Single-point and kinetics of peripheral residual disease by mass spectrometry to predict outcome in patients with high-risk smoldering multiple myeloma included in the GEM-CESAR trial

Noemí Puig,[1*] Cristina Agulló,[2*] Teresa Contreras,[2] José-Juan Pérez,[1] Irene Aires,[1] María-José Calasanz,[3] Ramón García-Sanz,[1] Sergio Castro,[1] Joaquín Martínez-López,[4] Paula Rodríguez-Otero,[3] Verónica González-Calle,[1] Marta S. González,[5] Albert Oriol,[6] Norma C. Gutiérrez,[1] Rafael Ríos-Tamayo,[7] Laura Rosiñol,[8] Miguel-Ángel Álvarez,[9] Joan Bargay,[10] Ana-Pilar González-Rodríguez,[11] Adrián Alegre,[12] Fernando Escalante,[13] María-Belén Iñigo,[14] Javier de la Rubia,[15] Ana-Isabel Teruel,[16] Felipe de Arriba,[17] Luis Palomera,[18] Miguel T. Hernández,[19] Javier López-Jiménez,[20] Marta Reinoso,[21] Aránzazu García-Mateo,[22] Enrique M. Ocio,[23] Joan Bladé,[8] Juan-José Lahuerta,[24] María-Teresa Cedena,[25] Bruno Paiva,[3] Jesús F. San Miguel[3] and María-Victoria Mateos[1]

[1]Hematology Department, Hospital Universitario de Salamanca-IBSAL, CIBERONC and Centro de Investigación del Cáncer, IBMCC (USAL-CSIC), Salamanca; [2]Biochemistry Department, University Hospital of Salamanca, Salamanca; [3]Hematology Department, Cancer Center Clínica Universidad de Navarra, Centro de Investigación Médica Aplicada (CIMA), IDISNA, CIBERONC, Pamplona; [4]Hematology Department, Hospital 12 de Octubre, Medicine Department, Medicine School of Complutense University (i+12) / Spanish National Cancer Research Center (CNIO), Madrid; [5]Hematology Department, Hospital Clínico Universitario Santiago de Compostela, Santiago de Compostela; [6]Clinical Hematology, Institut Català d'Oncologia and Josep Carreras Research Institute, Hospital Germans Trias I Pujol, Badalona; [7]Unidad de Gammapatías Monoclonales, Hospital Universitario Puerta de Hierro, Majadahonda; [8]Hematology Department, Hospital Clinic, IDIBARS, Barcelona; [9]Hematology Department, Hospital Universitario Reina Sofía, Córdoba; [10]Hematology Department, Hospital Universitario Son Llatzer, IdISBa (Institut d'Investigació Sanitaria Illes Balears), Palma; [11]Hematology Department, Hospital Universitario Central de Asturias, Oviedo; [12]University Hospital La Princesa & University Hospital QuironSalud, Autónoma University, Madrid; [13]Servicio de Hematología, Unidad i+i, Complejo Asistencial Universitario de León, León; [14]Hematology Department, Hospital Clínico San Carlos, Madrid; [15]Hematology Department, Hospital Universitario y Politécnico La Fe, Universidad Católica de Valencia, CIBERONC, Valencia; [16]Hematology Department, Hospital Clínico Universitario de Valencia, Valencia; [17]Hematology Department, Hospital Universitario Morales Meseguer, IMIB-Pascual; [18]Hematology Department, Hospital Clínico Universitario Lozano Blesa, Instituto de Investigación Sanitaria de Aragón, Zaragoza; [19]Hematology Department, Hospital Universitario de Canarias, Santa Cruz de Tenerife; [20]Hematology Department, Hospital Ramón y Cajal, Madrid; [21]Hematology Department, Hospital Universitario Virgen del Rocío, Instituto de Biomedicina de Sevilla (IBIS) / CSIC, Universidad de Sevilla, Sevilla; [22]Hematology Department, Complejo Asistencial de Segovia, Segovia; [23]Hematology Department, Hospital Universitario Marqués de Valdecilla (IDIVAL), Universidad de Cantabria, Santander; [24]Instituto de Investigación, Hospital Universitario 12 de Octubre, Madrid and [25]Hematology Department, Hospital Universitario 12 de Octubre, Instituto de Investigación (i+12), Madrid, Spain

**NP and CA contributed equally as first authors.*

Correspondence: M.-V. Mateos
mvmateos@usal.es

Received: April 23, 2024.
Accepted: July 3, 2024.
Early view: July 11, 2024.

https://doi.org/10.3324/haematol.2024.285742

Abstract

The value of quantitative immunoprecipitation mass spectrometry (QIP-MS) to identify the M-protein is being investigated in patients with monoclonal gammopathies but no data are yet available in high-risk smoldering myeloma (HRsMM). We have, therefore, investigated QIP-MS to monitor peripheral residual disease (PRD) in 62 HRsMM patients enrolled in the GEM-CESAR trial. After 24 cycles of maintenance, detecting the M-protein by MS or clonal plasma cells by next-generation flow cytometry (NGF) identified cases with a significantly shorter median progression-free survival (mPFS) (MS: not reached *vs.* 1.4 years, *P*=0.001; NGF: not reached *vs.* 2 years, *P*=0.0002) but reaching complete response (CR) + stringent CR (sCR) did not discriminate between patients with different outcome. With NGF as a reference, the combined results of NGF and MS showed a high negative predictive value (NPV) of MS: 81% overall and 73% at treatment completion. When sequential results were considered, sustained negativity by MS or NGF was associated with a very favorable outcome with an mPFS not yet reached *versus* 1.66 years and 2.18 years in cases never attaining PRD or minimal residual disease (MRD) negativity, respectively. We can, thus, conclude that: 1) the standard response categories of the International Myeloma Working Group do not seem to be useful for monitoring treatment in HRsMM patients; 2) MS could be used as a valuable, non-invasive, clinical tool with the capacity of guiding timely bone marrow evaluations (based on its high NPV with NGF as a reference); and 3) similarly to NGF, sequential results of MS are able to identify a subgroup of HRsMM patients with long-term disease control. This study was registered at www.clinicaltrials.gov (*clinicaltrials.gov identifier: 02415413*).

