Report of the European Myeloma Network on multiparametric flow cytometry in multiple myeloma and related disorders

Andy C. Rawstron*,[1] Alberto Orfao*,[2] Meral Beksac,[3] Ludmila Bezdickova,[4] Rik A. Brooimans,[5] Horia Bumbea,[6] Klara Dalva,[3] Gwenny Fuhler,[7] Jan Gratama,[5] Dirk Hose,[8] Lucie Kovarova,[9] Michael Lioznov,[10] Gema Mateo,[2] Ricardo Morilla,[11] Anne K. Mylin,[12] Paola Omedé,[13] Catherine Pellat-Deceunynck,[14] Martin Perez Andres,[2] Maria Petrucci,[15] Marina Ruggeri,[13] Grzegorz Rymkiewicz,[16] Alexander Schmitz,[17] Martin Schreder,[18] Carine Seynaeve,[19] Martin Spacek,[4] Ruth M. de Tute,[1] Els Van Valckenborgh,[19] Nicky Weston-Bell,[20] Roger G. Owen,[1] Jesús F. San Miguel,[2] Pieter Sonneveld,[21] Hans E. Johnsen,[22] on behalf of the European Myeloma Network[23]

[1]The HMDS, Leeds Teaching Hospitals NHS Trust, Leeds, UK; [2]General Cytometry Service, Department of Medicine and Cancer Research Center (IBMCC), University of Salamanca/CSIC and Department of Haematology, University Hospital of Salamanca, Salamanca, Spain; [3]Department of Haematology, Ankara University, Ankara, Turkey; [4]Department of Clinical Haematology, University Hospital Kralovske Vinohrady, Prague and 3rd Faculty of Medicine, Charles University, Prague, Czech Republic; [5]Department of Internal Oncology, Erasmus MC, Daniel den Hoed Cancer Center, Rotterdam, the Netherlands; [6]Emergency University Hospital, Department of Hematology, "Carol Davila" University of Medicine, Bucharest, Romania; [7]Dept. of Cell Biology Section Immunology University Medical Center Groningen, Groningen, The Netherlands; [8]Universitätsklinikum Heidelberg, Medizinische Klinik V, Sektion Multiples Myelom, Multiple Myeloma Research Laboratory, Heidelberg, Germany; [9]Dept. of Hematology, University Hospital, Brno, Czech Republic; [10]University Medical Centre Hamburg - Eppendorf Clinic for Stem Cell Transplantation, Hamburg, Germany; [11]Section of Haemato-Oncology, Institute of Cancer Research, Sutton, Surrey, UK; [12]Department of Haematology, Rigshospitalet University of Copenhagen, Denmark; [13]Flow Cytometry Lab., Divisione di Ematologia dell'Università di Torino, Italy; [14]Inserm U601 and Laboratoire d'Hématologie, CHU de Nantes, France; [15]University Rome La Sapienza, Dept. of Cellular Biotechnology and Hematology, Rome, Italy; [16]Flow Cytometry Lab, Department of Pathology, the Maria Skodowska-Curie Memorial Cancer Center and Institute of Oncology, Warsaw, Poland; [17]Department of Haematology, Aalborg Hospital Science and Innovation Center, Aarhus University Hospital, Aalborg, and Department for Molecular Biology, Aarhus University, Aarhus, Denmark; [18]Wilhelminen-Krebsforschungsinstitut 1st Department of Medicine Wilhelminenspital Wien, Vienna, Austria; [19]Vrije Universiteit Brussel, Brussels, Belgium; [20]University of Southampton, Southampton, UK; [21]Dept of Hematology, Erasmus MC, Rotterdam, the Netherlands; [22]Department of Haematology, the Research Laboratory, Aalborg Hospital Science & Innovation Center, Aarhus University Hospital, Aalborg, Denmark, and [23]The European Myeloma Network (www.myeloma-europe.org) Secretariat, Aalborg Hospital Science and Innovation Center, Aarhus University Hospital, Aalborg, Denmark

