98-106. doi: 10.1016/j.ejmp.2021.04.020. Epub 2021 May 12. PMID: 33991807.
- 81-Bush EC, Gifford A, Coolbaugh CL, Towse TF, Damon BM, Welch EB. Fat-Water Phantoms for Magnetic Resonance Imaging Validation: A Flexible and Scalable Protocol. *J Vis Exp*. 2018 Sep 7;(139):57704. doi: 10.3791/57704. PMID: 30247483; PMCID: PMC6235120.
- 82-Itri JN. Patient-centered Radiology. *Radiographics*. 2015 Oct;35(6):1835-46. doi: 10.1148/rg.2015150110. PMID: 26466190.
- 83-Dantendorfer K, Amering M, Bankier A, Helbich T, Prayer D, Youssefzadeh S, Alexandrowicz R, Imhof H, Katschnig H. A study of the effects of patient anxiety, perceptions and equipment on motion artifacts in magnetic resonance imaging. *Magn Reson Imaging*. 1997;15(3):301-6. doi: 10.1016/s0730-725x(96)00385-2. PMID: 9201677.
- 84- Adams HJ, Kwee TC, Vermoolen MA, Ludwig I, Bierings MB, Nievelstein RA. Whole-body MRI vs. CT for staging lymphoma: patient experience. *Eur J Radiol*. 2014 Jan;83(1):163-6. doi: 10.1016/j.ejrad.2013.10.008. Epub 2013 Oct 21. PMID: 24211036.
- 85-Oliveri S, Pricolo P, Pizzoli S, Faccio F, Lampis V, Summers P, Petralia G, Pravettoni G. Investigating cancer patient acceptance of Whole Body MRI. *Clin*

- Imaging. 2018 Nov-Dec;52:246-251. doi: 10.1016/j.clinimag.2018.08.004. Epub 2018 Aug 8. PMID: 30170274.
- 86- Busacchio D, Mazzocco K, Gandini S, Pricolo P, Masiero M, Summers PE, Pravettoni G, Petralia G. Preliminary observations regarding the expectations, acceptability and satisfaction of whole-body MRI in self-referring asymptomatic subjects. *Br J Radiol.* 2021 Feb 1;94(1118):20191031. doi: 10.1259/bjr.20191031. Epub 2020 Dec 1. PMID: 33237810; PMCID: PMC7934315.
- 87- Dyrberg E, Larsen EL, Hendel HW, Thomsen HS. Diagnostic bone imaging in patients with prostate cancer: patient experience and acceptance of NaF-PET/CT, choline-PET/CT, whole-body MRI, and bone SPECT/CT. *Acta Radiol.* 2018 Sep;59(9):1119-1125. doi: 10.1177/0284185117751280. Epub 2018 Jan 9. PMID: 29313360.
- 88- Ryder A, Parsons C, Hutchinson CE, Greaney B, Thake CD. A survey study investigating perceptions and acceptance of the whole-body imaging techniques used for the diagnosis of myeloma. *Radiography (Lond).* 2021 Nov;27(4):1149-1157. doi: 10.1016/j.radi.2021.06.002. Epub 2021 Jul 10. PMID: 34257014.
- 89- Evans R, Taylor S, Janes S, Halligan S, Morton A, Navani N, Oliver A, Rockall A, Teague J, Miles A; Streamline trials investigators. Patient experience and perceived acceptability of whole-body magnetic resonance imaging for staging colorectal and lung cancer compared with current staging scans: a qualitative study. *BMJ Open.* 2017 Sep 6;7(9):e016391. doi: 10.1136/bmjopen-2017-016391. PMID: 28882915; PMCID: PMC5588966.
- 90- Evans RE, Taylor SA, Beare S, Halligan S, Morton A, Oliver A, Rockall A, Miles A. Perceived patient burden and acceptability of whole body MRI for staging lung and colorectal cancer; comparison with standard staging investigations. *Br J Radiol.* 2018 Jun;91(1086):20170731. doi: 10.1259/bjr.20170731. Epub 2018 Mar 20. PMID: 29528257; PMCID: PMC6223281.
- 91- Rossi A, Prochowski Iamurri A, Diano D, Oboldi D, Sintuzzi E, Maurizio L, Andalò A, Cavallucci M, Ferroni F, Amadori E, Barone D, Petralia G. Patient centered radiology: investigating 3 Tesla whole body MRI acceptance in cancer patients. *Radiol Med.* 2023 Aug;128(8):960-969. doi: 10.1007/s11547-023-01665-y. Epub 2023 Jul 3. PMID: 37395842; PMCID: PMC10338551.