Introduction

The prognosis of patients with multiple myeloma (MM) achieving optimal responses to therapy is known to be heterogeneous, and different progression-free survival (PFS) and overall survival (OS) rates are observed among cases in complete response or better (CR+stringent CR [sCR]).[1] This has been explained in numerous studies by the persistence of minimal residual disease (MRD) detected in the bone marrow (BM) of some of these patients, which is responsible for the observed, unexpected progressions.[1] Acknowledging this situation, in 2016, the International Myeloma Working Group (IMWG) included guidelines for the definition of MRD negativity as a category of response, recommending the evaluation of a BM sample either with next-generation sequencing (NGS) or with next-generation flow cytometry (NGF) with a minimum sensitivity level of 10^{-5} in all patients achieving at least a CR.[2] A large body of data supports the clinical value of detecting residual disease in the BM using NGS and NGF in patients with MM.[3] In fact, attaining MRD negativity is currently selected as the main end-point in the majority of clinical trials. However, the use of BM samples to evaluate MRD requires an invasive procedure that cannot be performed very frequently and that could also limit the accuracy of the results. Although a logical alternative to the BM could be to use peripheral blood (PB), some correlative studies have revealed that 10^{-5} is insufficient for the results obtained in BM and PB to be comparable. In fact, using NGF in paired BM and PB samples from 137 newly diagnosed MM patients after active treatment, Sanoja-Flores *et al.* showed that in 55 out of the 91 (60%) patients with detectable MRD in the BM, circulating tumor plasma cells could not be identified.[4] Similarly, among the 28 MRD^+ follow-up samples detected by NGS by Mazzoti *et al.*, 18 (64%) did not present detectable circulating tumor DNA.[5] For this reason, highly sensitive approaches are being investigated for the detection of residual disease in PB, including mass spectrometry (MS). MS can identify the presence of M-protein in serum based on the specific aminoacidic sequence of each patient paraprotein, and thus its unique m/z ratio.[6] In fact, we recently reported that, compared to immunofixation electrophoresis (IFE), Immunoglobulin Isotypes (GAM) assay for the MS EXENT® analyzer is more sensitive to detect the M-protein in the serum of MM patients, both at baseline and during treatment monitoring, and that, most importantly, MS is more accurate than IFE to predict patients‘ outcome.[7] Different MS-based technologies are under development, although due to their high throughput, those based on matrix-assisted laser desorption ionization time of flight MS (MALDI-TOF-MS) are the most suitable for routine clinical practice.[6,8,9]

In the context of the GEM-CESAR trial for HRsMM patients, we have evaluated the presence of residual disease in serum samples by MS to ascertain its value as a complement to MRD that could overcome the limitations of BM-based disease evaluation.

Methods

Sixty-two eligible patients were included in this study. The diagnosis of SMM was based on the 2010 IMWG criteria and the high risk of progression was established according to either the Mayo 2008 and/or the PETHEMA models.[10-12] Since the trial was designed in 2013, ultra-high-risk SMM cases were also included. The GEM-CESAR trial included a total of 90 patients.[13] GEM-CESAR is a non-randomized, open-label, multicenter phase II trial in which patients

received induction with six 4-week cycles of KRd (carfilzomib, lenalidomide and dexamethasone), high-dose melphalan followed by autologous stem cell transplant (ASCT), consolidation with two further cycles of KRd and up to two years of maintenance with lenalidomide and dexamethasone (Rd). Patients presenting with biochemical progression were offered rescue therapy with daratumumab, pomalidomide and dexamethasone. Each study site's independent ethics committee approved the protocol, and informed consent forms were required prior to patient enrollment. The study was conducted according to the principles of the Declaration of Helsinki.

Next-generation flow cytometry

Minimal residual disease was assessed using the NGF method developed by EuroFlow for highly sensitive and standardized MRD detection in MM after induction, three months after ASCT and at the end of the treatment phase, after completing maintenance.[14] MRD studies were centralized in the 3 laboratories of the Spanish Myeloma Group.

Quantitative immunoprecipitation mass spectrometry

Serum samples were analyzed using Immunoglobulin Isotypes (GAM) assay for the MS EXENT® analyzer. First, the EXENT-iP500 liquid handler purified the immunoglobulins through paramagnetic beads coated with polyclonal sheep antibodies specific for human IgG, IgA or IgM heavy chains, and for total kappa and lambda light chains. Samples were then washed, eluted (20 mM TCEP in 5%(v/v) acetic acid) and spotted onto MALDI plates with HCCA matrix. Subsequently, analysis with the EXENT-iX500 MALDI-TOF device was carried out, and mass spectra from 5000 to 32000 mass to charge ratio (m/z) were collected. The +2 charge state was used for the interpretation of spectra by the EXENT-iQ software. The m/z of the M-protein identified in baseline samples was used as a patient-specific tumor marker in the subsequent samples.

End-points and assessments of the trial

The primary end-point of the trial was MRD negativity rate three months after-ASCT; secondary end-points included standard response rates and sustained MRD negativity rate at three, four and five years after-ASCT. We have assessed and compared the results and clinical value in terms of PFS of SPEP/IFE, NGF and MS to identify residual disease (the M-protein in serum samples by SPEP/IFE and MS and clonal plasma cells in bone samples by NGF) three months after ASCT and after two years of maintenance, at single time-points and considering the kinetics of the results.

Statistical analyses

GraphPad Prims v.9 was used for all statistical analyses. Differences in sensitivity were tested by Fishers' exact test. Confidence intervals were determined using exact binomial distribution. PFS was defined as the time since inclusion in the trial to the development to biochemical progression as defined by biochemical relapse/progressive disease according to the IMWG criteria; under this term we also included ultrasensitive MRD relapse defined by reappearance of MRD confirmed at least two months apart. Curves were constructed using the Kaplan-Meier method and two-sided, log-rank test. Dynamics of residual disease were analyzed from post-induction to treatment completion.

Results

Patients' characteristics

Baseline demographic and disease characteristics are shown in Table 1. Out of the 90 patients included in the trial, we had serum samples available for further analyses in 62 (68.8%), 61 (67.7%), 61 (67.7%), 51 (57%), and 35 (38.8%) of them at the time-points of post-induction, ASCT, consolidation, after one year of maintenance and at treatment completion, respectively. We focused our analysis in these groups of samples (and patients). Besides progression, some samples were missed due to the problems in dealing with the Covid-19 pandemic.

Table 1. Baseline demographic and disease characteristics of all study participants.

Characteristics	Results
Age in years, median (range)	59 (40-70)
Male, N (%)	34 (54.8)
Isotype, N (%) IgG IgA Light-chain	 42 (68) 20 (32) 6 (7)
Amount of M-protein, g/dL, mean (range)	2.9 (0.3-8.6)
Bone marrow infiltration, %, mean (range)	27.4 (5.5-78)
Cytogenetics, N (%) High-risk Standard risk Unknown	 18 (29) 40 (65) 4 (6)
High-risk, N (%) Mayo Pethema Both	 6 (10) 21 (34) 17 (27)
Ultra-high-risk, SLiM-CRAB criteria, N (%) BMPC ≥60% I/U FLC >100 >1 focal lesion by MRI	 5 (27.8) 10 (55.5) 3 (16.7)
M-spike by mass-spectrometry, %	100

N: number; BMPC: bone marrow plasma cell infiltration; FLC: serum free light chain; MRI: magnetic resonance imaging.

