Flow cytometry has become a highly valuable method to monitor minimal residual disease (MRD) and evaluate the depth of complete response (CR) in bone marrow (BM) of multiple myeloma (MM) after therapy. However, current flow-MRD has lower sensitivity than molecular methods and lacks standardization. Here we report on a novel next generation flow (NGF) approach for highly sensitive and standardized MRD detection in MM. An optimized 2-tube 8-color antibody panel was constructed in five cycles of design-evaluation-redesign. In addition, a bulk-lysis procedure was established for acquisition of $\geq 10^7$ cells/sample, and novel software tools were constructed for automatic plasma cell gating. Multicenter evaluation of 110 follow-up BM from MM patients in very good partial response (VGPR) or CR showed a higher sensitivity for NGF-MRD vs conventional 8-color flow-MRD-MRD-positive rate of 47 vs 34% ($P = 0.003$). Thus, 25% of patients classified as MRD-negative by conventional 8-color flow were MRD-positive by NGF, translating into a significantly longer progression-free survival for MRD-negative vs MRD-positive CR patients by NGF (75% progression-free survival not reached vs 7 months; $P = 0.02$). This study establishes EuroFlow-based NGF as a highly sensitive, fully standardized approach for MRD detection in MM which overcomes the major limitations of conventional flow-MRD methods and is ready for implementation in routine diagnostics.

Leukemia (2017) 31, 2094–2103; doi:10.1038/leu.2017.29

INTRODUCTION

In recent years, the landscape of treatment of multiple myeloma (MM) has substantially changed, leading to significantly increased complete response (CR) rates and survival. Therefore, CR has become a major goal in newly diagnosed MM, even in aged patients. Despite these advances, most CR patients will ultimately relapse. Consequently, better insight in depth of treatment response is required and more sensitive methods are needed for detection of minimal residual disease (MRD), particularly in cases that reached CR and stringent CR (sCR).

Conventional 4–8-color flow cytometry, and to a lesser extent also allele-specific oligonucleotide quantitative PCR (ASO-qPCR) and next generation sequencing (NGS), are progressively being used to monitor MRD in bone marrow (BM) of MM after therapy. These studies confirmed the relevance of MRD measurements for identification of MM patients at higher risk of relapse. However, despite the greater sensitivity of the MRD approaches (vs classical CR/sCR criteria), identification among MRD cases of patients that will eventually relapse vs those who are potentially cured still remains a challenge, implying that more sensitive MRD approaches are needed.

High sensitivity and broad applicability have both become mandatory requirements for MRD monitoring in MM. Early studies have shown that conventional 4–6-color flow-MRD is applicable in virtually all MM patients ($\geq 95\%$), whereas ASO-qPCR and NGS have a more restricted applicability (50–90% of cases), mainly due to the high number of somatic hypermutations, which cause variable levels of primer annealing with unpredictable amplification/quantitation results. However, the sensitivity of conventional flow-MRD ($< 10^{-9}$) remains (systematically) lower.
than that of ASOqPCR and NGS (1 × 10−4 to 10−6). More recently, several studies in MM have shown that conventional 8-color flow-MRD assays have an increased sensitivity – limit of detection (LOD) of between 1 × 10−4 and 1 × 10−5, leading to a significantly improved prediction of outcome. However, current flow-MRD assays and NGS still suffer from a major limitation: lack of standardization. Actually, different markers and antibody panels, distinct numbers of cells measured, and highly variable criteria for MRD positivity are currently applied worldwide. Therefore, standardization efforts have been made and consensus recommendations and guidelines have been recently proposed. However, such consensus recommendations still rely on subjective ‘expert-shared’ knowledge and experience, and do not completely solve the lack of technical standardization, whereas prospectively validated flow-MRD approaches are still missing.

Here we report on the design of a EuroFlow-based next generation flow (NGF) approach for highly sensitive and standardized detection of MRD in MM, and the results of its validation vs a conventional 8-color flow-MRD method and NGS. The novel NGF-MRD approach takes advantage of innovative tools and procedures recently developed by the EuroFlow Consortium for sample preparation, antibody panel construction (including choice of type of antibody and fluorochrome), and automated identification of plasma cells (PC) against reference databases of normal and patient BM.

Prospective validation of the whole procedure at two distinct centers confirmed its robustness and significantly greater sensitivity vs conventional 8-color flow-MRD approaches, comparable to current NGS methods, with an improved prediction of patient outcome.

**SUBJECTS AND METHODS**

Patients, controls and samples
A total of 375 BM and 10 peripheral blood EDTA-anticoagulated samples from 53 controls and 332 adult plasma cell disorder (PCD) patients were studied (Supplementary Table 1) to design Version 1 of the MM-MRD antibody panel (n = 94; 31 normal/reactive and 63 MM studied at diagnosis), to compare the performance of antibody reagents evaluated in Versions 2–5 of the NGF-MRD panel (19 BM diagnostic patient samples), to evaluate distinct sample preparation protocols (n = 10 peripheral blood and 8 BM samples) and to validate NGF (Version 5) against conventional 8-color flow-MRD - n = 244 consecutively recruited samples corresponding to 16 healthy donors, 66 PC samples studied at diagnosis and 162 MM patient samples investigated during follow-up, including 110 follow-up BM samples from MM patients evaluated at ≥ very good partial response (VGPR); 39 VGPR; 52 CR, and 19 SCR cases. Twenty-one ≥ VGPR BM samples and 10 additional BM samples -22/31 with very low MRD levels (≤10−5), were further used to blindly compare NGF vs NGS.

