ARTÍCULO 1

"Citometría de flujo de nueva generación para la detección sensible y estandarizada de enfermedad mínima residual en el mieloma múltiple"

Juan Flores-Montero*[1], **Luzalba Sanoja Flores***[1], Bruno Paiva*[2], Noemí Puig[3], Omar García-Sánchez[3], Sebastian Böttcher[4], Vincent H.J. van der Velden[5], José-Juan Pérez-Morán[3], María-Belén Vidriales[3], Ramón García-Sanz[3], Cristina Jimenez[3], Marcos González[3], Joaquín Martinez-López[6], Alba Corral Mateos[1], Georgiana-Emilia Grigore[7], Rafael Fluxá[7], Roberia Pontes[8], Joana Caetano[9], Lukasz Sedek[10], María-Consuelo del Cañizo[3], Joan Bladé[11], Juan-José Lahuerta[6], Carlos Aguilar[12], Abelardo Bárez[13], Aránzazu García-Mateo[14], Jorge Labrador[15], Pilar Leoz[1], Carmen Aguilera-Sanz[16], Jesús San-Miguel**[2], María-Victoria Mateos**[3], Brian Durie***[17], Jacques J.M. van Dongen***[5,18] y Alberto Orfao***[1] en nombre del Consorcio EuroFlow y de la Fundación Internacional del Mieloma.

*Estos autores han contribuido de igual manera al trabajo realizado y por tanto cada uno de ellos debe ser considerado como primer autor.
**Ambos autores han contribuido de igual manera a este trabajo y por tanto ambos deben ser considerados como penúltimo autor.
***Estos autores han contribuido de igual manera a este trabajo y por tanto cada uno de ellos debe ser considerado como último autor.

[1]Centro de Investigación del Cáncer (IBMCC-CSIC/USAL-IBSAL); Servicio General de Citometría (NUCLEUS) y Departamento de Medicina, Universidad de Salamanca, Salamanca, España. Centro de Investigación Biomédica en Red de Cáncer, Instituto Carlos III, Madrid, España; [2]Clínica Universidad de Navarra; Centro de Investigación de Medicina Aplicada (CIMA), IDISNA, Pamplona, España (UNAV); [3]Servicio de Hematología, Hospital Universitario de Salamanca, IBSAL, IBMCC (USAL-CSIC), Salamanca, España (HUSAL); [4]Segundo Departamento de Medicina, Hospital Universitario de Schleswig-Holstein, Campus de Kiel, Kiel, Alemania (UNIKIEL); [5]Departamento de Inmunología, Erasmus MC, Centro Médico Universitario de Rotterdam, Rotterdam, Países Bajos (EMC); [6]Servicio de Hematología, Hospital 12 de Octubre, Madrid, España; [7]Cytognos SL, Salamanca, España; [8]Servicio de Pediatría, Instituto de Pediatría y Puericultura, Universidad Federal de Río de Janeiro, Río de Janeiro, Brasil; [9]Servicio de Hematología, Instituto Portugués de Oncología, Lisboa, Portugal (IPOLFG); [10]Departamento de Hematología y Oncología Pediátrica, Universidad Médica de Silesia en Katowice, Zabrze, Polonia (SUM); [11]Servicio de Hematología, Hospital Clinic I Provincial, Barcelona, España; [12]Servicio de Hematología, Hospital General de Santa Bárbara, Soria, España; [13]Servicio de Hematología, Complejo Asistencial de Ávila, Ávila, España; [14]Servicio de Hematología, Complejo Asistencial de Segovia, Segovia, España; [15]Servicio de Hematología, Hospital Universitario de Burgos, Burgos, España; [17]Centro del Cáncer Cedars-Sinai Samuel Oschin, Los Angeles, Estados Unidos; [18]Servicio de Inmunohematología y Hemoterapia, Centro Médico Universitario de Leiden, Leiden, Países Bajos.

Leukemia (2017), 1–10
DOI: 10.1038/leu.2017.29

La información suplementaria correspondiente a este trabajo y citada en el mismo está disponible en formato electrónico en la página web de la revista, así como en el Anexo I y en el disco compacto disponibles al final de la presente memoria de tesis doctoral.

Introducción. La citometría de flujo constituye uno de los principales métodos empleados en la evaluación de la calidad de la respuesta al tratamiento y la monitorización de enfermedad mínima residual (EMR) en médula ósea (MO) de pacientes con mieloma múltiple (MM). Pese a sus ventajas, la detección de EMR mediante las técnicas convencionales de inmunofenotipado por citometría de flujo presenta niveles de sensibilidad inferiores a los de las técnicas moleculares y carece de estandarización.

Objetivo. En este trabajo nos hemos propuesto diseñar y estandarizar un método innovador y altamente sensible de citometría de flujo de nueva generación para la detección de EMR en MO de pacientes con MM tras tratamiento. Así mismo, validaremos la aplicabilidad, sensibilidad y especificidad de esta nueva herramienta inmunofenotípica frente a otras técnicas empleadas para el estudio de EMR, como la citometría de flujo convencional de 8-colores y herramientas moleculares de secuenciación masiva (de nueva generación) basadas en la identificación de reordenamientos del gen de las inmunoglobulinas.

Materiales y métodos. En total, estudiamos 375 muestras de MO y 10 muestras de sangre periférica procedentes de 332 pacientes con neoplasias de células plasmáticas (CP) y de 53 controles sanos, distribuidas entre las distintas etapas del trabajo. En un primer paso, diseñamos un panel de anticuerpos seleccionando los marcadores asociados a fenotipo de CP, que ofrecían mayor capacidad resolutiva para la identificación y discriminación entre las CP normales *vs.* tumorales, según su perfil inmunofenotípico. En paralelo, optimizamos un nuevo protocolo de preparación y tinción de muestras con el que fuese posible incrementar el número de células a analizar, aplicando, además, estrategias informáticas que permitiesen la identificación y cuantificación automática de las poblaciones celulares presentes en MO. Finalmente, validamos el diseño global de la nueva herramienta de citometría de flujo de nueva generación frente a la técnica de citometría de flujo convencional, en una amplia cohorte de muestras de MO de pacientes con MM (n=228), de las que 110 muestras procedían de pacientes que habían alcanzado al menos muy buena respuesta parcial al tratamiento administrado. Además, en un grupo de estos casos (n=31) analizamos también la presencia *vs.* ausencia de EMR en MO respecto a la técnica de secuenciación masiva.

Resultados. El desarrollo de la nueva técnica de detección de EMR basada en citometría de flujo de nueva generación descrita en este trabajo, se apoyó en tres pilares fundamentales: i) la optimización de un panel de anticuerpos para la detección de CP tumorales, compuesto por dos

tubos de ocho colores tras cinco ciclos de diseño, evaluación y rediseño de las combinaciones testadas, ii) el establecimiento de un nuevo protocolo de preparación y tinción de muestras que permite el análisis de $\geq 10^7$ células/muestra mediante un paso inicial de lisis masiva de hematíes, y iii) el desarrollo de herramientas informáticas innovadoras para la identificación y clasificación automática de poblaciones celulares de MO, incluidas las CP normales y patológicas.