Performance of serum protein electrophoresis / immunofixation electrophoresis, mass spectrometry and next-generation flow cytometry to detect residual disease

As a first step, we compared the ability of the 3 methods to detect the presence of residual disease at the different stages of the treatment schema (i.e., the M-protein in serum using MS or serum protein electrophoresis [SPEP] / IFE and clonal plasma cells in BM by NGF) (Figure 1). During intensive treatment (post-induction, after ASCT and post-consolidation), MS detected residual disease in the highest proportion of patients; after the first year of maintenance both NGF and MS detected residual disease in equal proportions of patients, and NGF showed the highest percentage of positive cases at treatment completion, after 24 cycles of maintenance. Notably, as stated above, the number of samples available for analysis was lower during maintenance.

Mass spectrometry identified the M-protein in 46/62 (74%) post-induction, 33/61 (54%) post-ASCT, and 29/61 (48%) post-consolidation, 11/51 (22%) 1st year after maintenance and 7/35 (20%) after the 2nd year of maintenance (Figure 1).

Analysis of the combined results of mass spectrometry with serum protein electrophoresis / immunofixation electrophoresis or next-generation flow cytometry post-autologous stem cell transplant and after 24 cycles of maintenance

We then paired the results obtained with MS with those from one of the alternative techniques (SPEP/IFE or NGF) and analyzed the combined results two by two, post-ASCT and after two years of maintenance. When we analyzed the results of SPEP/IFE and MS (Figure 2A), we observed that at the two time-points analyzed, both methods were highly in agreement and that, among discordances, all were due to IFE$^-$/MS$^+$ samples, except for 4 cases IFE$^+$/MS$^-$. In detail, 72% of the results were concordant (29% IFE$^+$MS$^+$ and 43% IFE$^-$MS$^-$) and 28% discordant (25% IFE$^-$MS$^+$ and 3% IFE$^+$MS$^-$) post-ASCT, and 82% were concordant (9% IFE$^+$MS$^+$ and 73% IFE$^-$MS$^-$) and 18% discordant (12% IFE$^-$MS$^+$ and 6% IFE$^+$MS$^-$) at the end of the treatment. Overall, out of the 274 samples analyzed, 74.8% of the results were concordant (26.3% IFE$^+$MS$^+$ and 48.5% IFE$^-$MS$^-$) and 25.2% discordant (4.4% IFE$^+$MS$^-$ and 20.8% IFE$^-$MS$^+$).

Similarly, we analyzed the results of NGF and MS (Figure 2B). Approximately two-thirds of the results were concordant; among discordances, the majority of them post-ASCT were MS$^+$/NGF$^-$ while MS$^-$/NGF$^+$ at treatment completion. These results reflect the respective sensitivities of NGF and MS at the two time-points investigated (see above). In detail, 71% of the results were concordant (34% NGF$^+$MS$^+$ and 37% NGF$^-$MS$^-$) and 29% discordant (8% NGF$^+$MS$^-$ and 21% NGF$^-$MS$^+$) post-ASCT and 74% were concordant (14% NGF$^+$MS$^+$ and 60% NGF$^-$MS$^-$) and 26% discordant (20% NGF$^+$MS$^-$ and 6% NGF$^-$MS$^+$) at end-of- treatment. Overall, out of the 274 samples analyzed, 75% of the results were concordant (33% NGF$^+$MS$^+$ and 42% NGF$^-$MS$^-$) and 23% discordant (9.5% NGF$^+$MS$^-$ and 14% NGF$^-$MS$^+$).

Taking the NGF results as reference, the negative predictive value (NPV) of MS was 81% overall. Furthermore, the NPV of MS *versus* NGF was 83% and 73% post-ASCT and after two years of maintenance, respectively (Figure 2B).

Response rates and clinical value of serum protein electrophoresis / immunofixation electrophoresis, mass spectrometry and next-generation flow cytometry at the main end-points of the trial

The primary end-point of the trial was the rate of undetectable residual disease three months after ASCT assessed by NGF with a sensitivity level of 10^{-5}, which in the intention-to-treat (ITT) analysis was reached in 62% of the patients (56/90), as recently reported by Mateos *et*

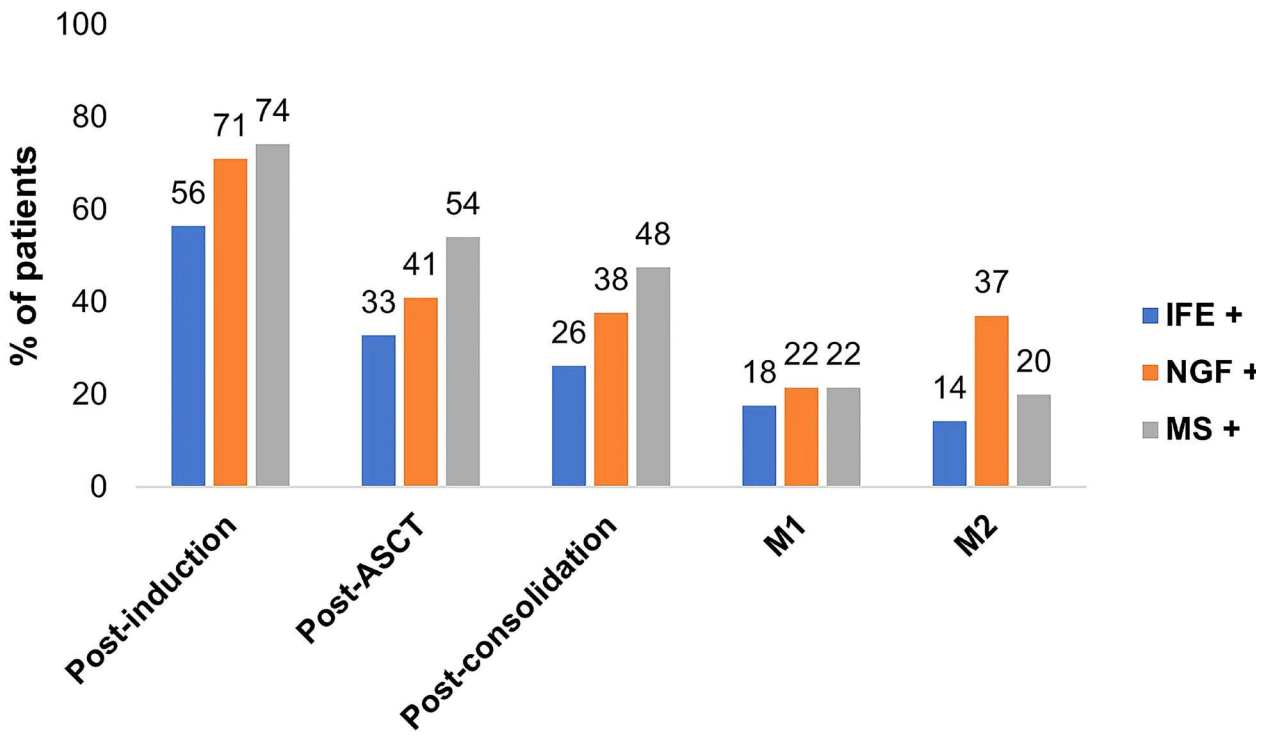
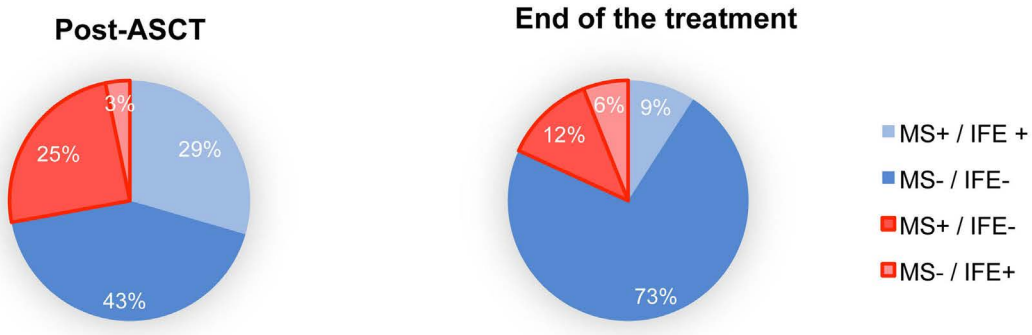