- 92-Croft J, Riddell A, Koh DM, Downey K, Blackledge M, Usher M, Boyd K, Kaiser M, Messiou C. Inter-observer agreement of baseline whole body MRI in multiple myeloma. *Cancer Imaging*. 2020 Jul 14;20(1):48. doi: 10.1186/s40644-020-00328-9. PMID: 32665028; PMCID: PMC7362571
- 93-Pricolo P, Ancona E, Summers P, Abreu-Gomez J, Alessi S, Jereczek-Fossa BA, De Cobelli O, Nolè F, Renne G, Bellomi M, Padhani AR, Petralia G. Whole-body magnetic resonance imaging (WB-MRI) reporting with the METastasis Reporting and Data System for Prostate Cancer (MET-RADS-P): inter-observer agreement between readers of different expertise levels. *Cancer Imaging*. 2020 Oct 27;20(1):77. doi: 10.1186/s40644-020-00350-x. PMID: 33109268; PMCID: PMC7590732.
- 94-Lai AYT, Riddell A, Barwick T, Boyd K, Rockall A, Kaiser M, Koh DM, Saffar H, Yusuf S, Messiou C. Interobserver agreement of whole-body magnetic resonance imaging is superior to whole-body computed tomography for assessing disease burden in patients with multiple myeloma. *Eur Radiol*. 2020 Jan;30(1):320-327. doi: 10.1007/s00330-019-06281-x. Epub 2019 Jul 2. PMID: 31267214; PMCID: PMC6890623
- 95-Lecouvet FE, Boyadzhiev D, Collette L, Berckmans M, Michoux N, Triqueneaux P, Pasoglou V, Jamar F, Vekemans MC. MRI versus ¹⁸F-FDG-PET/CT for detecting bone marrow involvement in multiple myeloma: diagnostic performance and clinical relevance. *Eur Radiol*. 2020 Apr;30(4):1927-1937. doi: 10.1007/s00330-019-06469-1. Epub 2019 Dec 16. PMID: 31844960.
- 96-Starekova J, Hernando D, Pickhardt PJ, Reeder SB. Quantification of Liver Fat Content with CT and MRI: State of the Art. *Radiology*. 2021 Nov;301(2):250-262. doi: 10.1148/radiol.2021204288. Epub 2021 Sep 21. PMID: 34546125; PMCID: PMC8574059.
- 97-Giles SL, Messiou C, Collins DJ, Morgan VA, Simpkin CJ, West S, Davies FE, Morgan GJ, deSouza NM. Whole-body diffusion-weighted MR imaging for assessment of treatment response in myeloma. *Radiology*. 2014 Jun;271(3):785-94. doi: 10.1148/radiol.13131529. Epub 2014 Jan 21. PMID: 24475858
- 98-Donners R, Candito A, Blackledge M, Rata M, Messiou C, Koh DM, Tunariu N. Repeatability of quantitative individual lesion and total disease multiparametric whole-body MRI measurements in prostate cancer bone metastases. *Br J Radiol*.