Citation: Rawstron AC, Orfao A, Beksac M, Bezdickova L, Brooimans RA, Bumbea H, Dalva K, Fuhler G, Gratama J, Hose D, Kovarova L, Lioznov M, Mateo G, Morilla R, Mylin AK, Omedé P, Pellat-Deceunynck C, Perez Andres M, Petrucci M, Ruggeri M, Rymkiewicz G, Schmitz A, Schreder M, Seynaeve C, Spacek M, de Tute RM, Van Valckenborgh E, Weston-Bell N, Owen RG, San Miguel JF, Sonneveld P, and Johnsen HE, on behalf of the European Myeloma Network. Report of the European Myeloma Network on multiparametric flow cytometry in multiple myeloma and related disorders. Haematologica 2008 Mar; 93(3):431-438. doi: 10.3324/haematol.11080.

Discrepancies in plasma cell enumeration

The major cause of discrepancy occurs because the marrow samples taken for morphological review are often the highest quality "first-pull" marrow aspirate samples while those sent for laboratory investigations are often secondary aspirate samples with a higher degree of peripheral blood contamination. A further cause of the discrepancy is that morphological assessment may overestimate the percentage of plasma cells in two ways. Firstly, morphological enumeration may begin from a field containing a high proportion of plasma cells rather than from a randomly chosen field, artificially increasing the plasma cell counts, particularly in cases with <15% plasma cells. Secondly, morphological enumeration may focus on bone marrow particles which have an increased concentration of plasma cells; although this may be more representative of bone marrow, it leads to an inaccurate comparison of enumeration approaches. In order to address this issue, plasma cell percentages detected by flow cytometry were compared with the percentages detected by morphological assessment of bone marrow films made from first-pull aspirate, bone marrow films from the laboratory EDTA sample and also cytocentrifuge slides made from the same sample that was used for flow cytometry. The results, shown in Figure 1, demonstrated that the major cause of numerical differences relates to sample quality, as there was a significant reduction in the percentage of plasma cells detected by flow cytometry compared to the first-pull aspirate morphology but a similar percentage to that detected on cytocentrifuged samples.

It is also acknowledged that assessment of trephine biopsies is helpful because aspirate samples may be dilute and/or non-representative. However, once a technical artefact has been excluded, there are still occasional discrepancies, further research being required to identify other factors (e.g.: homotypic adhesion, particularly through CD56), contributing to the numerical differences.[21-23]

Consensus existed about the need to continue morphological determination of the plasma cell percentage for the diagnosis of plasma cell disorders to maintain consistency with the current diagnostic criteria. Nevertheless, it was emphasized that recent studies indicate that the overall enumeration of plasma cells

from the whole BM aspirate sample by flow cytometry could be of higher prognostic utility in myeloma patients compared with the morphological plasma cells counts.[20]

Identification of prognostic markers and immunophenotypic screening of cytogenetic abnormalities in myeloma and MGUS

So far the prognostic value of immunophenotyping is not clearly established.[29] Recent data suggest that some antigenic profiles such as those defined by the patterns of expression of CD28 and CD117 may correlate with disease outcome. Nevertheless, these profiles are not independent prognostic factors due to their association with well-established genetic lesions.[33] For many years it has been well established that DNA hyperdiploidy in myeloma is associated with unique immunophenotypic features as reflected by a higher reactivity for CD56 and lower expression of CD10, CD15 and CD20.[27,34] CD20 expression has also been associated with the t(11;14)(q13;q32).[35] More recently, it has been shown that other balanced translocations and cytogenetic abnormalities are also closely related to characteristic immunophenotypic features of myeloma plasma cells.[24] Accordingly, DNA hyperdiploidy has been associated with a higher reactivity for both CD56 and CD117, while non-hyperdiploid cases more frequently express CD28 and CD20. In addition, IgH translocation to 11q13 in myeloma was reported as being more frequently associated with a CD20+, CD56- and CD117- myeloma plasma cell phenotype. In turn, IgH translocations involving chromosome partners other than 11q13 were almost exclusively found among CD20- and CD117- cases. Finally, del(13q) was also reported to be associated with a CD117- plasma cell immunophenotype. However, although a relationship has been found between specific genetic abnormalities and unique plasma cell phenotypes,[24,35] general agreement existed on the fact that such reported associations are not strong enough that they would support recommending the use of specific immunophenotypic profiles in the diagnostic screening of genetic abnormalities in myeloma. In this regard, it was concluded that additional studies are required to further explore the potential existence of stronger phenotypic/genetic associations. In a similar way, the diagnostic and/or prognostic utility of the identification and enumeration of circulating clonal plasma cells in myeloma patients,[36-40] was thought to require further confirmation in large series of patients by different clinical research groups, prior to recommending its routine clinical use.