La validación del nuevo procedimiento inmunofenotípico de detección de EMR, se llevó a cabo mediante un estudio multicéntrico sobre 110 muestras de MO de pacientes con MM que habían alcanzado una (muy buena) respuesta parcial o respuesta completa (RC) al tratamiento administrado. En este sentido, la técnica de citometría de flujo de nueva generación desarrollada mostró una mayor tasa de casos con EMR positiva respecto a la detectada con la citometría de flujo convencional de 8 colores (47% vs 34%, respectivamente; p=0,003), al identificar la presencia de EMR en un grupo de 18 pacientes con niveles de EMR indetectable mediante citometría convencional. De forma similar, la frecuencia de casos con EMR positiva mediante citometría de flujo de nueva generación fue también superior a la detectada mediante secuenciación masiva (70% vs 52%, respectivamente; p=0,06). Esta mayor sensibilidad alcanzada con la técnica inmunofenotípica de EMR puesta a punto en este trabajo se tradujo, a su vez, en una supervivencia libre de progresión significativamente más prolongada para aquellos pacientes que estando en RC, no presentaron EMR en MO mediante citometría de flujo de nueva generación, frente a aquellos en los que pese a estar en RC se les detectó persistencia de CP tumorales residuales en MO (EMR positiva).

Conclusión. La nueva estrategia desarrollada en este trabajo para la detección de EMR en MO de pacientes con MM posteriormente al tratamiento, basada en citometría de flujo de nueva generación, constituye un método más sensible y estandarizado que las técnicas convencionales de citometría de flujo, superando gran parte de las limitaciones actuales de estas últimas. Además, la presencia *vs.* ausencia de EMR en MO mediante la técnica desarrollada mostró un importante impacto en la predicción de la supervivencia libre de progresión de pacientes con MM que han alcanzado una buena respuesta al tratamiento, representando así una alternativa robusta frente a los métodos disponibles en la actualidad, lista para ser implementada en la rutina diagnóstica.

OPEN

Leukemia (2017), 1–10

www.nature.com/leu

ORIGINAL ARTICLE

Next Generation Flow for highly sensitive and standardized detection of minimal residual disease in multiple myeloma

J Flores-Montero[1,19], L Sanoja-Flores[1,19], B Paiva[2,19], N Puig[3], O García-Sánchez[3], S Böttcher[4], VHJ van der Velden[5], J-J Pérez-Morán[3], M-B Vidriales[3], R García-Sanz[3], C Jimenez[3], M González[3], J Martínez-López[6], A Corral-Mateos[1], G-E Grigore[7], R Fluxá[7], R Pontes[8], J Caetano[9], L Sedek[10], M-C del Cañizo[3], J Bladé[11], J-J Lahuerta[6], C Aguilar[12], A Bárez[13], A García-Mateo[14], J Labrador[15], P Leoz[1], C Aguilera-Sanz[16], J San-Miguel[2,20], M-V Mateos[3,20], B Durie[17,21], JJM van Dongen[5,18,21] and A Orfao[1,21]

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 $\geqslant 10^7$ 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 advance online publication, 10 March 2017; doi:10.1038/leu.2017.29

INTRODUCTION

In recent years, the landscape of treatment of multiple myeloma (MM) has substantially changed[1–4] leading to significantly increased complete response (CR) rates and survival.[5–11] Therefore, CR has become a major goal in newly diagnosed MM, even in aged patients.[5,12–15] Despite these advances, most CR patients will ultimately relapse.[7,10] 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).[16]

Conventional 4–8-color flow cytometry,[17–24] and to a lesser extent also allele-specific oligonucleotide quantitative PCR (ASOqPCR) and next generation sequencing (NGS),[25–29] are progressively being used to monitor MRD in bone marrow (BM) of MM after therapy.[16,30] These studies confirmed the relevance of MRD measurements for identification of MM patients at higher risk of relapse.[16,22] 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.[16,31]

High sensitivity and broad applicability have both become mandatory requirements for MRD monitoring in MM.[16] Early studies have shown that conventional 4–6-color flow-MRD is applicable in virtually all MM patients ($\geqslant$ 95%), whereas ASOqPCR 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.[16,25,32–34] However, the sensitivity of conventional flow-MRD ($< 10^{-4}$) remains (systematically) lower

[1]Cancer Research Center (IBMCC-CSIC/USAL-IBSAL); Cytometry Service (NUCLEUS) and Department of Medicine, University of Salamanca, Salamanca, Spain (USAL); [2]Clinica Universidad de Navarra; Applied Medical Research Center (CIMA), IDISNA, Pamplona, Spain (UNAV); [3]Department of Hematology, University Hospital of Salamanca, IBSAL; IBMCC (USAL-CSIC), Salamanca, Spain (HUSAL); [4]Second Department of Medicine, University Hospital of Schleswig-Holstein, Campus Kiel, Kiel, Germany (UNIKIEL); [5]Department of Immunology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands (EMC); [6]Department of Hematology, Hospital 12 de Octubre, Madrid, Spain; [7]Cytognos SL, Salamanca, Spain; [8]Faculty of Medicine, Federal University of Rio de Janeiro and Institute of Pediatrics and Childhood Care, Rio de Janeiro, Brazil; [9]Department of Hematology, Portuguese Institute of Oncology, Lisbon, Portugal (IPOLFG); [10]Department of Pediatric Hematology and Oncology, Medical University of Silesia in Katowice, Zabrze, Poland (SUM); [11]Department of Hematology, Hospital Clinic I Provincial, Barcelona, Spain; [12]Department of Hematology, Hospital General de Santa Bárbara, Soria, Spain; [13]Department of Hematology, Complejo Asistencial de Ávila, Ávila, Spain; [14]Department of Hematology, Complejo Asistencial de Segovia, Segovia, Spain; [15]Hematology Department, Hospital Universitario de Burgos, Burgos, Spain; [16]Department of Hematology, Hospital El Bierzo, Ponferrada, Spain; [17]Cedars-Sinai Samuel Oschin Cancer Center, Los Angeles, CA, USA and [18]Department of Immunohematology and Blood Transfusion, Leiden University Medical Center, Leiden, The Netherlands. Correspondence: Professor A Orfao, Centro de Investigación del Cáncer (CSIC-USAL), Avda. Universidad de Coimbra S/N, Campus Miguel de Unamuno, Salamanca 37007, Spain.
E-mail: orfao@usal.es
[19]These authors have all contributed equally to this study and should be considered as first authors.
[20]These authors have all contributed equally to this study and both should be considered as previous-to-last authors.
[21]These authors have all contributed equally to this study and the three should be considered as last authors.
Received 18 November 2016; revised 24 December 2016; accepted 3 January 2017; accepted article preview online 20 January 2017

than that of ASOqPCR and NGS ($<10^{-4}$–10^{-6}).[25,26,29,35] 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 $<10^{-4}$ and $<10^{-5}$-, leading to a significantly improved prediction of outcome.[24,36] However, current flow-MRD assays and NGS still suffer from a major limitation: lack of standardization.[37] Actually, different markers and antibody panels, distinct numbers of cells measured, and highly variable criteria for MRD positivity are currently applied worldwide.[38,39] Therefore, standardization efforts have been made and consensus recommendations and guidelines have been recently proposed.[40–42] 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.[39]

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[24] 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 automatic identification of plasma cells (PC) against reference databases of normal and patient BM.[43–46] 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 PCD samples studied at diagnosis and 162 MM patient samples investigated during follow-up, including 110 follow-up BM samples from MM patients evaluated at $\geqslant$very good partial response (VGPR): 39 VGPR; 52 CR; and 19 sCR cases-. Twenty-one $\geqslant$VGPR BM samples and 10 additional BM samples -22/31 with very low MRD levels ($\leqslant 10^{-4}$)- were further used to blindly compare NGF vs NGS (Supplementary Table 1).