- 2023 Nov;96(1151):20230378. doi: 10.1259/bjr.20230378. Epub 2023 Sep 3.
PMID: 37660399; PMCID: PMC10607420
- 99- Levey AS, Titan SM, Powe NR, Coresh J, Inker LA. Kidney Disease, Race, and GFR Estimation. *Clin J Am Soc Nephrol*. United States; 2020;15(8):1203–1212. doi: 10.2215/CJN.12791019
- 100- Bhatt P, Kloock C, Comenzo R. Relapsed/Refractory Multiple Myeloma: A Review of Available Therapies and Clinical Scenarios Encountered in Myeloma Relapse. *Curr Oncol*. 2023 Feb 15;30(2):2322-2347. doi: 10.3390/curroncol30020179. PMID: 36826140; PMCID: PMC9954856
- 101- Rasche L, Kortüm KM, Raab MS, Weinhold N. The Impact of Tumor Heterogeneity on Diagnostics and Novel Therapeutic Strategies in Multiple Myeloma. *Int J Mol Sci*. 2019 Mar 12;20(5):1248. doi: 10.3390/ijms20051248. PMID: 30871078; PMCID: PMC6429294
- 102- Westerland O, Amlani A, Kelly-Morland C, Fraczek M, Bailey K, Gleeson M, El-Najjar I, Streetly M, Bassett P, Cook GJR, Goh V; Myeloma Imaging Research Group at Guy's & St Thomas' Hospital, London and King's College London. Comparison of the diagnostic performance and impact on management of 18F-FDG PET/CT and whole-body MRI in multiple myeloma. *Eur J Nucl Med Mol Imaging*. 2021 Jul;48(8):2558-2565. doi: 10.1007/s00259-020-05182-2. Epub 2021 Jan 19. PMID: 33469686; PMCID: PMC8241666.
- 103- Mesguich C, Hulin C, Latrabe V, Lascaux A, Bordenave L, Hindié E, Marit G. Prospective comparison of 18-FDG PET/CT and whole-body diffusion-weighted MRI in the assessment of multiple myeloma. *Ann Hematol*. 2020 Dec;99(12):2869-2880. doi: 10.1007/s00277-020-04265-2. Epub 2020 Sep 19. PMID: 32951093.
- 104- Sun M, Cheng J, Ren C, Zhang Y, Li Y, Wang L, Zhang S, Lin L. Evaluation of Diffuse Bone Marrow Infiltration Pattern in Monoclonal Plasma Cell Diseases by Quantitative Whole-body Magnetic Resonance Imaging. *Acad Radiol*. 2022 Apr;29(4):490-500. doi: 10.1016/j.acra.2021.06.015. Epub 2021 Aug 4. PMID: 34362664.
- 105- Zwanenburg A, Vallières M, Abdalah MA, Aerts HJWL, Andrearczyk V, Apte A, Ashrafinia S, Bakas S, Beukinga RJ, Boellaard R, Bogowicz M, Boldrini L, Buvat I, Cook GJR, Davatzikos C, Depeursinge A, Desseroit MC, Dinapoli N, Dinh CV, Echegaray S, El Naqa I, Fedorov AY, Gatta R, Gillies RJ,

Goh V, Götz M, Guckenberger M, Ha SM, Hatt M, Isensee F, Lambin P, Leger S, Leijenaar RTH, Lenkiewicz J, Lippert F, Losnegård A, Maier-Hein KH, Morin O, Müller H, Napel S, Nioche C, Orlhac F, Pati S, Pfaehler EAG, Rahmim A, Rao AUK, Scherer J, Siddique MM, Sijtsema NM, Socarras Fernandez J, Spezi E, Steenbakkers RJHM, Tanadini-Lang S, Thorwarth D, Troost EGC, Upadhaya T, Valentini V, van Dijk LV, van Griethuysen J, van Velden FHP, Whybra P, Richter C, Löck S. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. *Radiology*. 2020 May;295(2):328-338. doi: 10.1148/radiol.2020191145. Epub 2020 Mar 10. PMID: 32154773; PMCID: PMC7193906.

106- Wennmann M, Ming W, Bauer F, Chmelik J, Klein A, Uhlenbrock C, Grözinger M, Kahl KC, Nonnenmacher T, Debic M, Hielscher T, Thierjung H, Rotkopf LT, Stanczyk N, Sauer S, Jauch A, Götz M, Kurz FT, Schlamp K, Horger M, Afat S, Besemer B, Hoffmann M, Hoffend J, Kraemer D, Graeven U, Ringelstein A, Bonekamp D, Kleesiek J, Floca RO, Hillengass J, Mai EK, Weinhold N, Weber TF, Goldschmidt H, Schlemmer HP, Maier-Hein K, Delorme S, Neher P. Prediction of Bone Marrow Biopsy Results From MRI in Multiple Myeloma Patients Using Deep Learning and Radiomics. *Invest Radiol*. 2023 Oct 1;58(10):754-765. doi: 10.1097/RLI.0000000000000986. Epub 2023 May 22. PMID: 37222527.

107- Latifoltojar A, Hall-Craggs M, Bainbridge A, Rabin N, Popat R, Rismani A, D'Sa S, Dikaios N, Sokolska M, Antonelli M, Ourselin S, Yong K, Taylor SA, Halligan S, Punwani S. Whole-body MRI quantitative biomarkers are associated significantly with treatment response in patients with newly diagnosed symptomatic multiple myeloma following bortezomib induction. *Eur Radiol*. 2017 Dec;27(12):5325-5336. doi: 10.1007/s00330-017-4907-8. Epub 2017 Jun 27. PMID: 28656463; PMCID: PMC567412.