Figure 1. Percentages of patients with detectable residual disease. Percentages of patients with detectable residual disease by serum protein electrophoresis / immunofixation electrophoresis (IFE) (blue), mass-spectrometry (MS) (gray) and next-generation flow cytometry (NGF) (orange) at the 5 time-points analyzed in the trial: post-induction, after high-dose chemotherapy and autologous stem cell transplant (ASCT), post-consolidation, after one year of maintenance (M1), and at treatment completion after two years of maintenance (M2).

A



IFE/MS AGREEMENT **72% (44/61)**
IFE/MS DISAGREEMENT **28% (17/61)**

P value	<0.0001
Sensitivity	0.9000
Specificity	0.6341
PPV	0.5455
NPV	0.9286

IFE/MS AGREEMENT **82% (27/33)**
IFE/MS DISAGREEMENT **18% (6/33)**

P value	0.0519
Sensitivity	0.6000
Specificity	0.8571
PPV	0.4286
NPV	0.9231

Post-ASCT

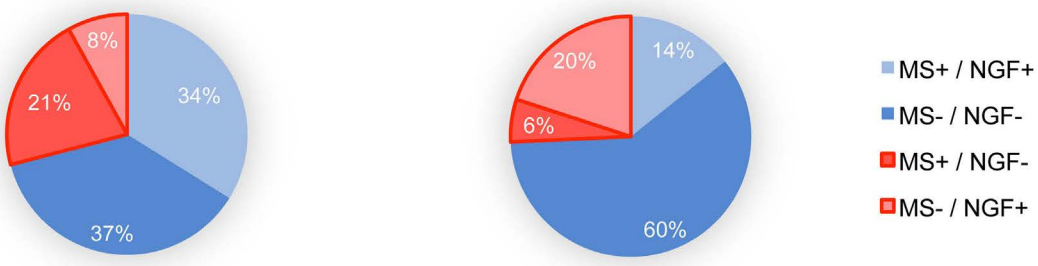
		IFE	
		Negative	Positive
MS	Negative	26 (43%)	2 (3%)
	Positive	15 (25%)	18 (29%)

End of the treatment

		IFE	
		Negative	Positive
MS	Negative	24 (73%)	2 (6%)
	Positive	4 (12%)	3 (9%)

Overall concordance: 74.8% (205/274); 72 (26.3%) $IFE^{+}MS^{+}$, 133 (48.5%) $IFE^{-}MS^{-}$
Overall discordance: 25.2% (69/274); 12 (4.4%) $IFE^{+}MS^{-}$, 57 (20.8%) $IFE^{-}MS^{+}$

B



NGF/MS AGREEMENT **71% (44/61)**
NGF/MS DISAGREEMENT **29% (18/61)**

P value	0.0007
Sensitivity	0.8077
Specificity	0.6389
PPV	0.6176
NPV	0.8214

NGF/MS AGREEMENT **74% (26/35)**
NGF/MS DISAGREEMENT **26%(9/33)**

P value	0.0331
Sensitivity	0.4167
Specificity	0.9130
PPV	0.7143
NPV	0.7500

Post-ASCT

		NGF	
		Negative	Positive
MS	Negative	23 (37%)	5(8%)
	Positive	13 (21%)	21 (34%)

End of the treatment

		NGF	
		Negative	Positive
MS	Negative	21 (60%)	7 (20%)
	Positive	2 (6%)	5 (14%)

Overall concordance: 75% (207/274); 90 (33%) $NGF^{+}MS^{+}$, 117 (42%) $NGF^{-}MS^{-}$
Overall discordance: 23.5% (67/274); 28 (9.5%) $NGF^{+}MS^{-}$, 39 (14%) $NGF^{-}MS^{+}$

Figure 2. Analysis of the combined results of mass spectrometry with serum protein electrophoresis / immunofixation electrophoresis or next-generation flow cytometry. Analysis of the combined results of mass spectrometry (MS) with serum protein electrophoresis (SPEP) / immunofixation electrophoresis (IFE) or next-generation flow cytometry (NGF) post-autologous stem cell transplant (ASCT) and after 24 cycles of maintenance. Percentages of concordant and discordant results, sensitivity, specificity, positive predictive value (PPV) and negative predictive values (NPV) of MS considering SPEP/IFE (A) and NGF (B) as a reference.

al. (M-V Mateos *et al.*, in press). The number of patients analyzed and rates of undetectable residual disease using SPEP/IFE, MS and NGF at the main time-points evaluated are detailed in Table 2.

Please note that, although serum samples were not available for MS analysis in all cases, underlining the difference between the two cohorts, the rates of CR+sCR and MRD negativity are quite comparable to those reported by Mateos *et al.* at the same time-points (M-V Mateos *et al.,* in press). Based on the previous results, we found that reaching CR+sCR according to the IMWG categories of standard response did not distinguish between the two cohorts of patients with significantly different outcome in terms of PFS at any of the main end-points analyzed (post-ASCT and after 2 years of maintenance) (Figure 3A).

According to the recommendations of the IMWG, we then analyzed the clinical value of MS and NGF in paired samples obtained from the group of patients reaching CR+sCR.[2] When focusing on these patients (42/61 [69%] post-ASCT and 28/35 [80%] at treatment completion), both PRD and MRD status were able to discriminate two cohorts with significantly different PFS, after ASCT and at treatment completion after two years of maintenance. As shown in Figure 3B, three months post-ASCT, the median PFS (mPFS) was not reached in PRD or MRD negative patients *versus* 4.1 years in PRD positive cases (P=0.0021) and 4.6 years in MRD positive cases (P=0.04); at treatment completion, the mPFS in PRD or MRD negative patients was not reached *versus* 1.1 years in PRD positive cases (P<0.0001) and 2.05 in MRD positive cases (P=0.0005).