In 1995, Pellat-Deceunynck *et al.*[21] reported that CD45 was lost by most malignant plasma cells as compared to normal plasma cells. More recently, the same authors showed that a subset of plasma cells, approximately 12% of total plasma cells are CD45+bright and that this subset display unique proliferative and phenotypic features. Accordingly, it was reported that proliferation was almost exclusively restricted to the CD45+bright plasma cells; in contrast to other myeloma plasma cells in the sample, CD45+bright plasma cells were CD11a+ with low reactivity for Bcl-2 (41). Although all participating groups could detect plasma cells with some CD45 expression in a significant proportion of patients, the plasma cells expressing CD45+bright could not reproducibly be detected by all groups. It was proposed that

Table 1. Monitoring of minimal residual disease (MRD) in multiple myeloma (MM): advantages and limitations of electrophoretic, molecular and immunophenotypic approaches.

	Immunofixation IgH-PCR	*RQ-ASO*	*Flow cytometry*
Invasive	No	Yes	Yes
Sampling Error	No	Possible	Possible but there are internal controls for marrow quality
Reproducible limit of detection	Low (~0.1%)	0.001%, although quantitative limit is 0.01%	0.01%
Direct measurement	No – only measure relative changes	Yes	Yes
Key problems	Up to 12 months to achieve negative result after plasma cells depleted due to long paraprotein half-life	Labor-intensive to design and test allele-specific oligonucleotides	Complex analysis procedure requires experienced operator
Acceptable sample age	Several days	Several weeks (possibly older depending on storage)	48 hours
Pre-treatment material required	Yes –to determine paraprotein quantity allele-specific and isotype	Yes –to design not essential primer	Preferable but
Rapid results	Yes	No	Yes
Applicability	90%	~75%	>95%
Cost	Very low	Initially high, follow-up low	Moderate
Optimal setting	Continuous monitoring over the course of the disease	Retrospective analysis in multi-center clinical trials	Clinical trials and single time-point response assessment

further studies, including the exchange of original flow cytometry data files and development of consensus gating strategies, could facilitate a better understanding of the reported findings.

Detection of minimal residual disease by flow cytometry

A complete response in myeloma requires the absolute disappearance of serum/urine monoclonal immunoglobulin detected by immunofixation with fewer than 5% plasma cells in the bone marrow as well as stable bone disease and normal calcium.[44] A critical problem with current approaches to disease monitoring is the long half-life of paraproteins, particularly IgG paraproteins, such that it is often not possible to determine whether the patient has had a complete response until three months to twelve months from the end of treatment .[45] Serum-free light chain has a short half-life and therefore responses are more rapid, but the assay is less sensitive for detection of residual disease.[46] Assessing the efficacy of maintenance strategies is even more problematic using conventional response criteria as patients often start maintenance in a near complete remission state. Without a direct measurement of disease levels, the only outcome measure for maintenance strategy is duration of response which takes many years to determine. Direct assessment of the bone marrow tumour load is far preferable and more sensitive, with several studies demonstrating that MRD detection in the bone marrow provides a better prediction of outcome than conventional response assessment by analysis of serum paraprotein levels.