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, UNIKIEL, IPOLFG, 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, Cylgκ, Cylgλ and β2-microglobulin), using previously described EuroFlow sample preparation protocols.[44,45] For data analysis, events from both 8-color tubes (per sample) were merged and the values of all parameters measured per tube were mathematically calculated for the individual PC events using the merge and calculation functions of the Infinicyt software (Cytognos SL, Salamanca, Spain).[43–45] Subsequently, phenotypic data on normal PC (nPC) from 31 normal/reactive BM samples were merged in a reference database. PC were identified based on their unique pattern of expression of CD38, CD138, CD45 and light scatter features, following consensus recommendations.[40–42] Principal component analysis[44,45] of single PC events was used to compare the reference database of nPC vs aberrant/clonal PC (aPC) from each of 63 MM BM studied at diagnosis, to identify the most discriminating markers that distinguished between (reference) nPCs and aPC 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.

Results on the comparison between antibody reagents directed against the same marker (Supplementary Table 2), was based on their staining profiles on nPC vs aPC (vs other non-PC BM populations) as defined by median fluorescence intensity (range: 0–262,144 arbitrary units) and stain index values, as previously described[47] (Supplementary text). The EuroFlow reagent evaluation criteria[44] were used to discard or accept individual reagents: (i) increased background fluorescence; (ii) dim fluorescence of PCs vs other BM populations (CD38, CD138, Cylgκ$^+$ or Cylgλ$^+$) 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,[44] and (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.

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).[24] 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-PacB CD138-OC515 CD38-FITC CD56-PE CD27-PerCPCy5.5 CD19-PECy7 CD117-APC CD81-APCH7), as previously described.[24] In turn, for the NGF approach a median volume of 1.5 ± 1.3 ml (range: 0.1–5.3 ml) was employed adding up to a median total sample volume of 1.8 ml (maximum of 5.6 ml). PC populations that coexisted in individual BM samples were identified based on a combination of the CD38, CD138, CD45 PC-associated markers and light scatter characteristics, the presence vs absence of myeloma-associated phenotypes, plus Cylg light chain restriction in case of NGF, as described elsewhere.[48] According to consensus recommendations,[42] the limit of quantitation and the LOD of the NGF MRD method was calculated at $<5\times10^{-6}$ and $<2\times10^{-6}$ aPC, based on the identification of $\geqslant50$ and $\geqslant20$ aPC among 10^7 events, respectively. More detailed information about instrument conditions, data acquisition and analysis, and the specific reagents used in the present study is provided as Supplementary Material.

Automatic identification and enumeration of aPC was performed in 110 MM BM follow-up (VGPR or CR/sCR) samples using the automatic gating function of the Infinicyt software and previously described procedures,[49,50] and the results were compared against the conventional expert-guided PC-identification/gating approach. For automatic gating, a database consisting of a subset of 14 normal BM samples stained with Version 5 of the antibody panel was constructed and used.[50]

In a subset of 31 MM follow-up BM samples with low MRD levels (for example, $\leqslant10^{-4}$) in which enough DNA was available, MRD was also evaluated by NGS. For this purpose, patient-specific *VDJH* rearrangements were amplified and directly sequenced from DNA extracted from diagnostic samples using the DNAzol reagent (MRC, Cincinnati, OH, USA) and *IGHV* family-specific primers that covered framework regions 1 (FR1) and FR2, plus a JH consensus primer, as described elsewhere.[26,32] *VDJH* rearrangements identified at diagnosis were used as MRD-targeted sequences for subsequent follow-up samples. Follow-up DNA samples were amplified using the LymphoTrack *IGH* Assay (InVivoScribe Technologies, San Diego, CA, USA) and sequenced in an Illumina MiSeq platform (Illumina, San Diego, CA, USA). To all reactions, a known amount of DNA from the MWCL-1 cell line was added as reference control for cell enumeration. The Fastq files generated were analyzed with the LymphoTrack/MiSeq Software (InVivoScribe/Illumina). The number of MRD cells was calculated from the number of reads for the diagnostic *VDJH* target rearrangements and the number of reads of the reference rearrangement;

