

Image interpretation criteria for FDG PET/CT in multiple myeloma: a new proposal from an Italian expert panel. IMPeTUs (Italian Myeloma criteria for PET USE)

Cristina Nanni¹ · Elena Zamagni² · Annibale Versari³ · Stephane Chauvie⁴ ·
Andrea Bianchi⁵ · Marco Rensi⁶ · Marilena Bellò⁷ · Ilaria Rambaldi¹ ·
Andrea Gallamini⁸ · Francesca Patriarca⁹ · Francesca Gay¹⁰ · Barbara Gamberi¹¹ ·
Michele Cavo² · Stefano Fanti¹

Received: 22 May 2015 / Accepted: 17 September 2015 / Published online: 16 October 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract

Purpose FDG PET/CT is able to detect active disease in patients with multiple myeloma (MM) and can be helpful for staging and assessing therapy response, but no standard interpretation criteria have been proposed for the evaluation of FDG PET/CT in MM.

Methods A group of Italian nuclear medicine physicians and haematologists met to propose new visual interpretation criteria to standardize FDG PET/CT evaluation in MM patients (Italian Myeloma criteria for PET USE; IMPeTUs) and

the reproducibility of these criteria was tested. This Italian multicentre protocol was set up as a subprotocol of EMN02, an international prospective multicentre trial of the European Myeloma Network. The criteria were agreed at multidisciplinary consensus meetings. They include a description of the metabolic state of the bone marrow (BM), number and site of focal PET-positive lesions, the number of osteolytic lesions, and the presence and site of extramedullary disease, paramedullary disease and fractures. A visual degree of uptake was defined for the target lesion and extramedullary lesions according to modified Deauville criteria. MM patients who had undergone FDG PET/CT at baseline (PET-0), after induction (PET-AI) and at the end of treatment (PET-EoT) were enrolled. The patients had been prospectively enrolled in EMN02 and their PET scans were a posteriori reinterpreted in a blinded independent central review process managed by WIDEN[®]. Five expert nuclear medicine physicians scored the scans according to the new criteria. A case was considered read when four out of the five reviewers completed the report. Concordance among reviewers on different metrics was calculated using Krippendorff's alpha coefficient.

Results A total of 17 consecutive patients were enrolled. On PET-0, the alpha coefficients for the BM score, the score for the hottest focal lesion, the number of focal lesions and the number of lytic lesions were 0.33 and 0.47, 0.40 and 0.32, respectively. On PET-AI, the alpha coefficients were 0.09 and 0.43, 0.22 and 0.21, respectively, and on PET-EoT, the alpha coefficients were 0.07, 0.28, 0.25 and 0.21, respectively. BM was generally difficult to score since grades 2 and 3 are difficult to discriminate. However, since neither of the two grades is related to BM myelomatous involvement, the difference was not

✉ Cristina Nanni
cnanni@aosp.bo.it

¹ Nuclear Medicine, AOU Policlinico S.Orsola-Malpighi, Bld 30, Via Massarenti, 9, 40138 Bologna, Italy

² Hematology, AOU Policlinico S.Orsola-Malpighi, Bologna, Italy

³ Nuclear Medicine, S. Maria Nuova Hospital, IRCSS, Reggio Emilia, Italy

⁴ Medical Physics Unit, Santa Croce e Carle Hospital, Cuneo, Italy

⁵ Nuclear Medicine, Santa Croce e Carle Hospital, Cuneo, Italy

⁶ Nuclear Medicine, AOU S.Maria della Misericordia, Udine, Italy

⁷ Nuclear Medicine, AO Città della Salute e della Scienza, Torino, Italy

⁸ Research and Innovation Department, A Lacassagne Cancer Center, Nice, France

⁹ Hematologic Clinic, Udine University, Udine, Italy

¹⁰ Myeloma Unit, Division of Hematology, University of Torino, Torino, Italy

¹¹ Hematology Unit, Azienda Ospedaliera ASMN, IRCCS, Reggio Emilia, Italy

clinically relevant. Agreement on focal lesion scores and on the number of focal lesions was good.

Conclusion The new visual criteria for interpreting FDG PET/CT imaging in MM patients, IMPeTUs, were found to be feasible in clinical practice.