Using allele-specific oligonucleotide (ASO) PCR approaches, approximately 30–70% of patients undergoing allogeneic transplantation become MRD-negative[47-49] whilst in an autologous setting up to 30% of patients achieve an MRD-negative status.[50-55] Such PCR approaches require the generation and validation of a specific oligonucleotide for each patient, which can be costly and time-consuming. Additionally they are only applicable to 70-80% of patients as it is not possible to generate clon-

Table 2. Identification of the optimal marker combination for gating plasma cells: overall performance of different combinations of plasma cell gating markers evaluated at the EMN workshop held in Leeds in May 2007.

Gating markers	*CD38*	*CD38 and CD45*	*CD38 and CD138*	*CD38, and CD45 CD138*
Proportion of cases with detectable disease	42%	28%	42%	61%
Median percentage of plasma cells in cases with detectable disease	8.1% (1.6-35%)	0.8% (0.2-26%)	7.6% (0.5-39%)	2.7% (0.07-33%)
Precision*	67%	67%	67%	92%

**percentage of cases with concordant results between participants.*

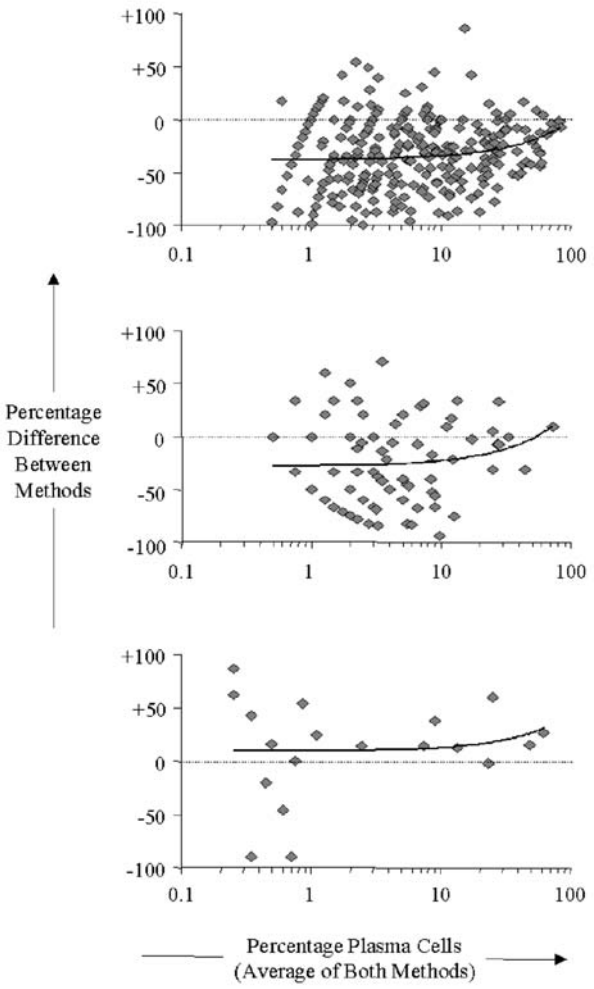

Figure 1. Comparison of flow cytometric detection of plasma cells with morphological analysis of bone marrow aspirate smears and cytocentrifuge slides. Bland-Altman plots are shown, plotting the difference in plasma cell percentage detected by flow cytometry and morphological assessment against the average. The average result for the percentage is plotted on a logarithmic axis to allow a better presentation of the spread of the results; as such, the regression line appears as a curve. Data was generated at the HMDS laboratory in Leeds. (A) Comparison of flow cytometric analysis with morphological assessment of smears made from the initial aspiration in 381 samples from patients at presentation or after treatment. The percentage by flow cytometry is approximately half of the morphologically determined percentage, but the analyses are performed on different samples from the same individual. (B) Comparison of morphological assessment of smears made from the initial aspiration and of smears made on arrival of the sample at the laboratory in 87 of the 381 cases. The percentage of plasma cells is also approximately half that of the result determined from smears derived from the initial aspiration demonstrating that the majority of the discrepancy occurs because of samples sent for laboratory analysis have a higher degree of peripheral blood contamination. (C) Comparison of flow cytometric analysis with morphological assessment of cytocentrifuge preparations using the same separated leucocyte sample for both analyses in 19 of the 381 cases. In this setting both flow cytometry and morphological assessment provide similar results except when there are very low levels of plasma cells (<0.5%) where morphological assessment is not sufficiently sensitive to generate a reliable result. In occasional cases, where the laboratory sample contained representative bone marrow, the percentage of plasma cells detected by flow cytometry was higher than that identified by morphology. The results demonstrate that the primary cause for discrepancies in the plasma cell count derived by flow cytometry compared to morphology is due to differences in sample quality.