Impact of peripheral residual disease and minimal residual disease dynamics in patient's outcome

Finally, we analyzed the dynamics of PRD by MS during the period of treatment and evaluated their impact on patient's outcome. Thirty-five patients with paired BM and PB samples available at at least 2 of the 5 time-points analyzed were included in this landmark analysis performed from the end of the treatment after completing the 24 cycles of maintenance. Sustained PRD negativity was observed in 12 patients (34%), and sustained positivity in 6 (17%). In 16 patients (46%), PRD converted from positive to negative and in one (3%) from negative to positive. Whereas sustained PRD$^+$ was associated with a very short mPFS of 1.66 years, patients with sustained PRD$^-$ or converting from PRD$^+$ to PRD$^-$ displayed a very favorable outcome with mPFS not reached and significantly different when compared with the sustained PRD$^+$ group (P<0.0001) (Figure 4). Interestingly, the only patient that converted from PRD$^-$ to PRD$^+$ had a very poor outcome with an mPFS of 3.8 months.

Sustained MRD negativity during the same period was observed in 8 patients (23%) and sustained positivity in 11 (31%). In 15 patients (43%), MRD converted from positive to negative, and in one (3%) from negative to positive. Whereas sustained MRD positivity was associated with a very short mPFS of 2.1 years, patients with sustained MRD$^-$ or who converted from MRD$^+$ to MRD$^-$ displayed a very favorable outcome with mPFS not reached and significantly different compared to those with sustained or converting to MRD$^+$. Please note that these figures cannot be directly compared with those from the whole cohort reported by Mateos *et al.* (M-V Mateos *et al.*, in press), where the analysis of the rate of sustained MRD negativity was calculated in patients MRD negative after ASCT that maintained the negative status four and five years after. Unfortunately, we do not have serum samples from the patients at those time-points to enable us to carry out an appropriate comparison.

Discussion

In this study, patients with HRsMM either by the Mayo or the Pethema models were treated with an intensive schema based on KRd induction and consolidation, ASCT and maintenance with Rd limited to two years.[11,12] Response assessment was based on MRD analysis in BM samples by NGF, and the primary end-point of the trial was the rate of MRD negativity three months post-ASCT in the ITT population. In 61 out of the total 90 patients included in the trial, we have analyzed and compared treatment response

Table 2. Number of patients analyzed and rates of undetectable residual disease using serum protein electrophoresis / immunofixation electrophoresis, mass spectrometry and next-generation flow cytometry.

	Post-induction N=62	Post-ASCT* N=61	Post-consolidation N=61	After 1 year maintenance N=51	After 2 years maintenance N=35
≥CR, N (%)	28 (45)	41 (67)	45 (74)	42 (82)	30 (86)
PRD^{neg}, N (%)	16 (26)	28 (46)	32 (52)	40 (78)	28 (80)
MRD^{neg}, N (%)	17 (27)	36 (58)	37 (61)	40 (78)	23 (66)

*Results corresponding to the main end-point of the trial. Minimal residual disease negativity rate (MRD^{neg}) refers to the percentage of cases with absence of phenotypically abnormal plasma cells in the bone marrow using EuroFlow recommendations (with a minimum sensitivity of 1 cell in 10^5 nucleated cells). Peripheral residual disease negativity rate (PRD^{neg}) refers to the percentage of patients with undetectable M-protein in serum samples using the Immunoglobulin Isotypes (GAM) assay for the MS EXENT analyzer; limit of detection: 0.015g/L. CR: complete response; ASCT: autologous stem cell transplant.

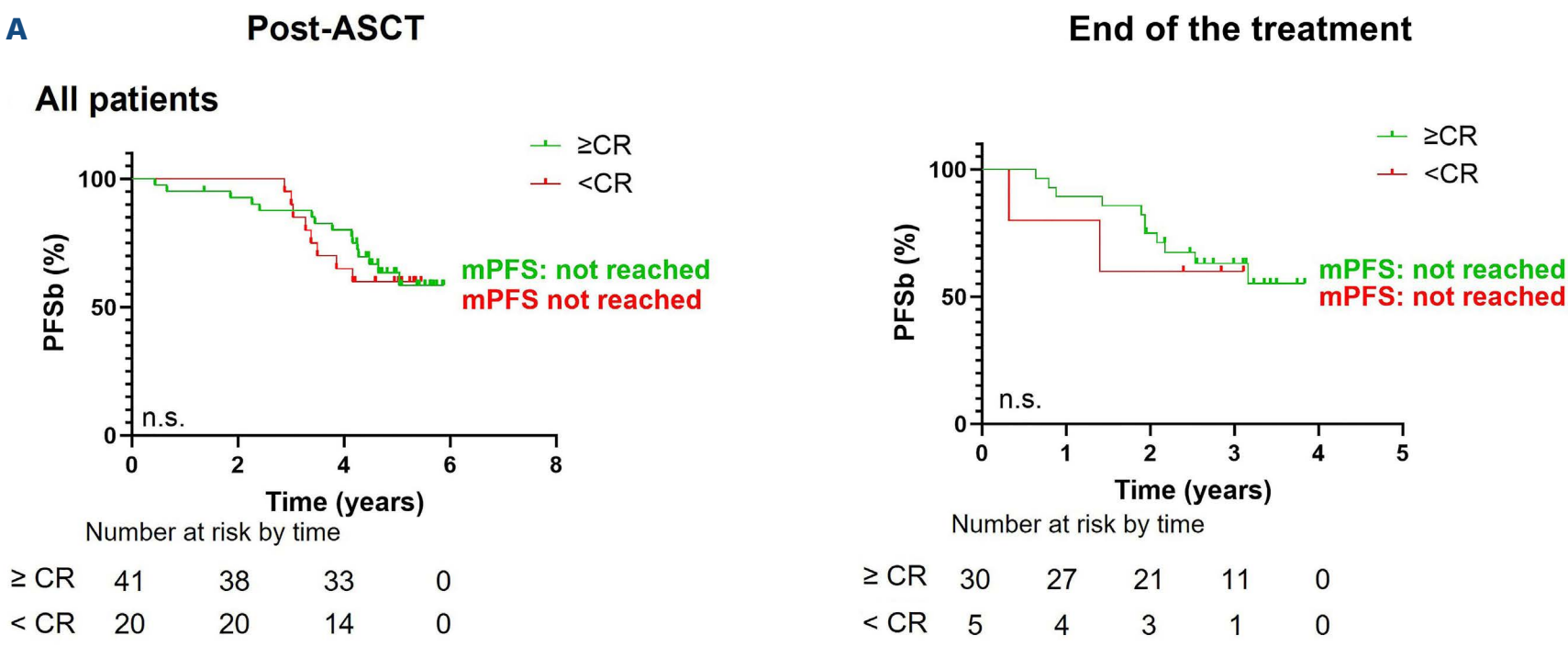


Figure 3. Progression-free survival after autologous stem cell transplant and at treatment completion after two years of maintenance. Progression-free survival (PFS) after autologous stem cell transplant (ASCT) (left) and at treatment completion after two years of maintenance (right) according to the results of (A) serum protein electrophoresis / immunofixation electrophoresis in all patients, (B) next-generation flow cytometry (NGF) in patients in complete response (CR) or stringent CR (sCR) and mass-spectrometry. MRD: minimal residual disease; PRD: peripheral residual disease; PFSb: biochemical PFS; mPFS: median PFS; HR: Hazard Ratio.