Keywords Multiple myeloma · FDG PET/CT · Interpretation criteria · Standardization · IMPeTUs

Introduction

Multiple myeloma (MM) is a malignant disease characterized by the autonomous proliferation of monoclonal plasma cells in the bone marrow (BM) and by the overproduction of either intact immunoglobulin molecules (M-component or M-protein) or immunoglobulin free light chains kappa or lambda. MM is rare among individuals younger than 40 years, but its incidence rises in subsequent decades and exhibits a slight male predominance. Localized proliferation of malignant plasma cells in the marrow and bone may produce painful osteolytic lesions visible on plain films. Bone involvement has an important clinical impact (e.g. pathological fractures) leading to increasing pain and skeletal instability. Neurological problems are a cause of morbidity, since vertebral osteolytic lesions may produce spinal cord compression or radiculopathy. Correct assessment of BM infiltration, identification of bone lesions and extramedullary (EM) disease is crucial in the assessment of MM [1].

Recently, FDG PET/CT has been used to stage MM patients [2–4], to accurately evaluate response to therapy [5], to detect the site of EM disease [6] and impending relapse [5, 7]. In recent years very promising and concordant results have been reported by different groups. FDG PET/CT has been proved to be prognostic in staging, in interim and final treatment monitoring, and during follow-up [4, 7]. FDG PET/CT has some recognized advantages in the evaluation of MM patients:

1. Extended field of view (including skull, ribs, upper limbs, femurs, pelvis and spine)
2. Absence of possible collateral effects or adverse reactions to FDG
3. Possibility to perform it even in patients with renal failure
4. Short image acquisition time with 3-D tomographs (important in patients with fractures, bone pain or vertebral collapse)
5. Free decubitus positioning
6. Possibility to evaluate soft tissues and organs at the same time to detect EM disease
7. Possibility to semiquantify disease activity using the maximum standardized uptake value (SUV_{max})

8. The low-dose CT images associated with the PET images allow the morphological appearance of the bone to be described
9. No restrictions in relation to metallic bone implants

However, FDG PET/CT may be difficult to interpret in some patients:

1. A significant percentage of patients are affected by MM-related anaemia that results in a significant increase in BM tracer uptake causing a hot background in bone
2. The amount of FDG uptake is variable and therefore a baseline study is needed for reference
3. Early PET-positive MM lesions may not correspond to an osteolytic area and may be difficult to call
4. The low spatial resolution of PET imaging does not allow “salt and pepper” BM infiltration to be accurately detected
5. Recent fractures may appear falsely positive
6. Bone metallic implants may cause significant artefacts on CT images and may cause infections with resulting non-specific FDG uptake
7. Therapy response criteria are not defined

No standard interpretation criteria have been proposed for the evaluation of FDG PET/CT scans in MM. Some groups base their image interpretation mainly on semiquantitation, others rely on visual assessment and others on both methods, and thus data reproducibility is prevented [7, 8]. For these reasons, a group of Italian nuclear medicine experts, haematologists and medical physicists defined new visual interpretation criteria (Italian Myeloma criteria for Pet Use; IMPeTUs) to standardize FDG PET/CT evaluation in MM patients, and tested them in a preliminary blinded independent central review process. This was done to provide standard image interpretation criteria to make clinical trial results reproducible.

Materials and methods

Criteria designing process

Criteria were agreed during interdisciplinary meetings attended by nuclear medicine physicians, haematologists and medical physicists from different Italian institutions, who combined their experience in the use of FDG PET/CT in MM. A synthesis of views was reached and mainly descriptive criteria were defined. Semiquantitative reporting was excluded owing to high variability in the measurement of SUVs on different tomographs in clinical practice, especially in PET centres not undergoing any clinical trial qualification process. However, semiquantitative data were measured and collected for future analysis. All PET/CT scanners used in patients in

this investigation underwent the clinical trial qualification endorsed by FIL (Fondazione Italiana Linfomi; Italian Lymphoma Foundation) in collaboration with the Italian Association of Nuclear Medicine (AIMN) and the Italian Association of Medical Physics (AIFM) in the Corelab at the Medical Physics Unit of Santa Croce e Carle Hospital, Cuneo, Italy. Variability among PET/CT scanners in the measurement of SUVs in anthropomorphic phantoms is guaranteed to be below 10 %.