The study was approved by the local ethics committees and written informed consent was given by each donor according to the Declaration of Helsinki. All samples were processed at each participating center (USAL/HUSAL, UNAV, EMC, UNIKEL, IPOFLG, SUM) within 24 h after collection.

Immunophenotypic studies for selection of plasma cell (PC)-associated markers
BM samples used to establish Version 1 (Table 1) of the MM-MRD panel were stained with the EuroFlow PCD 8-color panel for a total of 12 different markers (CD38, CD138, CD45, CD19, CD27, CD28, CD56, CD81, CD117, CyIgG, CyIgA and CyIgM) and used to set up a PC-identifiable monoclonal plasma cell (aPC) population:

- **CD38** - was used to distinguish plasma cells (PCs) from their reactive counterparts.
- **CD138** - is a specific marker for mature PCs.
- **CD45** - is a pan-leukocyte marker.
- **CD19** - is a marker for B cells.
- **CD27** - is a marker for mature PCs.
- **CD28** - is a marker for activated T cells.
- **CD56** - is a marker for natural killer (NK) cells.
- **CD81** - is a marker for PCs.
- **CD117** - is a marker for stem cells.
- **CyIgG**, **CyIgA**, and **CyIgM** - are markers for mature PCs.

These markers were used to select the reference database of nPC vs aPC from each of 63 MM BM studied at diagnosis, to identify the most discriminating markers that distinguished between (reference) nPCs and aPCs from individual MM patients (Figure 1), and establish the applicability of the method. Additionally, CD200 expression was also evaluated on PC in a subset of 28 MM.

Validation of the NGF MM-MRD method
Overall, 228 MM diagnostic (n = 66) and follow-up (n = 162) BM samples (n = 110 in VGPR or CR/SCR) were evaluated in parallel with the NGF MRD approach vs local routine flow-MRD methods (that is, conventional 8-color flow-MRD technique). Detailed description of these BM samples, related patient clinical data, disease status and time points at evaluation is provided in Supplementary Tables 1, 3 and 4. Briefly, conventional flow-MRD was based on staining of 300 µl of whole BM with a single 8-color antibody combination (CD45-PacO, CD56-PE, CD5-PerCPCy5.5, CD19-PECy7, CD3-APC) and scored as follows:

- **CD138** - was used to discard or accept individual reagents: (i) increased background fluorescence; (ii) dim fluorescence of PCs vs other BM populations (CD38, CD138, CyIgG or CyIgA) or the positive vs negative BM reference populations (CD19, CD27, CD56, CD45, CD81 vs CD117) and (iii) interaction with the staining pattern of other reagents.

Design and evaluation of sample preparation protocols
For the evaluation of different sample preparation protocols, eight BM and 10 peripheral blood samples were stained in parallel with the CD138-HV450, CD45-PacO, CD56-PE, CD5-PerCPCy5.5, CD19-PECy7, CD3-APC antibody combination under five different conditions:

- **(i)** the EuroFlow standard operating procedure (SOP) for staining of 50 µl of sample with cell surface markers,
- **(ii)** four different ammonium chloride-based bulk-lysis procedures followed by staining of 5 × 10^6 cells in 100 µl/tube (final concentration of 5 × 10^4 cells/µl), as described in detail in the Supplementary text.

Results on the comparison between antibody reagents directed against the same antigen were published elsewhere. 44

Next gen flow-minimal residual disease in myeloma
J Flores-Montero et al

Leukemia (2017) 2094 – 2103
The percentage MRD was calculated upon dividing the number (×100) of MRD cells by the total number of cells in the reaction.

Statistical methods
Number of samples required in the assay development and validation experiments were calculated using the hypothesis contrast strategy for the comparison of paired quantitative data (EPIDAT 4.0 software, Consellería de Sanidade, Xunta de Galicia, Spain). To evaluate the distribution of MRD data, the Kolmogorov-Smirnov test was used. The Wilcoxon or Friedman tests and the Mann-Whitney U or the Kruskal-Wallis tests were used to assess the (two-sided) statistical significance of differences observed between ≥2 groups for paired and unpaired variables, respectively. For correlation studies, the (two-sided) Spearman’s rho (ρ) for non-parametric
paired data was employed. The Kaplan–Meier method and the (two-sided) log-rank test were used to plot and compare progression-free survival (PFS) curves. PFS was defined as the time from MRD assessment to either disease progression or the last follow-up visit. Statistical significance was set at \( P < 0.05 \). All samples evaluated were blindly analyzed during the experimental phase.

RESULTS

Antibody panel construction for optimal identification of MM plasma cells at MRD levels

Comparison of the whole immunophenotypic profile of aPC from individual MM patients \((n = 63)\) vs the normal/reactive BM PC database \((n = 31)\) showed multiple aberrant phenotypes in every case (Figure 1). Eight of the 12 markers evaluated contributed most frequently to the discrimination between aPC and nPC based on principal component analysis of single PC phenotypes: CD19 \((97\%)\); CD45 \((89\%)\); CD56 \((86\%)\); CD81 \((86\%)\); CD38 \((59\%)\); CD38 \((73\%)\); CD27 \((71\%)\); CD117 \((60\%)\); and CyIg \((64\%)\). Re-evaluation of the utility of the combination of these eight markers alone in the same 63 MM BM confirmed clear-cut distinction between aPC and nPC in the database, in every case. Consequently, the six surface membrane markers of this list, plus the CD138 and CD38 PC-identification markers, were selected to be combined in a single 8-color tube. In a second 8-color antibody combination, CyIg and CyQL were added to the CD138 and CD38, together with CD229 as an extra PC-identification antigen, plus the three most informative markers (CD19, CD45, CD56), for parallel confirmation of immunoglobulin light chain \((\kappa v \lambda)\) restriction of PC suspected to be (clonal) myeloma PC (Versions 2 and 3).