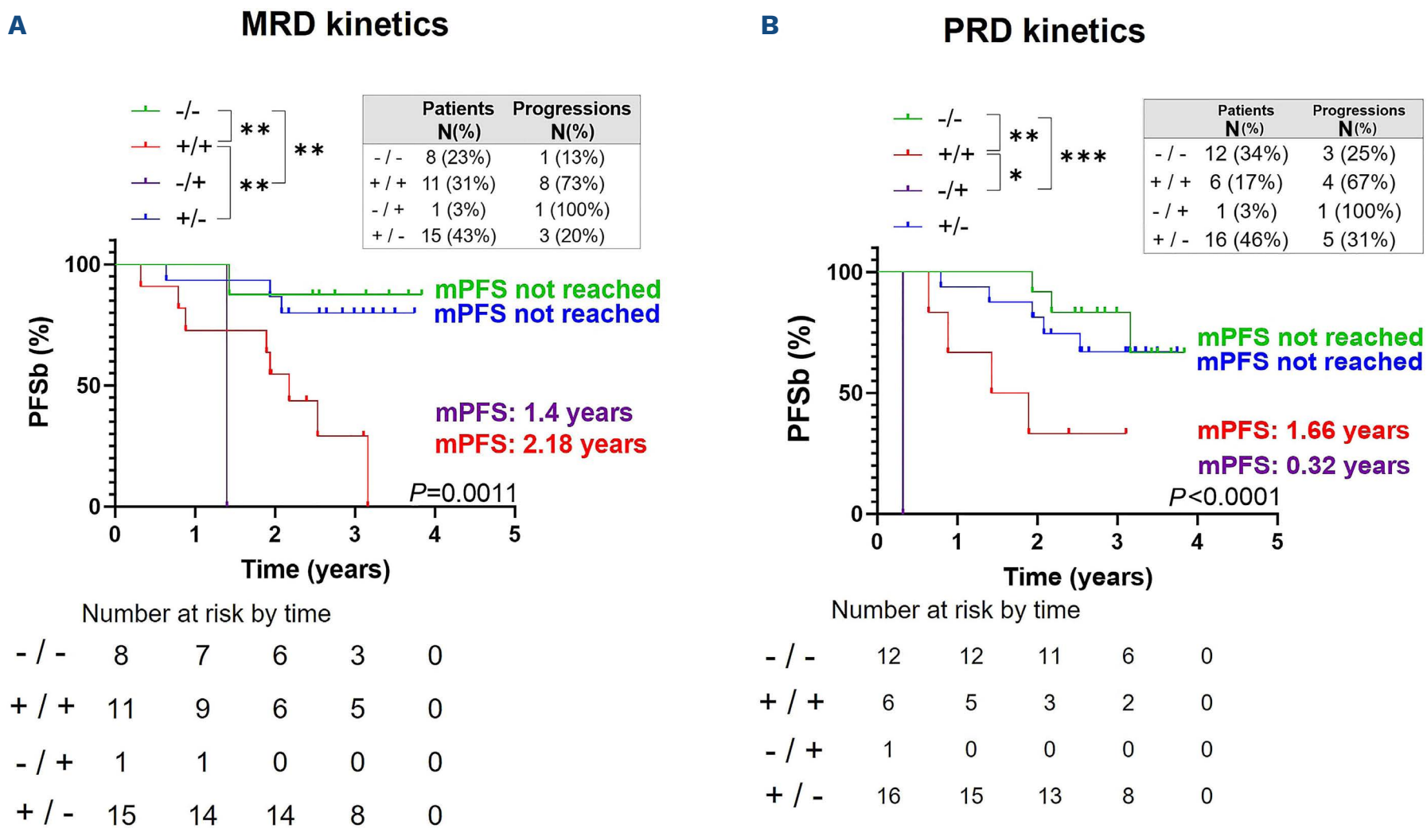


Figure 4. Landmark analysis of progression-free survival. Landmark analysis of progression-free survival (PFS) based on (A) minimal residual disease (MRD) or (B) peripheral residual disease (PRD) kinetics from randomization to treatment completion after 24 months of maintenance. PFSb: biochemical PFS; mPFS: median PFS. *P≤0.05, **P≤0.01, ***P≤0.001.

post-ASCT and at treatment completion using standard and MRD methods (SPEP/IFE in serum and NGF in BM, respectively) and MS, as a highly sensitive method to detect the presence of the patients' MP in serum samples.

From the results of our analysis, we have learnt that in this cohort of patients: 1) out of the 3 methods applied to detect residual disease, SPEP/IFE showed the lowest rates of positive results, MS the highest during intensive treatment, and both NGF and MS performed similarly during maintenance; 2) the results of SPEP/IFE do not associate with the clinical behavior of the disease in terms of PFS whereas MS and NGF both predict patient's outcome with similar accuracy; and 3) the clinical value of MRD and PRD analysis is improved when more than a single time-point is considered, but the dynamics of the results are taken into account.

In terms of the ability of the 3 methods to detect residual disease, SPEP/IFE showed the lowest rates of positive results, MS the highest during intensive treatment, and both NGF and MS behaved similarly during maintenance. This pattern (MS > NGF > SPEP/IFE) was maintained throughout, except at the end of the treatment, where the lower number of samples analyzed (N=35) could also have influenced our findings. We and others have shown the higher sensitivity of MS as compared to SPEP/IFE, which was further confirmed in this cohort of patients.[7,15-18] MS identified residual disease in a higher number of cases as compared to NGF at most of the time-points analyzed. This was an unexpected finding considering our previous results in patients included in the GEM2012-2014 clinical trials (N Puig *et al.,* 2024, submitted manuscript). Besides the potential limitations associated with MRD assessment in BM samples (i.e., hemodilution and patchy infiltration), more than 60% of the patients analyzed had an IgG MM, and the longer half-life of IgG could also be playing a role by increasing the detection rate of M-protein that had not yet been cleared at the initial stages of the treatment.[18] In this regard, analysis of MS^+/NGF^- cases (N=13) showed that 8 of them were obtained from patients with an IgG MM. Besides this, it is important to note that the higher detection rate of MS was not clinically misleading in this case, but also proved clinically valuable, similar to that observed with NGF. Finally, we acknowledge the fact that the target of serum-based techniques (the M-protein) is different to that of BM-based techniques (clonal/tumor plasma cells) and, therefore, discrepant results are expected, and their meaning should always be interpreted taking into account their clinical significance.