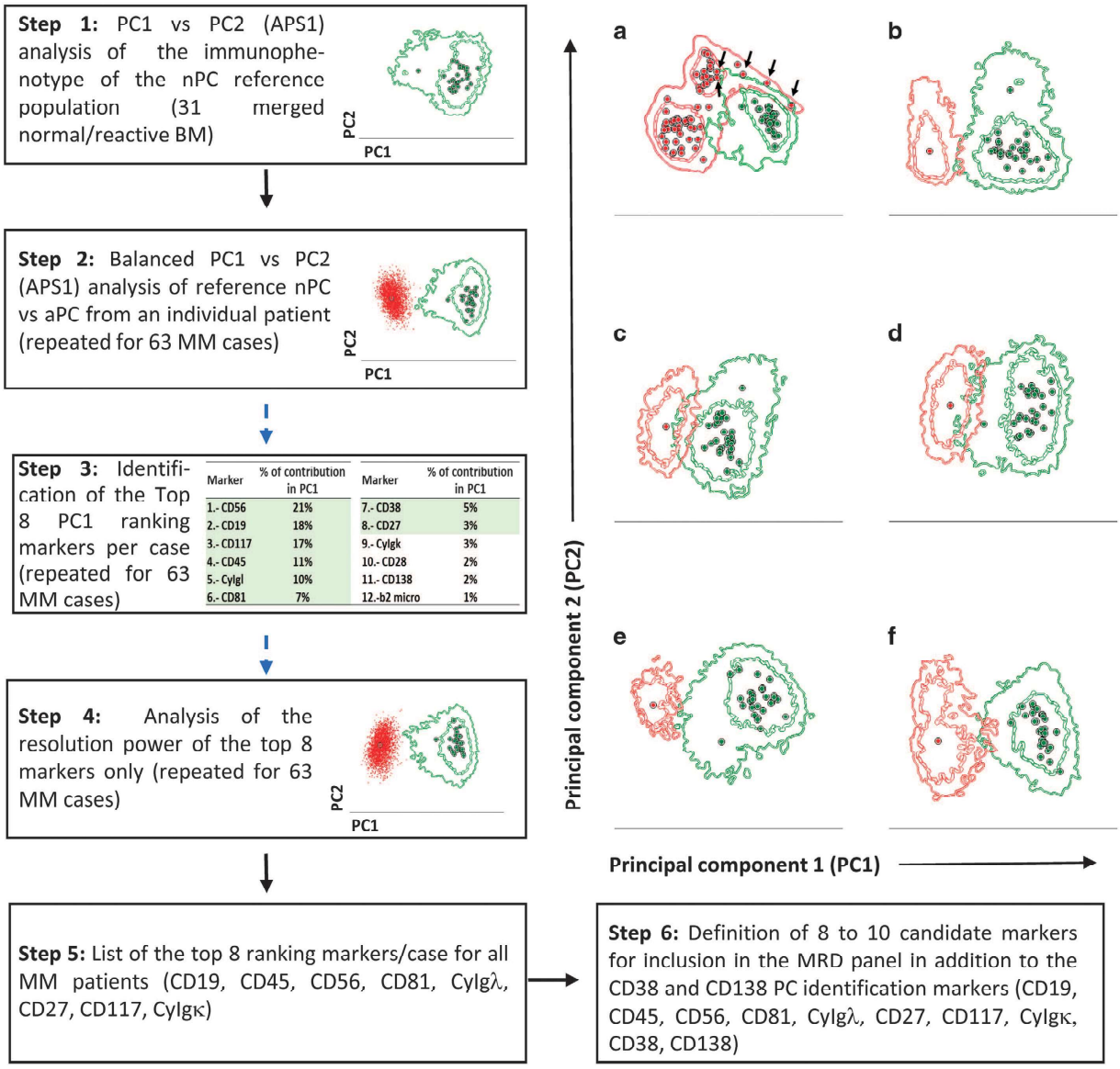


Figure 1. Diagram illustrating the process used for the selection and evaluation of markers for the NGF MM MRD panel. Left and lower right panels show the sequential steps followed to select those markers providing the best resolution between aPCs and nPC, including principal component 1 (PC1) vs PC2 —automatic population separator (APS1)— plots illustrating the described comparisons for steps 1, 2 and 4, respectively. (right; **a**) An APS1 plot corresponding to the simultaneous evaluation of 31 normal/reactive BM samples vs all 63 MM patients studied at diagnosis, in which the resolution power of the EuroFlow diagnostic PCD antibody panel combination is illustrated. Please note that PC from five samples (highlighted by the black arrows) showed suboptimal separation in the overall comparison; nonetheless, when individually compared vs the normal/reactive reference PC pool (**b**–**f**), these cases also showed sufficient phenotypic discrimination from nPC. Markers contributing to PC1 (and their percentage contribution) for each panel on the right include: (**a**) CD19(20%), CD56(17%), CD81(13%), CD45(11%), CD117(10%), CD27(8%), Cylgλ(5%), CD38(5%), Cylgκ(4%), β2 micro(3%), CD138(3%), CD28(1%); (**b**) Cylgκ(27%), Cylgλ(13%), β2 micro(12%), CD38(11%), CD56(10%), CD138(7%), CD28(16%), CD45(4%), CD27(4%), CD117(4%), CD19(2%), CD81(0%); (**c**) Cylgκ(20%), CD45 (16%), CD56(11%), CD28(10%), CD19(10%), CD27(8%), CD117(6%), CD81(6%), CD38(6%), Cylgλ(4%), CD138(3%), β2 micro(0%); (**d**), CD19(29%), β2 micro(13%), Cylgκ(13%), CD56(13%), CD45(11%), CD38(5%), CD27(4%), Cylgλ(3%), CD28(3%), CD138(2%), CD117(2%), CD81(2%); (**e**) CD19 (23%), β2 micro(13%), CD45(12%), Cylgλ(11%), CD81(9%), CD56(7%), Cylgκ(6%), CD38(6%), CD28(4%), CD138(4%), CD27(4%), CD117(1%); (**f**) CD19(20%), CD117(19%), CD81(14%), CD45(13%), CD56(12%), Cylgλ(9%), β2 micro(6%), CD38(3%), CD27(2%), CD28(1%), CD138(1%), Cylgκ (0%). In all PC1 vs PC2 plots, solid circles represent median values for the 12 fluorescence-associated parameters evaluated, inner (dotted) and outer (solid) lines represent the first and second standard deviations for individual PC. nPC populations are depicted in green while aPC are shown as red dots, circles and lines, respectively.

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 $\geqslant 2$ 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% of cases); CD45 (89%); CD56 (86%); CD81 (86%); CyIgλ (73%); CD27 (71%); CD117 (60%); and CyIgκ (56%). Re-evaluation of the utility of the combination of these eight top 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, CyIgκ and CyIgλ 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 (κ vs λ) 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/CyIgλ-APCH7).

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-APCH7 and CyIgλ-APCH7 reagents were both replaced by CD81-APCC750 and CyIgλ-APCC750. For Version 4, the CD138-HV500C reagent found to be still suboptimal was replaced, together with the CD27-PerCPCy5.5 and CD45-PacB fluorochrome positions, by CD138-HV450, CD27-HV500C and CD45-PerCPCy5.5, 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-lysing-fixation step (protocol A1 in Supplementary Figure 2), showed a significantly higher proportion of debris and dead cells ($P < 0.05$

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*	*PerCPCy5.5*	*PECy7*	*APC*	*APCH7*
1	**1**	CD45	CD138	CD38	CD56	CD27	CD19	CD117	CD81
	2	CD45	CD138	CD38				Cylgκ	Cylgλ
			HV500C						
2	**1**	CD45	*CD138*	CD38	CD56	CD27	CD19	CD117	CD81
	2	CD45	*CD138*	CD38	*CD56*	*CD229*	*CD19*	Cylgκ	Cylgλ
									APC C750
3	**1**	CD45	CD138	CD38	CD56	CD27	CD19	CD117	*CD81*
	2	CD45	CD138	CD38	CD56	CD229	CD19	Cylgκ	*Cylgλ*
		HV450							
4	**1**	*CD138*	*CD27*	CD38	CD56	*CD45*	CD19	CD117	CD81
	2	*CD138*	*CD27*	CD38	CD56	*CD45*	CD19	Cylgκ	Cylgλ
		BV421	***BV510***						
5	**1**	*CD138*	CD27	*CD38 (ME)*	CD56	CD45	CD19	CD117	CD81
	2	*CD138*	CD27	*CD38 (ME)*	CD56	CD45	CD19	Cylgκ	Cylgλ

10-color panel version	*BV421*	*BV510*	*BV605*	*FITC*	*PE*	*PerCPCy5.5*	*PECy7*	*APC*	*APCAF700/AF700*	*APC C750*
1	CD138	CD27	Cylgλ	CD38	CD56	CD45	CD19	CD117	Cylgκ	CD81
2	CD138	CD27	*CD117*	CD38	CD56	CD45	CD19	*Cylgκ*		*Cylgλ*
3	CD138	CD27	*CD45*	*CD38 (ME)*	CD56	*Cylgλ*	CD19	*CD117*	*CD81*	*Cylgκ*

Abbreviations: APC, allophycocyanin; A700, alexa fluor 700; BV421, brilliant violet 421; BV510, brilliant violet 510; BV605, brilliant violet 605; Cy, cytoplasmic; C750, C750 dye; Cy5.5, cyanin5.5; Cy7, cyanin7; 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.