Optimization of the two 8-color MM MRD antibody combinations

Subsequently, evaluation of the same 63 MM diagnostic samples using the two 8-color antibody combinations selected above, but focusing now on the detection of minimal numbers (that is, 0.02–0.1%) of MM PC, was performed using virtual (software) dilution experiments of decreasing numbers of PC in the nPC reference database, as previously described.\(^{51}\) This revealed suboptimal reagent performance (for example dim staining) for two fluorochrome positions (that is, CD138-PacO and CD81/CyIg-APC750).

From this initial Version 1, until the final version of the antibody panel (Version 5), four other versions of different fluorochrome-conjugated reagents of the same markers were tested (Table 1) as described in detail in Supplementary text. Briefly, in Version 2, the suboptimal CD138-PacO reagent was replaced by CD138-HV500C; in Version 3, suboptimal CD81-APC750 and CyIg-APC750 reagents were both replaced by CD81-APCC750 and CyIg-APC750. For Version 4, the CD138-HV500C reagent found to be still suboptimal was replaced, together with the CD27-PerCP Cy5.5 and CD45-PacB fluorochrome positions, by CD138-HV450, CD27-HV500C and CD45-PerCP Cy5.5, respectively. Finally, in Version 5, CD138-HV450, CD27-HV500C and CD38-FITC were replaced by the optimized CD138-BV421 and CD27-BV510 conjugates, and the multi-epitope CD38-FITC antibody, respectively. The later CD38 reagent showed an equivalent performance in diagnostic MM samples to that of the original CD38 antibody clone, but a much better resolution in BM samples from MM patients who had received Daratumumab therapy (Supplementary Figure 1). Moreover, CD229 was excluded from tube 2, since this marker did not identify all aPC in 4/49 (8%) MM cases tested, and it was not specific for PC, being also (strongly) expressed on plasmacytoid dendritic cells and a subset of lymphocytes.\(^{52}\)

Evaluation of sample preparation protocols

Overall, bulk-lysis procedures were systematically associated with acquisition of a significantly \((P < 0.05)\) greater number of cells vs the conventional BD FACS Lysing Solution (BD Biosciences, San Jose, CA, USA) based (FACS-lyse) SOP (Supplementary Figure 2A). However, all bulk-lysis conditions but that using low bovine serum albumin (0.5% bovine serum albumin) and a FACS-lysis-fixation step (protocol A1 in Supplementary Figure 2), showed a significantly higher proportion of debris and dead cells \((P < 0.05)\) from previous versions in markers or fluorochromes are highlighted in italics.

Table 1. Multiple myeloma NGF-MRD: 8- and 10-color antibody panels evaluated from the first (Version 1) to the final version (Version 5)

| 8-color panel version | Tube | PacB | PacO | FITC | PE | PerCP Cy5.5 | PECy7 | APC | APCC750 |
|-----------------------|------|------|------|------|----|-------------|-------|-----|---------|
| 1                     |      |      |      |      |    |             |       |     |         |
| 2                     |      |      |      |      |    |             |       |     |         |
| 3                     |      |      |      |      |    |             |       |     |         |
| 4                     |      |      |      |      |    |             |       |     |         |
| 5                     |      |      |      |      |    |             |       |     |         |

| 10-color panel version | BV421 | BV510 | BV605 | FITC | PE | PerCP Cy5.5 | PECy7 | APC | APCC750 |
|------------------------|-------|-------|-------|------|----|-------------|-------|-----|---------|
| 1                      |       |       |       |      |    |             |       |     |         |
| 2                      |       |       |       |      |    |             |       |     |         |
| 3                      |       |       |       |      |    |             |       |     |         |

Abbreviations: APC, allophycocyanin; A700, alexa fluor 700; BV421, brilliant violet 421; BV510, brilliant violet 510; BV605, brilliant violet 605; Cy, cytoplasmic; C750, C750 dye; CyS5, cyanin S5; Cy7, cyanin 7; FITC, fluorescein isothiocyanate; HV450, Horizon V450; HV500C, Horizon V500C; H7, Hillite 7; Ig, immunoglobulin; MRD, minimal residual disease; ME, multi-epitope; PacB, pacific blue; PacO, pacific orange; PE, phycoerythrin; PerCP, peridinin–chlorophyll–protein. Changes from previous versions in markers or fluorochromes are highlighted in italics.
vs FACS-lyse protocols) associated with similar numbers (P > 0.05) of cell doublets (Supplementary Figure 2B). Of note, detailed analysis of the specific recovery of the major leukocyte populations (that is, eosinophils, neutrophils, monocytes, mature lymphocytes and nucleated red cells) as well as of nPC and aPC showed no significant differences between the conditions evaluated (P > 0.05), except for higher nPC percentages for the bulk-lysis protocol B1 (Supplementary Figures 2C and D). Therefore, the bulk-lysis procedure including a FACS-lysing-fixation step and 0.5% bovine serum albumin (protocol A1 in Supplementary Figure 2) was selected as the reference SOP and used to further titrate the individual antibody reagents selected, for staining of ≥ 10^7 cells/tube (Supplementary Table 5).