Despite the significant increase in treatment efficacy seen over recent years in patients with MM, treatment response continues to be assessed following guidelines that have largely remained unchanged.[2] Only the relatively recent

introduction of the stringent complete response (sCR) category aimed to improve this situation, although its clinical value has not been fully consolidated.[19,20] Currently, a very high proportion of patients reach the best category of standard response (CR or sCR) during or after treatments, and the GEM/PETHEMA group has been able to show that, among them, only patients found to be MRD negative in the BM achieve a more favorable outcome.[1] The lack of any discriminative capacity (all patients reach CR), together with its limited sensitivity (highlighted by the fact that only the MRD results are clinically determinant), are both translating into a situation in which classical response categories can no longer be associated with different clinical value. This was also recently shown in the context of the GEM-2012MAIN trial where depth of standard response criteria after treatment intensification had limited prognostic value in transplant-eligible MM patients.[21] In this study, likely due to the same reasons stated above, standard response criteria did not associate with a significant prognostic value in terms of PFS at any of the time-points analyzed.

Analysis of the results obtained post-ASCT and after two years of maintenance with MS and SPEP/IFE show that concordances (MS^+/IFE^+ and MS^-/IFE^-) increase with the progress of the therapy due to an increasingly higher number of double negative cases induced by the treatment. Accordingly, a lower number of discordances is identified; these are mostly MS^+ cases but $SPEP/IFE^-$ due to the higher sensitivity of the former. The higher sensitivity of MS compared to SPEP/IFE explains why the results of MS are of greater clinical value than those of SPEP/IFE. When we analyzed the combined results of MS and NGF, the figures remain quite similar after ASCT and two years of maintenance: two-thirds of the cases concordant and one-third discordant. However, whereas the majority of discordances post-ASCT were due to MS^+/NGF^- cases, in contrast, at the end of the treatment, they were mostly due to MS^-/NGF^+ cases. Importantly, both methods were in agreement in the majority of cases and discordances are explained by the ability of both techniques to identify residual disease in this specific cohort of patients. Here, it is also worth underlining the fact that, despite the expected discordances between two methods with different targets, the results of both of them were associated with patients' outcome. Also, taking NGF as the gold standard method for MRD detection, the NPV of MS at the two time-points analyzed was high enough (82% post-ASCT and 75% at treatment completion) to consider the results of MS as a factor that can determine the most appropriate moment to carry out a BM aspiration with the aim of confirming MRD negativity in patients with MM.

Although the clinical value of reaching MRD negativity has been widely shown,[3] various papers have been recently published regarding the value of the dynamics of MRD.[22-24] Whereas achieving sustained MRD negativity seems to be the most accurate predictor for long-term disease control, conversions from positive to negative or vice versa have also been shown to have an added clinical value.[23] We have also confirmed this in our paper. In fact, maintaining or converting to MRD or PRD negativity identified a subgroup of cases with a very favorable outcome and with an mPFS not yet reached after a median follow-up of 65 months; in contrast, sustained positivity or conversions from negative to positive, especially by MS, identify patients with an imminent risk of progressive disease.

Due to the timing of the trial, there were 18 out of the 62 (29%) patients included in our study that presented with one of the SLiM-CRAB criteria: 5 (27.8%) had ≥60% clonal plasma cells in the BM, 10 (55.5%) a ratio of involved / uninvolved serum FLC >100, and 3 (16.7%) had more than one focal lesion on magnetic resonance imaging. This represents almost one-third of the analyzed cohort that, together with the limited number of samples / cases analyzed, could explain the short PFS found in the whole series, and specially in those MRD or PRD positive.

In the present work, we show for the first time that, as opposed to SPEP/IFE, MS is a valuable clinical monitoring tool in PB at all stages of treatment in patients with HRsMM. These findings mainly relate to the higher sensitivity of MS as compared to SPEP/IFE and they could justify the introduction of a new serological MS-based response category. Furthermore, due to the comparable clinical value of MS and NGF, and the high NPV of MS taking NGF as a reference method (specially at later stages of the treatment), MS could also be used as a gateway to perform a BM aspiration/biopsy for MRD assessment. Finally, MS and NGF dynamics during treatment are both able to identify a subgroup of HRsMM with imminent risk of progressive disease with long-term disease control.

Disclosures

NP declares consultancy, honoraria, and research funding for Janssen, consultancy and honoraria for Amgen, Celgene and Takeda, membership on an entity's Board of Directors or advisory committees, research funding, Speakers Bureau, and honoraria from The Binding Site. CA and TC declare honoraria from The Binding Site. RG-S declares honoraria from Janssen, Takeda, and Amgen, a consulting or advisory role for Janssen, research funding from Gilead (Inst) and Incyte (Inst), patents, royalties, other intellectual property for BIOMED 2 primers (Inst), and travel expenses from Janssen and Takeda (I). JM-L declares Speakers' Bureau for Novartis, Janssen-Cilag, Celgene, Bristol-Myers Squibb, and Incyte, and research funding from Bristol-Myers Squibb. PR-O declares consultancy, honoraria, and Speakers Bureau for Celgene Corporation, consultancy and honoraria for Janssen, consultancy for Takeda and Kite Pharma, and honoraria from BMS. VG-C declares honoraria from Janssen, Pfizer, Bristol Myers Squibb (BMS)/Celgene, and GSK, travel expenses from Janssen and GSK, and participation on data safety monitoring or advisory board for Janssen. AO

declares consultancy and Speakers Bureau for Amgen, Takeda and Celgene Corporation, and consultancy for Janssen. NCG declares honoraria from Janssen-Cilag. RR-T declares consultancy or advisory role for Janssen and Celgene. RGS declares consultancy or advisory role for BMS/Celgene, Janssen, Amgen, Takeda and GSK, and Speakers Bureau from BMS/Celgene, Janssen, Amgen, and GSK. LR declares honoraria from Janssen, Celgene, Amgen and Takeda. JBa declares honoraria from Janssen, Celgene, Takeda, Amgen and Oncopeptides. AA declares membership on an entity's Board of Directors or advisory committees for Takeda, Amgen, Janssen, and Celgene. JdlR declares consultancy for Amgen, Celgene, Takeda, Janssen, and AbbVie. FdA declares honoraria from Janssen, Celgene, Amgen, GlaxoSmithKline, and Takeda. MTH declares consulting or advisory role for Celgene, Janssen, Amgen, Takeda, and GSK, and Speakers Bureau for Janssen, Celgene, and Amgen. JL-J declares honoraria for Gilead Sciences. EMO declares consultancy, honoraria, and research funding from Celgene and Amgen, honoraria from BMS, consultancy and honoraria from Takeda, Janssen, and Novartis, research funding from Mundipharma, consultancy for AbbVie, Seattle Genetics, and Pharmamar, and research funding from Sanofi and Array Pharmaceuticals. JBl declares honoraria from Jansen, Celgene, Takeda, Amgen, and Oncopeptides. J-JL declares consultancy or advisory role for Celgene, Takeda, Amgen, Janssen and Sanofi, and travel expenses from Celgene. M-TC declares honoraria from Janssen, Celgene, and Abbvie. BP declares collaboration with Amgen, Bristol-Myers Squibb, Celgene, Janssen, Merck, Novartis, Roche, and Sanofi, unrestricted grants from Celgene, EngMab, Sanofi, and Takeda, consultancy for Celgene, Janssen, and consultancy, honoraria, research funding, and Speakers Bureau for Sanofi. JFSM declares consultancy and honoraria for Amgen, Bristol-Myers Squibb, Celgene, Janssen, MSD, Novartis, Roche, Sanofi, and Takeda. M-VM declares membership on an entity's Board of Directors or advisory committees for GSK, Pharmamar, EDO, and Abbvie, honoraria and membership on an entity's Board of Directors or advisory committees for Celgene, Takeda, Janssen and Amgen, and honoraria from Adaptive. All other authors have no conflicts of interest to disclose.