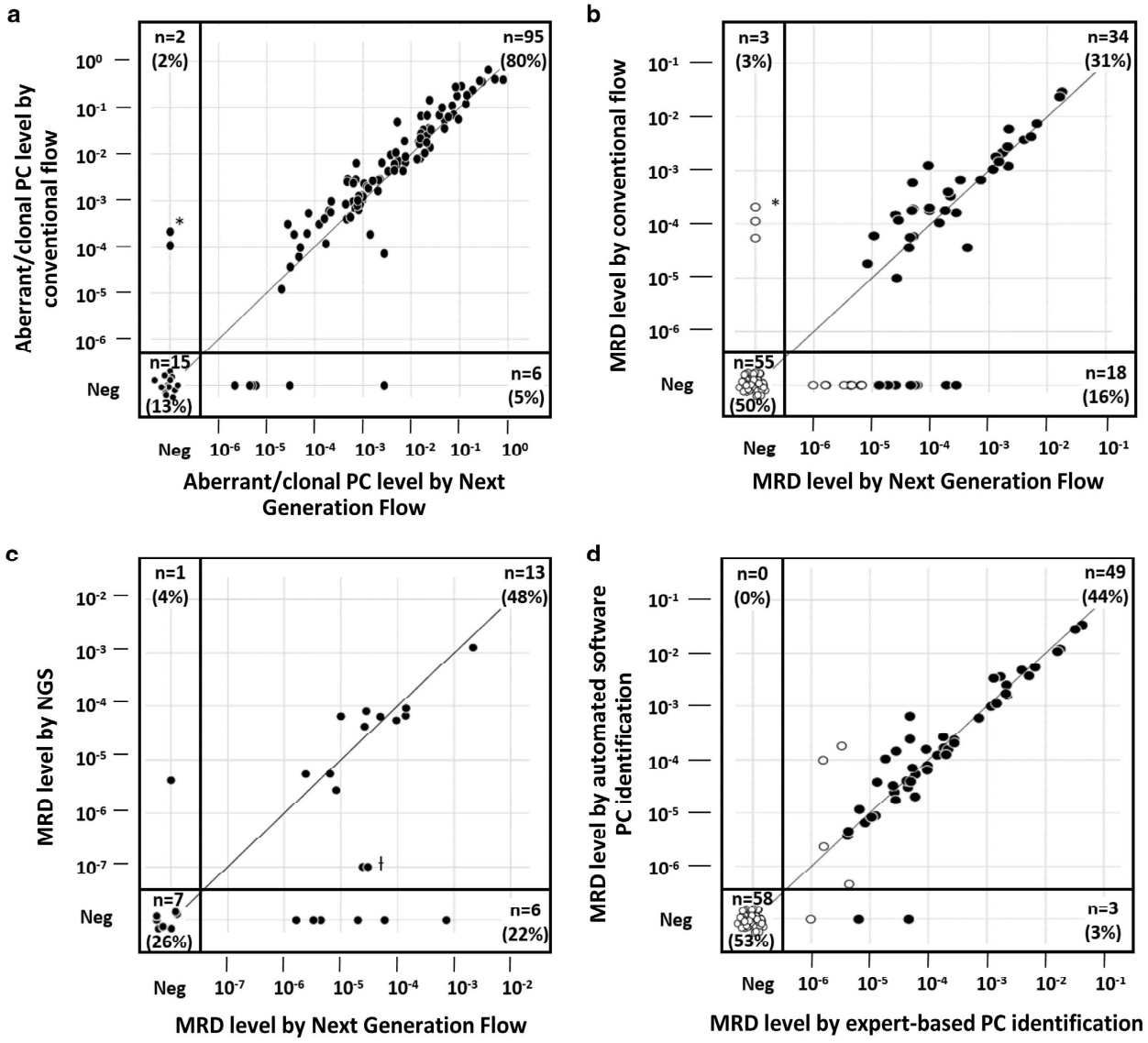
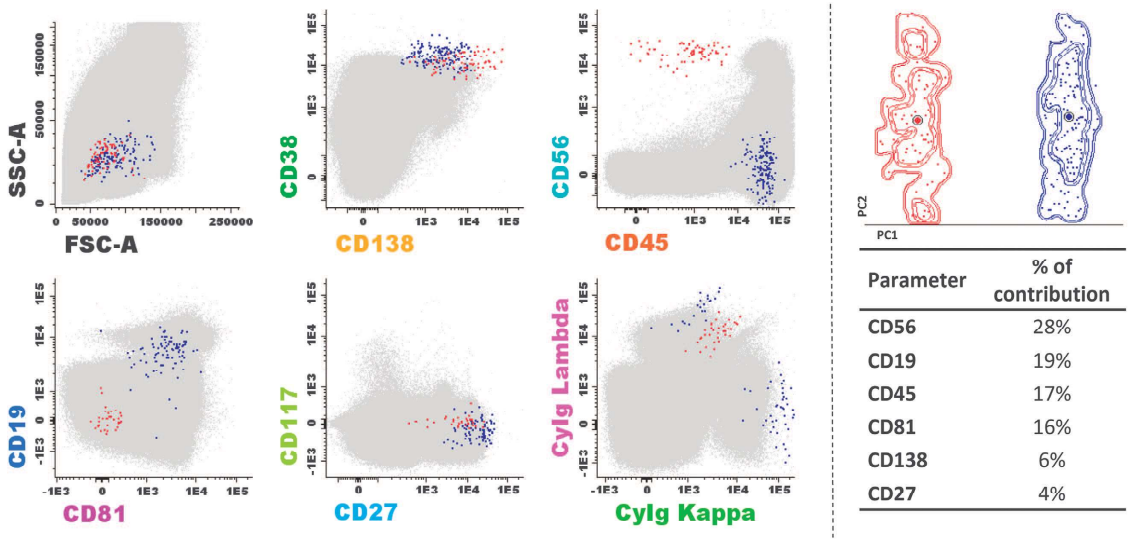


Figure 2. Validation of the new NGF method for MRD detection in MM against both conventional 8-color flow-MRD (**a**, **b**) and NGS (**c**), including expert-based vs automated NGF-MRD data analysis (**d**). In **a**, the comparison between NGF and conventional flow-MRD is shown for diagnostic and follow-up 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). [I]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.

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 $\geqslant 10^7$ 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 with relatively high tumor burden from patients in partial response, stable disease and progressive disease ($\rho = 0.96$; $P < 0.001$; Figure 2a). Most importantly, a fairly good overall correlation was also observed among cases in VGPR and CR/sCR ($\rho = 0.77$; $P < 0.001$; Figure 2b), albeit significantly different rates of MRD^+ samples were detected with both methods: 37/110 (34%) for conventional flow-MRD vs 52/110 (47%) for EuroFlow-NGF, respectively ($P = 0.003$; Figure 2b). This was due to a relatively



Parameter	% of contribution
CD56	28%
CD19	19%
CD45	17%
CD81	16%
CD138	6%
CD27	4%

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 ($>10^7$ 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) Igκ vs Cylgλ 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 (CD138^{hi}, CD38^{dim}, CD19$^-$, CD81$^-$, CD117$^-$ and CD27^{dim}) and a restricted pattern of expression of Cylgλ. 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 Cylgs, 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 $<5\times10^{-6}$.