Validation of the EuroFlow-based NGF MM-MRD method

The final NGF MM-MRD approach described above was validated against conventional 8-color flow-MRD in 228 MM BM samples studied at diagnosis (n = 66) or after therapy (n = 162), particularly focusing on 110 BM samples from patients in VGPR or CR/sCR. A strong correlation was found between both methods in diagnostic and follow-up samples from patients with stable/progressive disease and partial response (n = 118), while in the two flow methods are specifically compared for follow-up samples from MM patients in VGPR and CR/sCR (n = 110). (d) The correlation between expert-based vs automatic PC identification MRD levels in the same 110 BM samples as those of b. In turn, c shows the correlation between NGS and NGF MRD levels for those 27 (low level) MRD samples analyzed by both methods.*Samples proven polyclonal by Cy Ig κ/λ staining (2/2 and 2/3 in a and b, respectively). Samples positive by NGS at the limit of quantitation of the technique. White and black circles in b and d represent NGF MRD levels below and above the limit of quantitation of the technique, respectively.

Figure 2. Validation of the new NGF method for MRD detection in MM against both conventional 8-color flow-MRD (a, b) and NGS (c), including expert-based vs automated NGF-MRD data analysis (d). In a, the comparison between NGF and conventional flow-MRD is shown for diagnostic and follow-up samples from patients with stable/progressive disease and partial response (n = 118), while in b the two flow methods are specifically compared for follow-up samples from MM patients in VGPR and CR/sCR (n = 110). (d) The correlation between expert-based vs automatic PC identification MRD levels in the same 110 BM samples as those of b. In turn, c shows the correlation between NGS and NGF MRD levels for those 27 (low level) MRD samples analyzed by both methods. *Samples proven polyclonal by Cy Ig κ/λ staining (2/2 and 2/3 in a and b, respectively). Samples positive by NGS at the limit of quantitation of the technique. White and black circles in b and d represent NGF MRD levels below and above the limit of quantitation of the technique, respectively.
high number of discrepant cases, which were mostly MRD⁺ by EuroFlow-NGF but MRD⁻ by conventional flow: 18/21 (86%) vs 3/21 (14%) discrepant samples, respectively. Of note, such discrepant NGF cases typically showed MRD levels < 10⁻⁴ by EuroFlow-NGF with median (range) MRD levels of 0.0001% (0.0001–0.03%) vs 0.02% (0.0008–1.79%) for MRD⁺ cases by both methods (P < 0.001; Figures 2b and 3). Interestingly, in three MRD⁻ cases by EuroFlow-NGF, MRD⁺ results at relatively high levels (median of 0.01%; range: 0.006–0.02%) were observed by conventional MRD-flow (Table 2). Evaluation of cytoplasmic κ/λ expression in the suspicious PC from these three patients by EuroFlow-NGF, demonstrated the polyclonal nature of the suspected aPC in 2/3 cases, indicating false-positivity in conventional flow-MRD; in contrast, no clear explanation was found for the discrepant results observed in the other patient. Overall, the frequency of aberrant expression profiles for individual markers on clonal PC by NGF, was as follows: CD45, 96%; CD19, 96%; CD56, 96%; CD27, 89%; CD81, 79%; CD38, 77%; CD117, 48% and CD138, 37%. In 8/52 MRD⁻ cases confirmation of light chain restriction among small numbers of suspicious PC carrying slightly aberrant phenotypes was required. No significant differences were observed in the validation phase between centers with respect to rate and type of MRD discrepant cases (P = 0.63). Importantly, in all but 7/110 cases, > 7 x 10⁶ cells were evaluated (median 10.4 x 10⁶ cells) with an impact on the sensitivity of the method because of not reaching the LOD in only 2 cases (1.8%). Interestingly, an alarm for decreased percentage of CD117hi mast cells (≤ 0.002%) suggesting blood contamination, was observed in 17/110 samples, 11/17 MRD⁻ and 6/17 MRD⁺ samples (Table 3).

Automatic identification and enumeration of aPCs showed an excellent correlation with expert-based gating, also in VGPR and CR/sCR samples (p = 0.96; P < 0.001; Figure 2d). However, due to the minimum number of events required by the software algorithm, aPCs were not identified in 3/110 cases (2.7%) with low MRD levels by NGF (and MRD⁻ by conventional flow-MRD) –median % aPC (range): 0.006% (0.0001–0.005%: Figure 2d). Parallel assessment of MRD by NGF and NGS in a subset of 31 samples showed a higher applicability for the EuroFlow-NGF approach: 31/31 (100%) vs 27/31 (87%) cases, respectively (P < 0.001). Among those 27 cases assessed by both methods (22 of them with MRD levels ≤ 10⁻⁴), a good correlation was found between the percentage of residual aPC by NGF and NGS (p = 0.62; P = 0.001; Figure 2c). However, NGF showed a higher sensitivity than NGS with 19/27 (70%) vs 14/27 (52%) MRD⁺ samples (P = 0.06) with higher MRD levels –median percentage ± s.d. MRD⁺ cells of 0.01 ± 0.04% vs 0.006 ± 0.02% (P = 0.07), respectively-. This was due to 7/27 discrepant cases including six MRD⁺ by NGF and MRD⁻ by NGS: median (range) percentage aPC of 0.001% (0.0002–0.07%) and one MRD⁺ by NGS (0.0004% aPC) and negative (LOD < 0.0002% aPC) by NGF (Figures 2c and 3). In fact, 8/27 samples (30%) were NGF-positive (qualifiable), but NGS-negative or discrepantly low positive.

From the prognostic point of view, MM patients who were MRD⁺ by NGF had a significantly (P = 0.01) longer PFS vs MRD⁻ cases -75% PFS not reached (NR) vs 10 months; Figure 4a-, including also those that were MRD⁺ by NGF and MRD⁻ by conventional 8-color flow (75% PFS of 10 months; Figure 4b); similar results were observed when the analysis was restricted to patients in CR/sCR (P = 0.02; 75% PFS NR vs 7 and 5 months, respectively; Figures 4c and d). Of note 2/6 patients who were MRD⁺ by NGF and MRD⁻ by NGS also showed disease progression, while the only NGF⁻/NGS⁺ discrepant patient remained in continuous CR after 14 months follow-up.