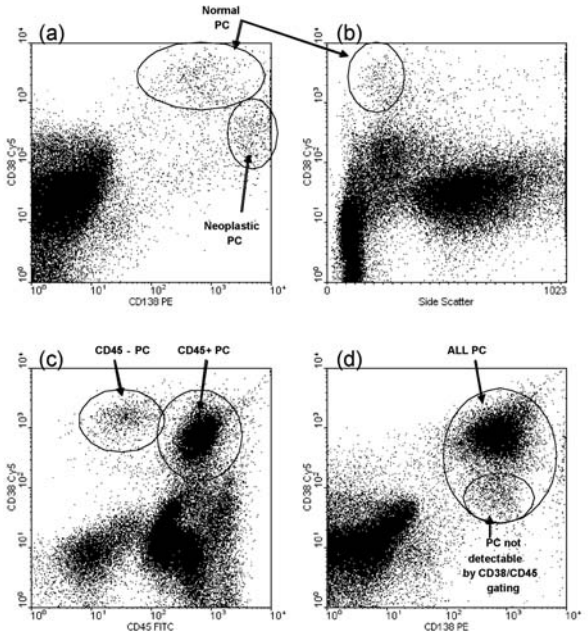

Figure 2. Common problems with immunological gating of plasma cells in bone marrow samples from patients with monoclonal gammopathies. Plasma cells have heterogeneous expression of the common gating markers. There is variation between normal plasma cells and neoplastic plasma cells, with the latter showing stronger expression of CD138, weaker expression of CD38 and CD45. There is also inter-patient variation. CD38 vs. side scatter gating is not suitable for residual disease analysis because neoplastic plasma cells express an average of one log lower levels. This figure shows bone marrow cells from patients with a monoclonal gammopathy prepared by ammonium chloride lysis. Plot (a) shows the CD38 vs. CD138 expression for a patient with approximately equal numbers of normal and residual neoplastic plasma cells. Plot (b) shows the CD38 vs. Side Scatter profile for the same data: the normal plasma cells are apparent as a discrete cluster. The neoplastic plasma cells are not apparent as they form a diffuse cluster overlapping with normal monocytes. Plasma cell gating requires incorporation of at least one other antigen, particularly for residual disease analysis. In the majority of cases, the CD38 and CD45 expression profile is sufficient to discriminate normal and neoplastic plasma cells from the background of normal leucocytes. However, CD45 expression is variable and a significant proportion of patients have a CD45+ fraction of neoplastic cells. Plot (c) shows leucocytes from a patient with several populations of neoplastic plasma cells: the majority of plasma cells express moderate levels of CD45. The CD45+ plasma cells are evident at presentation although a proportion with weak CD38 cannot be identified from CD38/CD45 expression alone and the remainder would be difficult to discriminate in a minimal disease setting when granulocytes, monocytes, B-cells and T-cells with overlapping levels of CD38/CD45 expression are in excess. The CD138 expression is uniformly increased above background for all the plasma cells. Ideally plasma cell gating should utilise CD38, CD45 and CD138 expression. In most cases two of these antigens are sufficient although it is not possible to determine which two without assessing all three in at least one test. When using four colour analyses to determine plasma cell clonality, it is critical to ensure that the gates used are sufficient to identify all plasma cells.

al IgH rearrangements in all patients.[54] ASO-PCR approaches can detect clonal B-lineage cells at the 0.001% level but sensitivity can vary from patient to patient. Additionally two studies have demonstrated that the critical level for predicting early relapse is above the 0.01% level.[50,54] This level is within the quantitative range of disease-specific flow cytometry approaches, and two independent studies have demonstrated the efficacy of flow cytometric disease monitoring for overall response assessment in clinical trials[56,57] while a further study has demonstrated utility in the evaluation of maintenance strategies.[58] The technical issues involved in performing MRD analysis, as well as diagnostic immunophenotyping, are described in the following section.