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 0.001% (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 the participating centers with respect to rate and type of MRD discrepant cases ($P=0.63$). Importantly, in all but 7/110 cases, $>7\times10^6$ cells were evaluated (median 10.4×10^6 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 CD117^{hi} 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 cases ($\rho=0.96$; $P<0.001$; Figure 2d). However, due to the minimum number of events required by the software algorithm, aPC were not identified in 3/110 cases (2.7%) with low MRD levels by NGF (and MRD$^-$ by conventional flow-MRD) -median % aPC (range): 0.0006% (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 $\leqslant10^{-4}$), a good correlation was found between the percentage of residual aPC by NGF and NGS ($\rho=0.62$; $P=0.001$; 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 -mean 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 (quantifiable), 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.

DISCUSSION

MRD detection in BM has proven clinically meaningful for MM monitoring after therapy, particularly to predict outcome of patients that reach CR[12–16,29] independently of therapy.[22,53] Currently, several different flow- and PCR-based MRD approaches are available.[30] Flow-MRD has clear advantages vs PCR-based methods because of its relative simplicity, high speed, greater clinical applicability and worldwide availability.[16,25] However, major concerns have been recently raised about the lack of standardization of flow-MRD in MM[38] and its lower sensitivity vs PCR-based approaches, particularly NGS.[25,29,53–55]

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 $< 5 \times 10^{-6}$ (ability to accurately quantity tumor PC percentages at levels down to five cells per million cells, that is, 0.0005%), calculated following consensus recomendations.[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 cytoplasmic κ/λ 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^+ vs MRD^-) by NGF vs conventional flow-MRD assays

Case ID	*Disease status*	*Conventional flow*			*NGF*		
		% tPC	*% nPC*	*% aPC*	*% tPC*	*% nPCs*	*% aPC*
1	CR	0.01	0.01	0	0.008	0.007	0.0004
2	CR	0.01	0.01	0	0.02	0.02	0.0002
3	CR	0.007	0.007	0	0.007	0.007	0.0003
4	CR	0.04	0.04	0	0.02	0.02	0.0004
5	sCR	0	0	0	0.003	0.002	0.0006
6	CR	0.1	0.1	0	0.1	0.1	0.002
7	CR	0.2	0.2	0	0.2	0.2	0.02
8	sCR	0	0	0	0.002	0.003	0.0002
9	CR	0.9	0.9	0	0.6	0.6	0.005
10	CR	0.2	0.2	0	0.06	0.06	0.0004
11	VGPR	0	0	0	0.01	0.01	0.0007
12	CR	0	0	0	0.01	0.01	0.006
13	sCR	0.2	0.2	0	0.2	0.2	0.005
14	sCR	0.03	0.03	0	0.03	0.03	0.002
15	CR	0.1	0.1	0	0.07	0.04	0.03
16	VGPR	0	0	0	0.009	0.008	0.001
17	CR	0.1	0	0	0.004	0.003	0.001
18	CR	0.04	0	0	0.01	0.01	0.0002
19	CR	0.09	0.08	0.01	0.05	0.05	0
20	CR	0.1	0.1	0.006	0.1	0.1	0
21	CR	0.4	0.3	0.02	0.3	0.3	0

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^- and MRD^+ cases with decreased levels

BM-associated cell population (phenotype[a])	*Normal BM range (n = 16), hemodiluted BM cutoff*	*% of cases with low levels*		*NGF MRD^- patients showing disease progression TOTAL*	
		MRD^-	*MRD^+*	*Case 1*	*Case 2*
% Mast cells ($CD117^{hi}$, $CD45^+$, $CD81^{dim}$, $CD38^+$, SSC/FSC^{int-hi})	0.006% (0.002–0.03%)	10%	5%	0.0017%	0.0006%
% NRBC (SSC/FSC^{lo}, all markers$^-$)	6.4% (2–11.5%)	3%	1%	8%	5%
% $CD19^-$ nPC ($CD38^{hi}$, $CD138^+$, $CD45^{-/+}$, $CD56^{-/+}$, $CD81^{-/+}$, $CD27^+$, FSC/SSC^{int})	0.05% (0.003–0.2%)	14%	14%	0.0038%	0.0036%
% $CD27^+$ B Cell precursors ($CD19^+$; $CD38^{++}$, $CD45^{dim}$, $CD81^{++}$, $CD27^+$, $CyIg^-$, FSC/SSC^{lo})	0.08% (0.004–0.4%)	12%	11%	0.0001%	0.003%
% $CD27^-$ B Cell precursors ($CD19^+$; $CD38^{++}$, $CD45^{dim}$, $CD81^{++}$, $CD27^-$, $CyIg^-$, FSC/SSC^{lo})	0.4% (0.05–2.2%)	11%	13%	0.001%	0.01
% Mature B cells ($CD19^+$; $CD38^+$, $CD45^{dim/+}$, $CD81^+$, $CD27^{-/+}$, $CyIg^+$, FSC/SSC^{lo})	1.6% (0.6–3.5%)	26%	23%	0.1%	0.1%
% Myeloid precursors ($CD117^+$, $CD45^{dim}$, $CD38^{dim/+}$; $CD81^+$, SSC/FSC^{int})	1.8% (0.2–3.6%)	8%	8%	0.3%	0.3%
% Endothelial and mesenchymal cells ($CD81^{hi}$, $CD45^-$, $SSC/FSC^{int/hi}$)	0.01% (0.0005–0.08%)	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. [a]Markers used in the gating strategy.