Figure 3. Illustrating graphical representations of the performance of the NGF method based on the analysis of (merged) data files corresponding to a BM MM sample (> 10⁶ cells) with low level MRD stained with the NGF-MM MRD panel (Version 5). Left panels show classical bivariate dot plot representations in which PC (blue and red dots) were gated using a conventional manual analysis strategy. nPC (blue dots) display characteristic normal patterns of expression for the surface membrane markers used, with a cytoplasmic (Cy) Ig vs Cy Igλ ratio of 1.6. In contrast, clonal/aberrant PC (red dots) can be clearly discriminated from nPC based on their more homogeneous phenotypic profile, the presence of myeloma-associated phenotypes (CD138hi, CD38dim, CD19-, CD81-, CD117- and CD27dim) and a restricted pattern of expression of Cy Igλ. Other non-PC BM populations are depicted as gray dots. In turn, the top right panel shows the results of principal component analysis –automatic population separator 1 (APS1) view of principal component 1 (PC1) vs PC2—demonstrating a clearly different overall immunophenotypic profile of normal and aPCs in this sample. In this later plot, circles represent median values for all phenotypic parameters measured in the two tubes but Cy Iggs, while inner (doted) and outer (solid) lines represent the first and second s.d. of the distribution of the PC events in the multidimensional space, respectively. The table in the right illustrates the top 6 parameters contributing to the separation between nPC and aPC in the above PC1 vs PC2 plot and their percentage contribution to the separation. Please note that, in this sample, PC corresponded to 0.005% of all nucleated BM cells; in turn, aberrant PC (127 PC events) corresponded to 0.001% of the whole BM cellularity with an assay sensitivity (in the quantitative range) of < 5 x 10⁻⁶.
DISCUSSION
MRD detection in BM has proven clinically meaningful for MM monitoring after therapy, particularly to predict outcome of patients that reach CR\textsuperscript{12–16,29} independently of therapy.\textsuperscript{22,53}

Here we describe an innovative EuroFlow-based high-sensitive, standardized and validated NGF-MRD method which can be applied to virtually every MM patient for MRD monitoring in BM after therapy. Overall, our results show a similar applicability but a significantly increased sensitivity for the novel EuroFlow-NGF approach vs conventional flow-MRD, with around one fourth of all MRD-negative samples by conventional flow becoming MRD-positive by NGF. To the best of our knowledge, this is the first time that a validated high-sensitive flow-MRD assay is described with a LOD close to \(10^{-6}\) (ability to identify down to 20 tumor PCs among 10\(^7\) evaluated BM cells) and a limit of quantitation of <5 x 10\(^{-10}\) (ability to accurately quantity tumor PC percentages at levels down to five cells per million cells, that is, 0.00005%), calculated following consensus recommendations.\textsuperscript{42} Importantly, the EuroFlow-NGF approach provided similar results in different centers, which further confirms the high standardization level of the method. From the clinical point of view, despite the (still) limited follow-up time, the NGF approach already showed a significant prognostic impact on PFS of MM patients, even among those in CR/sCR, significantly improving the predictive clinical value of conventional flow-MRD.

The greater sensitivity of NGF vs conventional flow-MRD was mostly due to the use of both an optimized combination of fluorochromes and antibody reagents for increased specificity at very low MRD levels, and the 10-fold increase in the number of cells evaluated, in the context of fully standardized laboratory protocols. The two-tube approach proposed also allows confirmation in a second independent measurement, of the clonal nature of suspicious (low numbers of) PC, through evaluation of the clonal nature of suspicious (low numbers of) PC, through evaluation of the cytoplasmic k/\(\lambda\) restriction of phenotypically aberrant PC which proved to be required in a significant number of cases. Interestingly, the phenotypic markers selected as most informative

| Table 2. Distribution of aberrant (aPC) and normal (nPC) plasma cells in BM samples from MM patients in VGPR, CR and sCR with discrepant MRD results (MRD\textsuperscript{+} vs MRD\textsuperscript{−}) by NGF vs conventional flow-MRD assays |
|---|---|---|---|---|---|
| Case ID | Disease status | Conventional flow | NGF |
| | | % tPC | % nPC | % aPC | % tPC | % nPC | % aPC |
| 1 | CR | 0.01 | 0.01 | 0.008 | 0.007 | 0.004 |
| 2 | CR | 0.01 | 0.01 | 0.002 | 0.002 | 0.002 |
| 3 | CR | 0.007 | 0.007 | 0.007 | 0.007 | 0.003 |
| 4 | CR | 0.04 | 0.04 | 0.002 | 0.002 | 0.004 |
| 5 | sCR | 0 | 0 | 0.003 | 0.002 | 0.006 |
| 6 | CR | 0.1 | 0.1 | 0.1 | 0.1 | 0.002 |
| 7 | CR | 0.2 | 0.2 | 0.2 | 0.2 | 0.02 |
| 8 | sCR | 0 | 0 | 0.002 | 0.003 | 0.002 |
| 9 | CR | 0.9 | 0.9 | 0.6 | 0.6 | 0.005 |
| 10 | CR | 0.2 | 0.2 | 0.06 | 0.06 | 0.004 |
| 11 | VGPR | 0 | 0 | 0.01 | 0.01 | 0.007 |
| 12 | CR | 0 | 0 | 0.01 | 0.01 | 0.006 |
| 13 | sCR | 0.2 | 0.2 | 0.2 | 0.2 | 0.005 |
| 14 | sCR | 0.03 | 0.03 | 0.03 | 0.03 | 0.002 |
| 15 | CR | 0.1 | 0.1 | 0.07 | 0.04 | 0.03 |
| 16 | VGPR | 0 | 0 | 0.009 | 0.008 | 0.001 |
| 17 | CR | 0 | 0 | 0.004 | 0.003 | 0.001 |
| 18 | CR | 0.04 | 0.04 | 0.01 | 0.01 | 0.002 |
| 19 | CR | 0.09 | 0.08 | 0.01 | 0.05 | 0.00 |
| 20 | CR | 0.1 | 0.1 | 0.006 | 0.1 | 0.1 |
| 21 | CR | 0.4 | 0.3 | 0.02 | 0.3 | 0.03 |