Contributions

NP, CA, TC, MVM and JSM conceived the analysis and designed the analysis protocol. JML, PRO, VGC, MSG, AO, NCG, RRT, LR, MAA, JBa, APGR, AA, FE, MBI, JdlR, AIT, FdA, LP, MTH, JLJ, MR, AGM, EMO, BP, NP, MTC, JBl and JJL provided study samples or patients. NP, CA, TC and SC analyzed the mass spectrometry data and the M-protein kinetics. NP, CA, TC, MVM, JML, LR, JdlR, BP, NP, MTC, JB, JJL and JSM analyzed and interpreted data. All authors contributed to the writing and approved the final manuscript.

Funding

This study was supported by the Instituto de Salud Carlos III/Subdirección General de Investigación Sanitaria, Spain and by the Leukemia and Lymphoma Society, grant n. PI21/01917 (to M-VM, NP and TC), n. PI21/01751 (to MTC), PI21/01816 (to MJC), and LLS Grant ID 6660-23 (to M-VM, NP and CA).

Data-sharing statement

If required, these data can be obtained through the corresponding author.

References

1. Lahuerta JJ, Paiva B, Vidriales MB, et al. Depth of response in multiple myeloma: a pooled analysis of three PETHEMA/GEM clinical trials. J Clin Oncol. 2017;35(25):2900-2910.
2. 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(8):e328-346.
3. Munshi NC, Avet-Loiseau H, Anderson KC, et al. A large meta-analysis establishes the role of MRD negativity in long-term survival outcomes in patients with multiple myeloma. Blood Adv. 2020;4(23):5988-5999.
4. Sanoja-Flores L, Flores-Montero J, Puig N, et al. Blood monitoring of circulating tumor plasma cells by next generation flow in multiple myeloma after therapy. Blood. 2019;134(24):2218-2222.
5. Mazzotti C, Buisson L, Maheo S, et al. Myeloma MRD by deep sequencing from circulating tumor DNA does not correlate with results obtained in the bone marrow. Blood Adv. 2018;13(2):2811-2813.
6. Zajec M, Langerhorst P, VanDuijn MM, et al. Mass spectrometry for identification, monitoring, and minimal residual disease detection of M-proteins. Clin Chem. 2020;66(3):421-433.
7. Puig N, Contreras MT, Agulló C, et al. Mass spectrometry vs immunofixation for treatment monitoring in multiple myeloma. Blood Adv. 2022;6(11):3234-3239.
8. Fan H, Wang B, Shi L, et al. Monitoring minimal residual disease in patients with multiple myeloma by targeted tracking serum M-protein using mass spectrometry (EasyM). Clin Cancer Res. 2024;30(6):1131-1142.
9. Martins CO, Huet S, Yi SS, et al. Mass spectrometry-based method targeting Ig variable regions for assessment of minimal residual disease in multiple myeloma. J Mol Diagnostics. 2020;22(7):901-911.
10. Kyle RA, Durie BG, Rajkumar SV, et al. Monoclonal gammopathy of undetermined significance (MGUS) and smoldering (asymptomatic) multiple myeloma: IMWG consensus perspectives risk factors for progression and guidelines for monitoring and management. Leukemia. 2010;24(6):1121-1127.
11. Kyle RA, Remstein ED, Therneau TM, et al. Clinical course and prognosis of smoldering (asymptomatic) multiple myeloma. N Engl J Med. 2007;356(25):2582-2590.
12. Pérez-Persona E, Vidriales MB, Mateo G, et al. New criteria to

identify risk of progression in monoclonal gammopathy of uncertain significance and smoldering multiple myeloma based on multiparameter flow cytometry analysis of bone marrow plasma cells. Blood. 2007;110(7):2586-2592.
13. Rajkumar SV, Dimopoulos MA, Palumbo A, et al. International Myeloma Working Group updated criteria for the diagnosis of multiple myeloma. Lancet Oncol. 2014;15(12):e538-548.
14. Flores-Montero J, Sanoja-Flores L, Paiva B, et al. Next generation flow for highly sensitive and standardized detection of minimal residual disease in multiple myeloma. Leukemia. 2017;31(10):2094-2103.
15. Santockyte R, Jin C, Pratt J, et al. Sensitive multiple myeloma disease monitoring by mass spectrometry. Blood Cancer J. 2021;11(4):78.
16. Dispenzieri A, Krishnan A, Arendt B, et al. Mass-Fix better predicts for PFS and OS than standard methods among multiple myeloma patients participating on the STAMINA trial (BMTCTN0702/07LT). Blood Cancer J. 2022;12(2):27.
17. Campbell L, Simpson D, Ramasamy K SR. Using quantitative immunoprecipitation mass spectrometry (QIP-MS) to identify low level monoclonal proteins. Clin Biochem. 2021;95:81-83.
18. Mai EK, Huhn S, Miah K, et al. Implications and prognostic impact of mass spectrometry in patients with newly-diagnosed multiple myeloma. Blood Cancer J. 2023;13(1):1-10.
19. Durie BGM, Harousseau JL, Miguel JS, et al. International uniform response criteria for multiple myeloma. Leukemia. 2006;20(9):1467-1473.
20. Martínez-López J, Paiva B, López-Anglada L, et al. Critical analysis of the stringent complete response in multiple myeloma: contribution of sFLC and bone marrow clonality. Blood. 2015;126(7):858-862.
21. Jiménez-Ubieto A, Paiva B, Puig N, et al. Validation of the International Myeloma Working Group standard response criteria in the PETHEMA/GEM2012MENOS65 study: are these times of change? Blood. 2021;138(19):1901-1905.
22. San-Miguel J, Avet-Loiseau H, Paiva B, et al. Sustained minimal residual disease negativity in newly diagnosed multiple myeloma and the impact of daratumumab in MAIA and ALCYONE. Blood. 2022;139(4):492-501.
23. Paiva B, Manrique I, Dimopoulos MA, et al. MRD dynamics during maintenance for improved prognostication of 1280 patients with myeloma in the TOURMALINE-MM3 and -MM4 trials. Blood. 2023;141(6):579-591.
24. Avet-Loiseau H, San-Miguel J, Casneuf T, et al. Evaluation of sustained minimal residual disease negativity with daratumumab-combination regimens in relapsed and/or refractory multiple myeloma: analysis of POLLUX and CASTOR. J Clin Oncol. 2021;39(10):1139-1149.