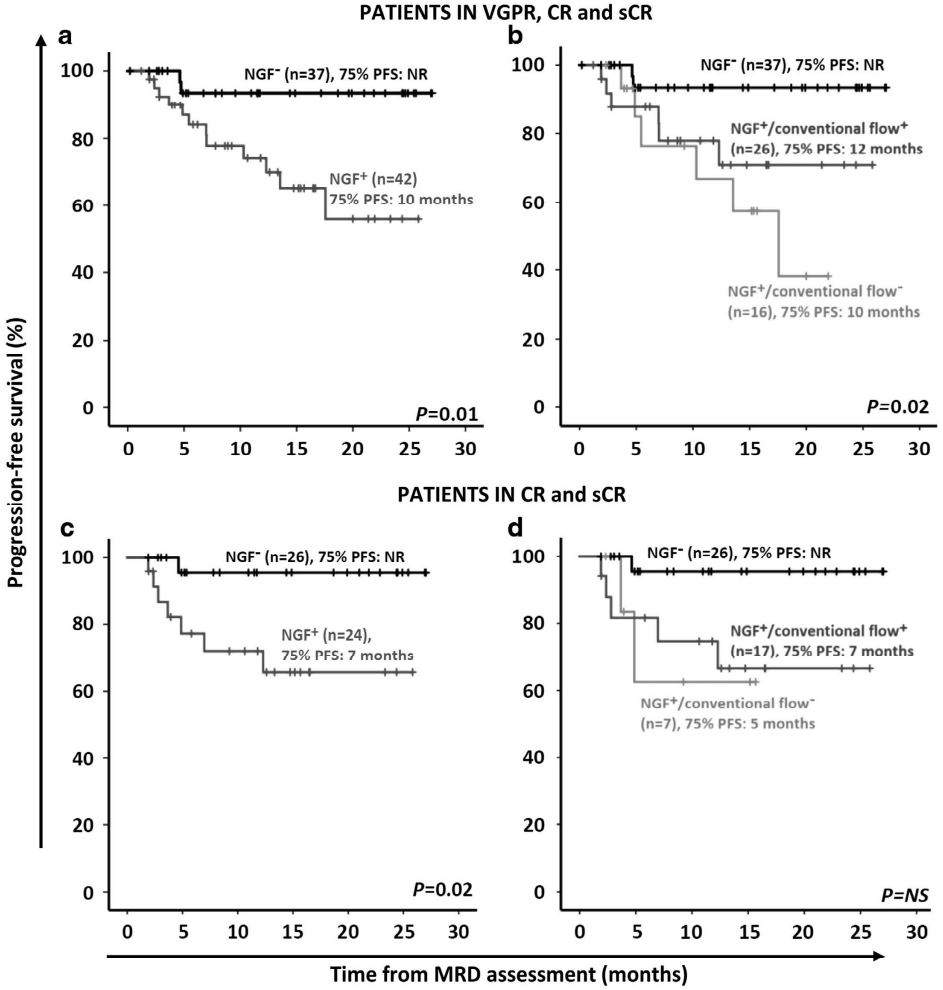


Figure 4. PFS curves of MM patients grouped according to their BM MRD status as assessed by NGF (**a**, **c**) and both NGF and conventional flow-MRD (**b**, **d**). In **a** and **b** the impact of the MRD status is shown for MM patients in VGPR, CR and sCR ($n = 79$), while in **c** and **d**, PFS analyses was restricted to MM patients who were in CR and sCR at the moment of MRD assessment ($n = 50$).

did not differ from those considered to be essential by expert consensus.[40,41,48] However, we proved here that selection of optimal fluorochrome-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.[42,56] To the best of our knowledge, this is the first report in which consistently $> 5 \times 10^6$ cells/patient sample (usually $> 10^7$ cells) were investigated. For the most frequently used stain-lyse-and-then-wash sample preparation flow-MRD procedures, hundred thousand to 1–2 million cells in $\leqslant 300\,\mu l$ BM are analyzed. In contrast, the here described EuroFlow SOP assured acquisition of $> 10^7$ 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 $< 5 \times 10^{-6}$ 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,[16,32,34] and deserves further investigation.

Independently of its potentially greater sensitivity, NGF has other additional advantages over NGS:[37] it is faster (< 4 h), standardized and reproducible, it has a greater applicability ($\geqslant 98\%$), 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.[16] However, EuroFlow-based NGF required fresh material analyzed within 24 h after sampling; this is

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).[16]

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, $CD117^{hi}$ mast cells, $CD45^{-}$ sideward-$scatter^{lo}$ nucleated red blood cells, $CD117^{+}$ myeloid precursors, $CD19^{+}CD38^{hi}CD45^{lo}$ 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 ($CD19^{+}CD38^{hi}CD45^{lo}$ and $CD45^{int}$), immature B-lymphocytes ($CD19^{+}CD45^{hi}CD38^{lo}$), naive B cells ($CD19^{+}CD38^{-}CD27^{-}$), memory B-lymphocytes ($CD19^{+}CD38^{-}CD27^{+}$), and nPC ($CD19^{+}CD56^{-}$, $CD19^{-}CD56^{-}$ and $CD19^{-}CD56^{lo}$) in addition to myeloma PC (Supplementary Figure 3). Of note, decreased numbers of mast cells were found in 17/110 $\geqslant$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.[57] 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

G-EG and RF are employees of Cytognos SL, Salamanca, Spain. JJMvD received research support and traveling 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, MG, RP, M-CdC, JB, J-JL, CA, AB, AG-M, JL, PL, CA-S, JS-M, BD and AO 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 DTS15/00119 from Instituto de Salud Carlos III, Ministerio de Economía y Competitividad, Madrid, Spain.

REFERENCES

1 Ocio 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, Durie 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, Martinez-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* 2008; **26**: 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, Bringhen 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, Bornhauser M. Multiple myeloma. *Lancet* 2015; **385**: 2197–2208.
11 Mateos MV, Ocio EM, Paiva B, Rosinol L, Martinez-Lopez 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, Martinez-Lopez 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, Schaafsma 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-Lopez 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-Lopez J, Vidriales MB, Mateos MV, Montalban 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, Montalbán MA, Puig N, Cordon L, Martinez-Lopez J, Ocio 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 transplantation. *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, Wuilleme S. A single-tube multiparameter seven-colour flow cytometry strategy for the detection of malignant plasma cells in multiple myeloma. *Blood Cancer J* 2013; **3**: e134.