The key advantage of flow cytometric MRD analysis is that it uses disease-specific rather than patient-specific markers. Therefore the same basic panel of antibodies can be used for each patient, and it is possible to perform an MRD assessment without knowledge of the presentation phenotype. A single time-point analysis by flow cytometry is considerably cheaper than PCR. Furthermore, the results can be provided rapidly, and therefore can potentially be used to guide therapy. A comparison of the advantages and disadvantages of the different approaches for residual disease detection is outlined in Table 1.

Primary gating antibodies

A series of twelve cases from patients with MGUS or myeloma at presentation or after treatment were analysed with a single six-colour assay reported previously[59] incorporating CD38, CD138, CD45 for plasma cell gating as well as CD19 and intracellular kappa and lambda expression. Four data sets were created from the files: the first data set was created by electronically manipulating the files to remove both CD138 and CD45 information, such that participants could only identify plasma cells according to CD38 and light scatter characteristics. The second data set was created by removing CD138 information, such that participants could only identify plasma cells according to CD38, CD45 and light scatter characteristics. The third data set was created by removing CD45 information, such that participants could only identify plasma cells according to CD38, CD138 and light scatter characteristics. The fourth data set contained all the information. Participants were then requested to analyse the data sets in order, and identify the numbers of CD19+κ+λ-, CD19+κ-λ+, CD19-κ+λ- and CD19-κ-λ+ plasma cells in each file. For the purpose of comparison, cases were defined as having abnormal cells present if there were CD19- plasma cells with a

kappa:lambda ratio > 10:1 or <0.1:1 and at least 50 events in the gate. The results are shown in Table 2.

Using CD38 vs. CD45 to define the primary gate resulted in the lowest detection rate and the lowest median percentage of plasma cells in the cases classified as positive. This approach results in false negative results because disease levels are underestimated in patients with CD45+ disease.

Using either CD38 only or CD38 vs. CD138 as the primary gate resulted in a higher detection rate and the highest median percentage of plasma cells in the cases classified as positive. Both approaches resulted in complete concordance in detection of abnormal cells in only 8/12 cases. Using CD38 as the primary gate resulted in missing cases with relatively low levels of neoplastic plasma cells with weak CD38 expression and in some cases inclusion of non-plasma cell populations such as B-progenitors and monocytes. Using CD38 vs. CD138 to create the primary gate resolved some of these problems, allowing identification of small plasma cell populations, but resulted in the inclusion of events with high-levels of non-specific binding which often skewed the cytoplasmic kappa:lambda ratio in cases with a low percentage of abnormal cells.

Using CD38, CD138 and CD45 resulted in detection of an abnormal plasma cell population in the highest proportion of cases, with the highest precision (concordance in 11/12 cases). The median percentage of plasma cells in positive cases was intermediate to the other approaches, reflecting the fact that using all three reagents for gating results in the highest sensitivity, primarily due to the inclusion of all plasma cell events, and highest specificity as contamination by non-plasma cells is minimized. When using bi-variate plots to create gates, it is critical to create the primary gate using CD38 vs CD138 data rather than CD38 vs. CD45 as the latter approach results in the same problems, i.e. exclusion of CD45+ plasma cells, as using CD38 and CD45 alone to identify plasma cells.

The eventual approach will depend critically on the number of fluorescence detectors available for analysis in the flow cytometer. Examples where the combined use of CD38, CD138 and CD45 was useful for the identification of neoplastic plasma cells in BM samples from patients with monoclonal gammopathies are illustrated in Figure 2.