Abbreviations: aPC, aberrant plasma cell; BM, bone marrow; CR, complete response; MM, multiple myeloma; MRD, minimal residual disease; NGF, next generation flow; nPC, normal plasma cell; sCR, stringent CR; tPC, total plasma cells; VGPR, very good partial response.

| Table 3. Distribution of distinct BM-associated populations as identified by the NGF antibody combination (version 5) and percentages of MRD\textsuperscript{−} and MRD\textsuperscript{+} cases with decreased levels |
|---|---|---|---|---|---|
| BM-associated cell population (phenotype\textsuperscript{a}) | Normal BM range (n=16), hemodiluted BM cutoff | % of cases with low levels | NGF MRD\textsuperscript{−} patients showing disease progression |
| | | MRD\textsuperscript{−} | MRD\textsuperscript{+} | Case 1 | Case 2 |
| % Mast cells (CD117\textsuperscript{hi}, CD45\textsuperscript{−}, CD81\textsuperscript{dim}, CD38\textsuperscript{−}, SSC/FSC\textsuperscript{int−hi}) | 0.006% (0.002–0.03%) | 10% | 5% | 0.0017% | 0.0006% |
| % NRBC (SSC/FSC\textsuperscript{lo}, all markers\textsuperscript{b}) | 6.4% (2–11.5%) | 3% | 1% | 8% | 5% |
| % CD19\textsuperscript{−} nPC (CD38\textsuperscript{lo}, CD138\textsuperscript{−}, CD45\textsuperscript{−/−}, CD81\textsuperscript{−/−}, CD27\textsuperscript{−}, SSC/FSC\textsuperscript{high}) | 0.05% (0.003–0.2%) | 14% | 14% | 0.0038% | 0.0036% |
| % CD27\textsuperscript{+} B cell precursors (CD19\textsuperscript{−}, CD38\textsuperscript{++}, CD45\textsuperscript{dim}, CD81\textsuperscript{−/−}, CD27\textsuperscript{+}, CyIg\textsuperscript{−}, SSC/FSC\textsuperscript{lo}) | 0.08% (0.004–0.4%) | 12% | 11% | 0.0001% | 0.003% |
| % CD27\textsuperscript{−} B cell precursors (CD19\textsuperscript{−}, CD38\textsuperscript{−}, CD45\textsuperscript{dim/−}, CD81\textsuperscript{−/−}, CD27\textsuperscript{−}, CyIg\textsuperscript{−}, SSC/FSC\textsuperscript{lo}) | 0.4% (0.05–2.2%) | 11% | 13% | 0.001% | 0.01 |
| % Mature B cells (CD19\textsuperscript{−}, CD38\textsuperscript{−}, CD45\textsuperscript{dim/−}, CD81\textsuperscript{−/−}, CD27\textsuperscript{−}, CyIg\textsuperscript{−}, SSC/FSC\textsuperscript{lo}) | 1.6% (0.6–3.5%) | 26% | 23% | 0.1% | 0.1% |
| % Myeloid precursors (CD117\textsuperscript{−}, CD45\textsuperscript{dim}, CD81\textsuperscript{−/−}, SSC/FSC\textsuperscript{low/low}) | 1.8% (0.2–3.6%) | 8% | 8% | 0.3% | 0.3% |
| % Endothelial and mesenchymal cells (CD81\textsuperscript{hi}, CD45\textsuperscript{−}, SSC/FSC\textsuperscript{high}) | 0.01% (0.0005–0.008%) | 1% | 0% | 0.02% | 0.02% |

Abbreviations: BM, bone marrow; FSC, forward scatter; MRD, minimal residual disease; NGF, next generation flow; nPC, normal plasma cells; NRBC, nucleated red blood cells; SSC, sideward scatter. \textsuperscript{a}Markers used in the gating strategy.
did not differ from those considered to be essential by expert consensus. However, we proved here that selection of optimal fluochrome-conjugated antibody clones per marker could not be predicted by pre-existing (shared) expertise. Thus, identification of the optimal marker combinations for the 2-tube 8-color antibody panel required five rounds of optimization of what we already considered initially to be a potentially 'optimal' panel (that is, Version 1). Major limitations of suboptimal reagents were: (i) increased non-specific and background fluorescence; (ii) too dim fluorescence intensity; (iii) specific interactions among mixed reagents and/or; (iv) suboptimal staining or reduced reactivity on nPC vs aPCs, particularly on Daratumumab-treated vs non-treated patients. Altogether, these results indicate that (extensive) prospective testing is mandatory to define optimal combinations of reagents for flow-MRD monitoring in MM, due to the problems encountered and the significantly different staining profiles obtained with distinct combinations of reagents of the same CD markers. As an example, only two of the many (n = 9) CD38 antibody clones evaluated proved efficient for detecting CD38 on PC from Daratumumab-treated MM patients (Supplementary Figure 1); even more, only one of these two clones proved to be effective for detecting CD38 on MM PC treated in vitro with the Isatuximab antibody (Supplementary Figure 1).