Image acquisition protocol

PET/CT scans were acquired according to a local protocol (applying EANM PET procedure guidelines for FDG studies), but the following conditions were required for inclusion of scan data: (a) studies had to be carried out using full-ring PET/CT; (b) iterative reconstruction was applied to PET images; (c) CT and attenuation-corrected PET images at baseline (PET-0), after induction therapy (PET-AI) and the end of treatment (PET-EoT) were available for central review; (d) PET scans of the same patient had to be performed on the same scanner, and (e) both PET and CT scans had to cover the region from the top of the skull to the lower third of the femurs.

Administered activity was 279 ± 68 MBq on average. The uptake time was uniformly in the range 58 – 88 min. The average liver SUV was 2.3 ± 0.5 . The differences in parameter values among PET scans at different time points were never significant ($p > 0.60$).

Image revision process

After anonymization, PET/CT scans were uploaded by the participating PET centres into WIDEN[®] (diXit srl, Torino, Italy). PET scans were excluded from the study if: (a) they were poor-quality images with low statistics that were not considered suitable for diagnostic interpretation, (b) the image dataset was incomplete, and (c) large violations with respect to uptake time and administered dose were found analysing DICOM headers of PET scans. Upon image upload, the WIDEN platform automatically distributed images to five expert reviewers (M.B., A.B., C.N., M.R. and A.V.) who downloaded the PET images onto their own workstation and independently reviewed the scans by completing the online form, patient by patient.

The review was considered complete when four of the five reviewers fulfilled the criteria. A two-step process for the reviews was set up. A preliminary set of 30 scans (belonging to patients not enrolled in the protocol) was distributed to the experts who reviewed them blindly using their own workstation as a preliminary test of the criteria. The results were compared by calculating the interobserver variabilities of all the parameters. The reviewers then met in a consensus review

session to report jointly the cases that showed discrepant results during the previous independent review. When needed, the criteria were adjusted. A new set of rules was established. This new set of rules was used to review blindly a new set of patients enrolled in the protocol. The main change included simplification in the number of focal lesions described, that were merged into the following groups: 0 lesions, 1 – 3 lesions, 4 – 10 lesions, >10 lesions.

IMPETUs criteria

The final version of the criteria is only descriptive and is based on five-point scales. This is important, since features correlated with the presence of active disease can be recognized “a posteriori” in the light of follow-up data; this process resembles that used in validation of the Deauville criteria. The final version includes the description, using a five-point scales, of: metabolic state of the BM, number and site of focal PET-positive lesions with or without osteolytic characteristics, presence and site of EM disease, presence of paramedullary (PM) disease, and presence of fractures. The visual degree of uptake is defined for the target lesion and EM lesions according to the schema proposed in the Deauville criteria for the evaluation of lymphoma patients [8] (Fig. 1).

The IMPETUs criteria are summarized in Table 1, and in detail are as follows:

Bone marrow (BM):

Deauville criteria:

- 1 No uptake at all
- 2 \leq mediastinal blood pool uptake (SUV_{max})
- 3 $>$ mediastinal blood pool uptake, \leq liver uptake
- 4 $>$ liver uptake +10 %
- 5 \gg liver uptake (twice)
“A” appended if there is hypermetabolism in limbs and ribs