24 Paiva B, Cedena MT, Puig N, Arana P, Vidriales MB, Cordon L *et al*. Minimal residual disease monitoring and immune profiling in multiple myeloma in elderly patients. *Blood* 2016; **127**: 3165–3174.
25 Puig N, Sarasquete ME, Balanzategui A, Martinez J, Paiva B, Garcia H *et al*. Critical evaluation of ASO RQ-PCR for minimal residual disease evaluation in multiple myeloma. A comparative analysis with flow cytometry. *Leukemia* 2014; **28**: 391–397.
26 Sarasquete ME, Garcia-Sanz R, Gonzalez D, Martinez J, Mateo G, Martinez P *et al*. Minimal residual disease monitoring in multiple myeloma: a comparison between allelic-specific oligonucleotide real-time quantitative polymerase chain reaction and flow cytometry. *Haematologica* 2005; **90**: 1365–1372.
27 Ladetto M, Donovan JW, Harig S, Trojan A, Poor C, Schlossnan R *et al*. Real-Time polymerase chain reaction of immunoglobulin rearrangements for quantitative evaluation of minimal residual disease in multiple myeloma. *Biol Blood Marrow Transplant* 2000; **6**: 241–253.
28 Puig N, Sarasquete ME, Alcoceba M, Balanzategui A, Chillon MC, Sebastian E *et al*. Kappa deleting element as an alternative molecular target for minimal residual disease assessment by real-time quantitative PCR in patients with multiple myeloma. *Eur J Haematol* 2012; **89**: 328–335.
29 Martinez-Lopez J, Lahuerta JJ, Pepin F, Gonzalez M, Barrio S, Ayala R *et al*. Prognostic value of deep sequencing method for minimal residual disease detection in multiple myeloma. *Blood* 2014; **123**: 3073–3079.
30 Mailankody S, Korde N, Lesokhin AM, Lendvai N, Hassoun H, Stetler-Stevenson M *et al*. Minimal residual disease in multiple myeloma: bringing the bench to the bedside. *Nat Rev Clin Oncol* 2015; **12**: 286–295.
31 Kumar S, Paiva B, Anderson KC, Durie B, Landgren O, Moreau P *et al*. International Myeloma Working Group consensus criteria for response and minimal residual disease assessment in multiple myeloma. *Lancet Oncol* 2016; **17**: e328–e346.
32 van Dongen JJ, Langerak AW, Bruggemann M, Evans PA, Hummel M, Lavender FL *et al*. Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98-3936. *Leukemia* 2003; **17**: 2257–2317.
33 Ladetto M, Bruggemann M, Monitillo L, Ferrero S, Pepin F, Drandi D *et al*. Next-generation sequencing and real-time quantitative PCR for minimal residual disease detection in B-cell disorders. *Leukemia* 2014; **28**: 1299–1307.
34 Bai Y, Wong KY, Fung TK, Chim CS. High applicability of ASO-RQPCR for detection of minimal residual disease in multiple myeloma by entirely patient-specific primers/probes. *J Hematol Oncol* 2016; **9**: 107.
35 Martinez-Lopez J, Fernandez-Redondo E, Garcia-Sanz R, Montalban MA, Martinez-Sanchez P, Pavia B *et al*. Clinical applicability and prognostic significance of molecular response assessed by fluorescent-PCR of immunoglobulin genes in multiple myeloma. Results from a GEM/PETHEMA study. *Br J Haematol* 2013; **163**: 581–589.
36 Rawstron AC, de Tute RM, Haughton J, Owen RG. Measuring disease levels in myeloma using flow cytometry in combination with other laboratory techniques: lessons from the past 20 years at the leeds haematological malignancy diagnostic service. *Cytometry B Clin Cytom* 2016; **90**: 54–60.
37 van Dongen JJ, van der Velden VH, Bruggemann M, Orfao A. Minimal residual disease diagnostics in acute lymphoblastic leukemia: need for sensitive, fast, and standardized technologies. *Blood* 2015; **125**: 3996–4009.
38 Flanders A, Stetler-Stevenson M, Landgren O. Minimal residual disease testing in multiple myeloma by flow cytometry: major heterogeneity. *Blood* 2013; **122**: 1088–1089.
39 Salem D, Stetler-Stevenson M, Yuan C, Landgren O. Myeloma minimal residual disease testing in the United States: evidence of improved standardization. *Am J Hematol* 2016; **91**: E502–E503.
40 Stetler-Stevenson M, Paiva B, Stoolman L, Lin P, Jorgensen JL, Orfao A *et al*. Consensus guidelines for myeloma minimal residual disease sample staining and data acquisition. *Cytometry B Clin Cytom* 2016; **90**: 26–30.
41 Rawstron AC, Orfao A, Beksac M, Bezdickova L, Brooimans RA, Bumbea H *et al*. Report of the European Myeloma Network on multiparametric flow cytometry in multiple myeloma and related disorders. *Haematologica* 2008; **93**: 431–438.
42 Arroz M, Came N, Lin P, Chen W, Yuan C, Lagoo A *et al*. Consensus guidelines on plasma cell myeloma minimal residual disease analysis and reporting. *Cytometry B Clin Cytom* 2016; **90**: 31–39.
43 Pedreira CE, Costa ES, Barrena S, Lecrevisse Q, Almeida J, van Dongen JJ *et al*. Generation of flow cytometry data files with a potentially infinite number of dimensions. *Cytometry A* 2008; **73**: 834–846.
44 Kalina T, Flores-Montero J, van der Velden VH, Martin-Ayuso M, Bottcher S, Ritgen M *et al*. EuroFlow standardization of flow cytometer instrument settings and immunophenotyping protocols. *Leukemia* 2012; **26**: 1986–2010.
45 van Dongen JJ, Lhermitte L, Bottcher S, Almeida J, van der Velden VH, Flores-Montero J *et al*. EuroFlow antibody panels for standardized n-dimensional flow cytometric immunophenotyping of normal, reactive and malignant leukocytes. *Leukemia* 2012; **26**: 1908–1975.
46 Pedreira CE, Costa ES, Lecrevisse Q, van Dongen JJ, Orfao A, EuroFlow C. Overview of clinical flow cytometry data analysis: recent advances and future challenges. *Trends Biotechnol* 2013; **31**: 415–425.
47 Maecker HT, Frey T, Nomura LE, Trotter J. Selecting fluorochrome conjugates for maximum sensitivity. *Cytometry A* 2004; **62**: 169–173.
48 Flores-Montero J, de Tute R, Paiva B, Perez JJ, Bottcher S, Wind H *et al*. Immunophenotype of normal vs. myeloma plasma cells: Toward antibody panel specifications for MRD detection in multiple myeloma. *Cytometry B Clin Cytom* 2016; **90**: 61–72.
49 Orfao De Matos Correia e Vale A, Sobral da Costa E, Pedreira CE. Method for the multidimensional detection of aberrant phenotypes in neoplastic cells to be used to monitor minimal disease levels using flow cytometry, European patent 1724581B1 (2014).
50 Orfao De Matos A, Pedreira CE, Sobral Da Costa E. Generation of flow cytometry datafiles with a potentially infinite number of dimensions derived from the fusion of a group of separate flow cytometry datafiles and their multidimensional reconstruction with both actually measured and estimated flow cytometry data, European patent EP1770387B1 (2011).
51 Pedreira CE, Costa ES, Almeida J, Fernandez C, Quijano S, Flores J *et al*. A probabilistic approach for the evaluation of minimal residual disease by multiparameter flow cytometry in leukemic B-cell chronic lymphoproliferative disorders. *Cytometry A* 2008; **73A**: 1141–1150.
52 Pojero F, Flores-Montero J, Sanoja L, Perez JJ, Puig N, Paiva B *et al*. Utility of CD54, CD229, and CD319 for the identification of plasma cells in patients with clonal plasma cell diseases. *Cytometry B Clin Cytom* 2016; **90**: 91–100.
53 Paiva B, Puig N, Garcia-Sanz R, San Miguel JF. Is this the time to introduce minimal residual disease in multiple myeloma clinical practice?. *Clin Cancer Res* 2015; **21**: 2001–2008.
54 Rawstron AC, Paiva B, Stetler-Stevenson M. Assessment of minimal residual disease in myeloma and the need for a consensus approach. *Cytometry B Clin Cytom* 2016; **90**: 21–25.
55 Silvennoinen R, Lundan T, Kairisto V, Pelliniemi TT, Putkonen M, Anttila P *et al*. Comparative analysis of minimal residual disease detection by multiparameter flow cytometry and enhanced ASO RQ-PCR in multiple myeloma. *Blood Cancer J* 2014; **4**: e250.
56 Wood BL. Principles of minimal residual disease detection for hematopoietic neoplasms by flow cytometry. *Cytometry B Clin Cytom* 2016; **90**: 47–53.
57 Royston DJ, Gao Q, Nguyen N, Maslak P, Dogan A, Roshal M. Single-Tube 10-fluorochrome analysis for efficient flow cytometric evaluation of minimal residual disease in plasma cell myeloma. *Am J Clin Pathol* 2016; **146**: 41–49.

Supplementary Information accompanies this paper on the Leukemia website (http://www.nature.com/leu)