Although the specific combination(s) of markers used is a key factor for optimal identification of BM PC and discrimination between nPC and aPC, another critical factor in building a sensitive flow-MRD technique is the number of cells analyzed. To the best of our knowledge, this is the first report in which consistently > 5 × 10⁶ cells/patient sample (usually > 10⁷ cells) were investigated. For the most frequently used stain-lyse-and-then-wash sample preparation flow-MRD procedures, hundreds thousand to 1–2 million cells in ≤300 μl BM are analyzed. In contrast, the here described EuroFlow SOP assured acquisition of > 10⁷ events in most MM BM MRD samples, by means of staining a median of 1.5 ml of total BM sample, with a proven limit of quantitation for the NGF MRD method of < 5 × 10⁻⁶ and a LOD of < 2 cells per million. These features contribute to explain the higher sensitivity of EuroFlow-based NGF vs conventional flow-MRD, and NGS. The relatively high frequency of discrepant results (30%) in the NGF vs NGS comparison among cases with low (< 10⁻⁴) MRD levels, might be caused by suboptimal annealing of the NGS-PCR primers due to high levels of somatic hypermutations in the Ig genes of nPC and aPC, and deserves further investigation.

Independently of its potentially greater sensitivity, NGF has other additional advantages over NGS: it is faster (< 4 h), standardized and reproducible, it has a greater applicability (≥ 98%), and it does not require a diagnostic sample or patient-specific probes, which potentially lead to a lower variability in the sensitivity reached per patient. However, EuroFlow-based NGF required fresh material analyzed within 24 h after sampling; this is

![Figure 4](40x40)
feasible in virtually all countries, since the standardized EuroFlow procedures have now been implemented worldwide. In addition, the costs of the NGF reagents (~100 USD) and assay (~350 USD), are estimated to be lower than those of NGS (~350 USD and 700 USD, respectively). Importantly, NGF can also provide an overall assessment of the quality of the patient sample through identification of a significant decrease in non-PC BM cell populations (for example, CD117hi mast cells, CD45−/sideward-scatterlo nucleated red blood cells, CD117+ myeloid precursors, CD19+CD38lo/CD45lo B-cell precursors and CD19+ nPC) in hemodiluted BM samples (Table 3) and through providing full insight in the normal B-cell compartment via the identification of normal BM residual B-cell precursors (CD19+CD38+CD45+), immature B-lymphocytes (CD19+CD38−CD45lo), naive B cells (CD19+CD38+CD27−), memory B-lymphocytes (CD19+CD38−CD27+), and nPC (CD19+CD56−, CD19+CD56+ and CD19+CD56lo) in addition to myeloma PC (Supplementary Figure 3). Of note, decreased numbers of mast cells were found in 17/110 ≥VGPR BM samples, particularly among MRD cases, including the only two patients that showed disease progression (Table 3), pointing out the need for careful evaluation of MRD cases for blood contamination. Whether evaluation of an additional sample from the same patient is required in the such cases, still remains to be established.

The relatively short time needed for the complete EuroFlow-NGF procedure (<4 h) can be further reduced by implementing automatic sample preparation procedures, pre-mixed and dried antibody cocktails, and software algorithms for automatic data analysis. Such improvements are ongoing and will further contribute to prevent diagnostic errors.

A major challenge we faced during the design phase was to determine whether the 2-tube 8-color NGF approach could be replaced by a single 10-color tube to decrease reagent costs and data acquisition time, as suggested by others. Direct comparison of Version 5 of the two 8-color tube antibody panel vs three different versions of a single 10-color antibody combination (Supplementary text) showed quite comparable results for both formats. However, the 2-tube 8-color method emerged as a more robust assay because (i) higher numbers of cells were measured; (ii) the confirmatory value of the second tube in case of small populations of suspicious PC found in the first tube; and (iii) the increased consistency and precision of replicate vs single measurements.

In summary, here we describe a novel validated EuroFlow-NGF assay for high-sensitive, fast and standardized quantification of MRD in MM that overcomes previous limitations of conventional flow-MRD approaches and improves prediction of patient outcome. This method is ready-to-use and well-suited for implementation in clinical trials to establish the diagnostic role of MRD in MM.

CONFLICT OF INTEREST

E-G and RF are employees of Cytognos SL, Salamanca, Spain. JIMvD received research support and travel support from Roche, Amgen and Becton Dickinson. SB received research support from Roche, Celgene, Becton Dickinson and AbbVie as well as honoraria from Roche and AbbVie. JF-M, LS-F, OG-S, J-JP-M, AC-M, CJ, JM-L, M-VM, VvdV, JC, LS, NP, M-BV, RGS, MP, M-CdcC, JB, J-JL, CA, AB, AG-M, JL, PL, CA-S, JS-M, BD and AD declare no conflict of interest.

ACKNOWLEDGEMENTS

This work has been supported by the International Myeloma Foundation-Black Swan Research Initiative, the Red Temática de Investigación Cooperativa en Cáncer (RTICC); grant SA079U14 from the Consejería de Educación, Junta de Castilla y León, Valladolid, Spain and; grant DSTS15/00119 from Instituto de Salud Carlos III, Ministerio de Economía y Competitividad, Madrid, Spain.