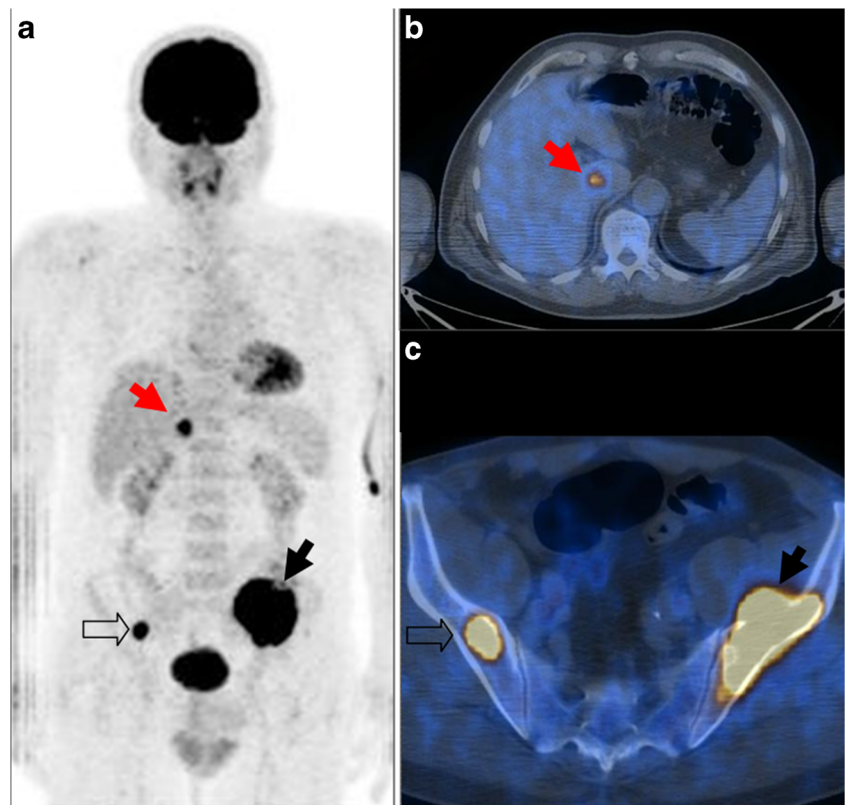
Focal bone lesions (F):

Lesion number group (x):

- x = 1: no lesions
- x = 2: 1 to 3 lesions
- x = 3: 4 to 10 lesions
- x = 4: >10 lesions
- S: skull
- Sp: spine
- ExtraSp: all the rest
- Target lesion is the hottest area

Plus Deauville criteria:

Fig. 1 FDG PET/CT imaging in a patient with MM. **a** MIP image, **b-c** transaxial fused images (*red arrows* extramedullary lesion (liver), *filled and unfilled arrows* extraspinal bone lesions). In this patient the descriptive criteria (IMPETUs) are: BM3, F2 ExtraSp (5), L, EM EN (5), where BM3 indicates bone marrow uptake is < liver but > mediastinum, F2 indicates one to three lesions, ExtraSp indicates outside the spine with (5) indicating reference lesion uptake >> liver, L indicates at least one lesion is also lytic, EM indicates at least one extramedullary lesion, EN indicates the extramedullary lesion is extranodal (liver) with (5) indicating extramedullary lesion uptake >> liver



- 1 No uptake at all
- 2 ≤ mediastinal blood pool uptake (SUV_{max})
- 3 > mediastinal blood pool uptake, ≤ liver uptake
- 4 > liver uptake +10 %
- 5 >> liver uptake (twice or more)

- x = 1: no lesions
- x = 2: 1 to 3 lesions
- x = 3: 4 to 10 lesions
- x = 4: >10 lesions

Presence of at least one lytic lesion (L):

Presence of at least one fracture on CT images (Fr)

Table 1 Summary of IMPETUs criteria

Lesion type	Site	Number of lesions (x)	Grading
Diffuse	Bone marrow ^a		Deauville five-point scale
Focal (F)	Skull (S)	x = 1 (no lesions)	Deauville five-point scale
	Spine (SP)	x = 2 (1 to 3 lesions)	
	Extraspinal (ExP)	x = 3 (4 to 10 lesions)	
		x = 4 (>10 lesions)	
Lytic (L)		x = 1 (no lesions)	
		x = 2 (1 to 3 lesions)	
		x = 3 (4 to 10 lesions)	
		x = 4 (>10 lesions)	
Fracture (Fr)	At least one		
Paramedullary (PM)	At least one		
Extramedullary (EM)	At least one	N/EN (nodal/extranodal) ^b	Deauville five-point scale

^a“A” if hypermetabolism in limbs and ribs

^bFor nodal disease (N): C cervical, SC supraclavicular, M mediastinal, Ax axillary, Rp retroperitoneal, Mes mesenteric, In inguinal; For extranodal disease (EN): Li liver, Mus muscle, Spl spleen, Sk skin, Oth other

Presence of paramedullary disease (PM): a bone lesion involving surrounding soft tissues with bone cortical interruption

Extramedullary disease (EM):

Nodal disease (N) plus site:

LC: laterocervical

SC: supraclavicular

M: mediastinal

Ax: axillary

Rp: retroperitoneal

Mes: mesentery

In: inguinal

Extranodal disease (EN) plus site:

Li: liver

Mus: muscle

Spl: spleen

Sk: skin

Oth: other

Plus Deauville criteria for target EM lesions:

- 1 No uptake at all
- 2 \leq mediastinal blood pool uptake (SUV_{max})
- 3 $>$ mediastinal blood pool uptake, \leq liver uptake
- 4 $>$ liver uptake +10 %
- 5 \gg liver uptake (twice or more)

Semiquantitative data for physiological areas are also added. The semiquantitative indexes are measured using a region of interest (ROI) with a radius greater than 3 cm in the central portion of the liver far away from its edge and a ROI completely encompassed in the lumen of the aorta, taking care to avoid the edge of the vessel wall or areas of calcification, for the mediastinal blood pool structures. The following semiquantitative parameters are annotated:

- BM SUV_{max} of the hottest lesion per macro area
- Focal reference SUV_{max} of the hottest lesion per macro area
- Focal reference SUV_{mean} of the hottest lesion (five pixels) per macro area
- Fr on REF: present/absent
- REF in EM disease: present/absent
- Liver SUV_{max}
- Liver SUV_{mean}
- Mediastinal blood pool SUV_{max}
- Mediastinal blood pool SUV_{mean} :
- Comments: ...

Statistics

Krippendorff's alpha was used for measurement of overall consensus agreement [9]. This coefficient is 0 for random coincidences and below 0 for concordances lower than random coincidence.

Results

The first exploratory PET/CT series included 30 scans performed in ten patients. This dataset was used for a preliminary check of the applicability of the criteria and was therefore carried out in patients not enrolled in the multicentre protocol. Overall a good concordance was found for BM evaluation. A low concordance rate among the reviewers was found for focal lesion number, focal lesion score, presence of fracture, presence of lysis, and presence of PM and EM disease (Table 2).

After the criteria were modified on the basis of the findings in the first ten patients (in particular categorizing the number of lesions into different groups), another 17 consecutive patients were enrolled within the Italian multicentre protocol. They had all undergone a baseline PET scan (PET-0), a PET scan after induction therapy (PET-AI) and a PET scan at the end of therapy (PET-EoT). The patient characteristics are presented in Table 3.

PET-0

Concordance in the evaluation of BM was 71.5 %. Alpha values for BM score and focal score were 0.33 and 0.47, respectively. Alpha values for number of focal lesions and lytic lesions were 0.40 and 0.32. Skull, spine and extraspinal involvement was found in 2, 14 and 12 patients, respectively, with mean concordances of 37 %, 78 % and 71 %. Four patients had a fracture by one reviewer and one patient had a fracture by two reviewers. Two patients had a PM lesion by one reviewer and two patients had a PM lesion by all reviewers. Four patients had nodal and extranodal EM lesions.

PET-AI

Concordance in the evaluation of BM was 64 %. Alpha values for BM score and focal score were 0.09 and 0.43, respectively. Alpha values for number of focal lesions and lytic lesions were 0.22 and 0.21, respectively. Skull, spine and extraspinal involvement was found in 1, 13 and 5 patients with a mean concordance of 25 %, 67 % and 79 %. Four patients had a fracture by one reviewer and one patient had a fracture by all reviewers. Four patients had a PM lesion by one reviewer and two patients had a PM lesion by all reviewers. No patients had EM disease.

Table 2 Percentage agreement on the first reading (ten patients, 30 scans)

PET scan	Agreement (%)						
	Bone marrow	Focal score	Number of focal lesions	Presence of fracture	Presence of bone lysis	Presence of paramedullary disease	Presence of extramedullary disease ^a
Staging	71.5	80	40	20	40	40	–
After induction therapy	64	68.3	40	50	40	30	–
End of therapy	68	46.7	73.3	70	73	40	–

^a Too few to calculate agreement

PET-EoT

Concordance in the evaluation of BM was 68 %. Alpha values for BM score and focal score were 0.07 and 0.28, respectively. Alpha values for number of focal lesions and lytic lesions were 0.25 and 0.21, respectively. Skull, spine and extra-spinal involvement was found in 1, 9 and 5 patients with a mean concordance of 25 %, 47 % and 70 %. Five patients had a fracture by one reviewer. No patients had PM or EM disease (Table 4).

Results regarding the agreement among the reviewers as to the presence/absence of bone lesions are presented in Table 5.

Discussion

In clinical practice, oncohaematologists rely on imaging to obtain prognostic data, to stratify patients into risk groups

Table 3 Patient characteristics

Characteristic	Value
No. of patients	17
Age (years), median (range)	59 (32 – 79)
Creatinine	
Serum concentration (mg/dL), median (range)	1.1 (0.5 – 10)
Patients with >2 mg/dL (%)	11
Calcium	
Serum concentration (mg/dL), median (range)	9.6 (2.5 – 14.9)
Patients with >10 mg/dL (%)	14
Patients with stage II/III (%) ^a	53
Patients with del(13q) (%)	43
Patients with del(17p) (%)	12
Patients with t(4;14) (%)	18
C-reactive protein (mg/L), median (range)	4 (0 – 39)
Lactate dehydrogenase (IU/L), median (range)	297 (103 – 2325)
Autologous stem cell transplantation (%)	
Single	35
Double	65

^a International staging system, ISS

and to identify the best and personalized treatment for each patient. FDG PET/CT is a sensitive but rather nonspecific diagnostic method in oncology. Moreover, standard and shared interpretation criteria to ensure data reproducibility are still lacking. Some nuclear medicine physicians, for example, tend to report only findings with a very high positive predictive value (specificity higher than sensitivity), others tend to report even tiny and faint findings (sensitivity higher than specificity), and others may change their reporting approach depending on the disease under evaluation, clinical history, patient age, risk of relapse etc. Consequently, the same set of images with equivocal findings may result in a positive report, inconclusive report or negative report according to the approach of the nuclear medicine physician, which is usually based on his or her own practical experience and relationship with referring clinicians. This is why a standardized interpretation algorithm is of basic importance. Understanding which finding is relevant, which is collateral, which is prognostic and which remains equivocal is the key to speaking the same language in both clinical practice and scientific clinical research.

At about the turn of the millennium, imaging generally played a marginal role in MM. Only skeletal plain radiography was included in staging systems, but was unsuitable for assessing response to therapy [10]. More recently, MRI was included in staging procedures because of its exceptional sensitivity in the detection of tumour spread, especially in the spine [11–14]. However, MRI is also insensitive for the interim assessment of therapy response [15]. The use of PET in MM started relatively late in comparison to PET in lymphoma. Several reasons account for this: (1) the presence in MM of a reliable, readily available and cheap marker for diagnosis and treatment monitoring, the M-protein; (2) low availability of PET scanning facilities in multicentre settings for a disease associated with serious morbidity in more than 30 % of patients at disease onset; (3) lack of standard criteria for PET scan interpretation. MM is a very complicated disease from an imaging point of view. To increase sensitivity of PET the field of view must be extended to also include the long bones and the skull, which are involved by disease in a significant percentage of patients. Moreover, the skeleton can be affected with a diffuse/infiltrative pattern or with focal lesions. Finally,

Table 4 Krippendorff's alpha values in 17 patients for agreement among four reviewers

PET scan	Krippendorff's alpha				
	Bone marrow	Focal score	Number of focal lesions	Presence of bone lysis	Presence of extramedullary disease ^a
Staging	0.33	0.47	0.40	0.32	–
Post Induction Therapy	0.09	0.43	0.22	0.21	–
End of therapy	0.07	0.28	0.25	0.21	–

^a Too few to calculate agreement

extraskelatal lesions may be present. Concomitant myeloma-related conditions (including anaemia and fractures) are common and complicate image interpretation.

In recent years several studies have focused on FDG PET/CT for MM staging, restaging and treatment monitoring. Many studies have provided interesting preliminary results on the accuracy and predictive value of PET/CT, and its value in the early assessment of therapy response [4, 7, 16]. However, the interpretation criteria used in these studies are different. In some purely visual criteria were used, in some semi-quantitative criteria and in some mixed criteria. Such criteria are rather concordant in providing a clear definition of a positive or a negative lesion, but may be strongly discordant if the findings are equivocal such as in the presence of mild focal bone uptake (especially during or after therapy), leading to variable performance of FDG PET/CT [4, 7, 16].