REFERENCES

1. Ocío EM, Richardson PG, Rajkumar SV, Palumbo A, Mateos MV, Orlowski R et al. New drugs and novel mechanisms of action in multiple myeloma in 2013: a report from the International Myeloma Working Group (IMWG). Leukemia 2014; 28: 525–542.

2. Bianchi G, Richardson PG, Anderson KC. Promising therapies in multiple myeloma. Blood 2015; 126: 300–310.

3. Lonial S, Dure B, Palumbo A, San-Miguel J. Monoclonal antibodies in the treatment of multiple myeloma: current status and future perspectives. Leukemia 2016; 30: 526–535.

4. Rajkumar SV, Kumar S. Multiple myeloma: diagnosis and treatment. Mayo Clin Proc 2016; 91: 101–119.

5. Lahuerta JJ, Mateos MV, Martínez-Lopez J, Rosinol L, Sureda A, de la Rubia J et al. Influence of pre- and post-transplantation responses on outcome of patients with multiple myeloma: sequential improvement of response and achievement of complete response are associated with longer survival. J Clin Oncol 2006; 24: 5775–5782.

6. Sonneveld P, Goldschmidt H, Rosinol L, Blade J, Lahuerta JJ, Cavo M et al. Bortezomib-based versus nonbortezomib-based induction treatment before autologous stem-cell transplantation in patients with previously untreated multiple myeloma: a meta-analysis of phase III randomized, controlled trials. J Clin Oncol 2013; 31: 3279–3287.

7. Usmani SZ, Crowley J, Hoering A, Mitchell A, Waheed S, Nair B et al. Improvement in long-term outcomes with successive total therapy trials for multiple myeloma: are patients now being cured? Leukemia 2013; 27: 226–232.

8. Palumbo A, Biringer S, Larocca A, Rossi D, Di Raimondo F, Magarotto V et al. Bortezomib-melphalan-prednisone-thalidomide followed by maintenance with bortezomib-thalidomide compared with bortezomib-melphalan-prednisone for initial treatment of multiple myeloma: updated follow-up and improved survival. J Clin Oncol 2014; 32: 634–640.

9. Mateos MV, Leleu X, Palumbo A, San Miguel JF. Initial treatment of transplant-ineligible patients in multiple myeloma. Expert Rev Hematol 2014; 7: 67–77.

10. Rollig C, Knop S, Bornhäuser M. Multiple myeloma. Lancet 2015; 385: 2197–2208.

11. Mateos MV, Ocío EM, Paiva B, Rosinol L, Martínez-Juárez J, Blade J et al. Treatment for patients with newly diagnosed multiple myeloma in 2015. Blood Rev 2015; 29: 387–403.

12. Fulciniti M, Munshi NC, Martínez-Juárez J. Deep response in multiple myeloma: a critical review. BioMed Res Int 2015; 2015: Article ID 832049.

13. Lonial S, Anderson KC. Association of response endpoints with survival outcomes in multiple myeloma. Leukemia 2014; 28: 258–268.

14. Gay F, Larocca A, Wijermans P, Cavallo F, Rossi D, Schaadusa R et al. Complete response correlates with long-term progression-free and overall survival in elderly myeloma treated with novel agents: analysis of 1175 patients. Blood 2011; 117: 3025–3031.

15. Barlogie B, Mitchell A, van Rhee F, Epstein J, Morgan GJ, Crowley J. Curing myeloma at last: defining criteria and providing the evidence. Blood 2014; 124: 3043–3051.

16. Paiva B, van Dongen JJ, Orfao A. New criteria for response assessment: role of minimal residual disease in multiple myeloma. Blood 2015; 125: 3059–3068.

17. Paiva B, Gutierrez NC, Rosinol L, Vidriales MB, Montalban MA, Martinez-Juárez J et al. High-risk cytogenetics and persistent minimal residual disease by multiparameter flow cytometry predict unsustained complete response after autologous stem cell transplantation in multiple myeloma. Blood 2012; 119: 687–691.

18. Paiva B, Martinez-Juárez J, Vidriales MB, Mateos MV, Montalbán MA, Fernandez-Redondo E et al. Comparison of immunofixation, serum free light chain, and immunophenotyping for response evaluation and prognostication in multiple myeloma. J Clin Oncol 2011; 29: 1627–1633.

19. Paiva B, Montalban MA, Puig N, Cordon L, Martinez-Juárez J, Ocío EM et al. Clinical significance of sensitive Flow-MRD monitoring in elderly multiple myeloma patients on the Pethema/GEM2010MAS65 trial. Blood 2014; 124: 3390.

20. Rawstron AC, Child JA, de Tute RM, Davies FE, Gregory WM, Bell SE et al. Minimal residual disease assessed by multiparameter flow cytometry in multiple myeloma: impact on outcome in the Medical Research Council Myeloma IX Study. J Clin Oncol 2013; 31: 2540–2547.

21. Rawstron AC, Davies FE, DasGupta R, Ashcroft AJ, Patmore R, Drayson MT et al. Flow cytometric disease monitoring in multiple myeloma: the relationship between normal and neoplastic plasma cells predicts outcome after transplanta
tion. Blood 2002; 100: 3095–3100.

22. Rawstron AC, Gregory WM, de Tute RM, Davies FE, Bell SE, Drayson MT et al. Minimal residual disease in myeloma by flow cytometry: independent prediction of survival benefit per log reduction. Blood 2015; 125: 1932–1935.

23. Robillard N, Bene MC, Moreau P, Wullemme S. A single-tube multiparameter seven-color flow cytometry strategy for the detection of malignant plasma cells in multiple myeloma. Blood Cancer J 2013; 3: e134.
Supplementary Information accompanies this paper on the Leukemia website (http://www.nature.com/leu)