In consideration of the increasing use of FDG PET/CT and of upcoming international multicentre trials on PET/CT in MM, a group of Italian nuclear medicine physicians, haematologists and medical physicists proposed new standardized interpretation criteria for PET/CT image reading. Moving from the positive experience with PET in lymphoma, it was decided to adopt interpretation rules similar to the Deauville criteria based on visual assessment [17], including a semiquantitative analysis (SUV_{max} of lesions and normal structures) for clarification of eventual equivocal results within the limitations of quantitative assessment in the multicentre setting. A descriptive readout was adopted for the proposed criteria and the follow-up criterion (e.g. the disappearance of the abnormality at the end of treatment)

Table 5 Agreement among the reviewers as to the presence/absence of bone lesions

PET scan	No. of patients with concordant reviewers		
	Skull	Spine	Extraspinal
Staging	2/37	14/78	12/71
Post induction therapy	1/25	13/67	5/79
End of therapy	1/25	9/47	5/70

was used to define the described finding as attributable to disease. This method for setting up new diagnostic criteria is longer than a standard validation of predefined a priori criteria, but is guaranteed to be more applicable and reproducible in use. Moreover, it allows nonspecific FDG uptake patterns that may not have a clinical impact or meaning to be ruled out (for example, increased BM activity or mild residual uptake in lytic lesions after therapy) but may be difficult to interpret. Accordingly, each abnormality is described but its clinical relevance is assessed a posteriori.

Nonetheless, the agreement among the reviewers seems at a first glance to be suboptimal, even though there was a certain improvement from the first version of the criteria to the final version that was simplified particularly in terms of the number of the categorized lesions. This is because statistical analysis carried out with only one coefficient takes into account perfect concordance of different degrees of positivity, and is not only a dichotomous judgment in terms of positivity/negativity as reported in previous studies on Deauville criteria in lymphoma patients [8]. Finally, the agreement among the reviewers was comparable to that obtained following the first revision in previous studies on lymphoma [8, 18, 19]. In this perspective a consensual revision of the nonconcordant cases was very helpful for clarifying the most critical points, and the overall set of rules awaits prospective validation in a larger cohort of patients.

As expected, a lower concordance rate was found in the detection of skull lesions and of lesions active after therapy. The skull is a problematic bone area in MM. Lesions are usually very small there, and their detectability is impaired by the physiological high uptake of the underlying brain. After therapy the detectability of lesions strongly depends on the background activity of “normal” BM, and many lesions with minimal residual uptake may be hardly detectable in patients with reactive BM.

The current set of interpretation criteria will be adopted in a forthcoming study including a larger series of patients, which, according to the planned protocol, is expected to reach a sample size of 100 to 130 patients. Subsequently, in the light of follow-up data, a visual cut-off for positivity will finally be defined and provided.

Conclusion

This preliminary study demonstrated that these new, easily applied and descriptive criteria for FDG PET/CT in MM seem to be sufficiently reproducible and may become the shared basis for clinical international multicentre protocols.

Compliance with Ethical Standards

Conflicts of interest S.C. is cofounder of diXit srl, a spin-off from the University of Torino and INFN (Italian Institute of Nuclear Physics). The other authors have no conflicts of interest to declare.

Ethical approval All procedures performed in this study involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the principles of the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards. This Italian multicentre protocol was set up as a subprotocol of EMN02, an international prospective multicentre trial of the European Myeloma Network, and was coordinated by the imaging committee of GIMEMA (Gruppo Italiano Malattie Ematologiche dell'Adulto). The ethics committees of each institution involved in the project approved it.

Informed consent All the enrolled patients signed a specific informed consent.

References

- Kyle RA, Rajkumar SV. Criteria for diagnosis, staging, risk stratification and response assessment of multiple myeloma. *Leukemia*. 2009;23(1):3–9.
- Nanni C, Zamagni E, Farsad M, Castellucci P, Tosi P, Cangini D, et al. Role of 18F-FDG PET/CT in the assessment of bone involvement in newly diagnosed multiple myeloma: preliminary results. *Eur J Nucl Med Mol Imaging*. 2006;33(5):525–31.
- Zamagni E, Nanni C, Patriarca F, Englaro E, Castellucci P, Geatti O, et al. A prospective comparison of 18F-fluorodeoxyglucose positron emission tomography-computed tomography, magnetic resonance imaging and whole-body planar radiographs in the assessment of bone disease in newly diagnosed multiple myeloma. *Haematologica*. 2007;92(1):50–5.
- Zamagni E, Patriarca F, Nanni C, Zannetti B, Englaro E, Pezzi A, et al. Prognostic relevance of 18-F FDG PET/CT in newly diagnosed multiple myeloma patients treated with up-front autologous transplantation. *Blood*. 2011;118(23):5989–95.
- Nanni C, Zamagni E, Celli M, Caroli P, Ambrosini V, Tacchetti P, et al. The value of 18F-FDG PET/CT after autologous stem cell transplantation (ASCT) in patients affected by multiple myeloma (MM): experience with 77 patients. *Clin Nucl Med*. 2013;38(2):e74–9.
- Zamagni E, Nanni C, Tacchetti P, Pantani L, Marzocchi G, Zannetti B, et al. Positron emission tomography with computed tomography-based diagnosis of massive extramedullary progression in a patient with high-risk multiple myeloma. *Clin Lymphoma Myeloma Leuk*. 2014;14(3):e101–4.
- Bartel TB, Haessler J, Brown TLY, Shaughnessy JDJ, van Rhee F, Anaissie E, et al. F18-fluorodeoxyglucose positron emission tomography in the context of other imaging techniques and prognostic factors in multiple myeloma. *Blood*. 2009;114(10):2068–76.
- Derlin T, Peldschus K, Munster S, Bannas P, Herrmann J, Stubig T, et al. Comparative diagnostic performance of 18F-FDG PET/CT versus whole-body MRI for determination of remission status in multiple myeloma after stem cell transplantation. *Eur Radiol*. 2013;23(2):570–8.
- Hayes AF, Krippendorff K. Answering the call for a standard reliability measure for coding data. *Commun Method Meas*. 2007;1:77–89.
- Durie BG, Kyle RA, Belch A, Bensinger W, Blade J, Boccadoro M, et al. Myeloma management guidelines: a consensus report from the Scientific Advisors of the International Myeloma Foundation. *Hematol J*. 2003;4(6):379–98.
- Durie BG. The role of anatomic and functional staging in myeloma: description of Durie/Salmon plus staging system. *Eur J Cancer*. 2006;42(11):1539–43.
- Durie BG, Waxman AD, D'Agnolo A, Williams CM. Whole-body (18)F-FDG PET identifies high-risk myeloma. *J Nucl Med*. 2002;43(11):1457–63.
- Bauerle T, Hillengass J, Fechtner K, Zechmann CM, Grenacher L, Moehler TM, et al. Multiple myeloma and monoclonal gammopathy of undetermined significance: importance of whole-body versus spinal MR imaging. *Radiology*. 2009;252(2):477–85.
- Baur A, Stabler A, Nagel D, Lamerz R, Bartl R, Hiller E, et al. Magnetic resonance imaging as a supplement for the clinical staging system of Durie and Salmon? *Cancer*. 2002;95(6):1334–45.
- Baur-Melnyk A, Buhmann S, Becker C, Schoenberg SO, Lang N, Bartl R, et al. Whole-body MRI versus whole-body MDCT for staging of multiple myeloma. *AJR Am J Roentgenol*. 2008;190(4):1097–104.
- Bannas P, Hentschel HB, Bley TA, Treszl A, Eulenburg C, Derlin T, et al. Diagnostic performance of whole-body MRI for the detection of persistent or relapsing disease in multiple myeloma after stem cell transplantation. *Eur Radiol*. 2012;22(9):2007–12.
- Meignan M, Gallamini A, Meignan M, Gallamini A, Haioun C. Report on the First International Workshop on Interim-PET-Scan in Lymphoma. *Leuk Lymphoma*. 2009;50(8):1257–60.
- Meignan M, Gallamini A, Haioun C, Polliack A. Report on the Second International Workshop on interim positron emission tomography in lymphoma held in Menton, France, 8-9 April 2010. *Leuk Lymphoma*. 2010;51(12):2171–80.
- Chauvie S, Biggi A, Stancu A, Cerello P, Cavallo A, Fallanca F, et al. WIDEN: a tool for medical image management in multicenter clinical trials. *Clin Trials*. 2014;11(3):355–61. Originally published online 7 April 2014.