Journal of Immunological Methods 475 (2019) 112294

Contents lists available at ScienceDirect

Journal of Immunological Methods

journal homepage: www.elsevier.com/locate/jim

Research paper

Lot-to-lot stability of antibody reagents for flow cytometry

Sebastian Böttcher [a,b,*], Vincent H.J. van der Velden [c], Neus Villamor [d], Matthias Ritgen [b], Juan Flores-Montero [e], Hugo Murua Escobar [a], Tomas Kalina [f], Monika Brüggemann [b], Georgiana Grigore [g], Marta Martin-Ayuso [g], Quentin Lecrevisse [e], Carlos E. Pedreira [h], Jacques J.M. van Dongen [i], Alberto Orfao [e,**]

[a] Rostock University Medical Center, Division of Internal Medicine, Medical Clinic III, Hematology, Oncology and Palliative Medicine, Special Hematology Laboratory, Rostock, Germany
[b] Second Department of Medicine, University Hospital of Schleswig Holstein (UKSH), Campus Kiel, Kiel, Germany
[c] Department of Immunology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands
[d] Hospital Clínic de Barcelona, Barcelona, Spain
[e] Cancer Research Center (IBMCC-CSIC/USAL-IBSAL), Cytometry Service (NUCLEUS) and Department of Medicine, University of Salamanca, Salamanca, Spain
[f] Department of Pediatric Hematology and Oncology, 2nd Faculty of Medicine, Charles University (DPH/O), Prague, Czech Republic
[g] Cytognos SL, Salamanca, Spain
[h] Systems and Computing Department (PESC), COPPE, Federal University of Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil
[i] Dept. of Immunohematology and Blood Transfusion (IHB), Leiden University Medical Center (LUMC), Leiden, The Netherlands

ARTICLE INFO

Article history:
Received 18 February 2017
Received in revised form 25 March 2017
Accepted 28 March 2017
Available online 30 March 2017

Keywords:
Fluorescence intensity
Fluorochrome-to-antibody ratio
Flow cytometry
Variability

ABSTRACT

The fluorescence detected using fluorochrome-labelled monoclonal antibodies depends not only on the abundance of the target antigen, but amongst many other factors also on the effective fluorochrome-to-antibody ratio. The diagnostic approach of the EuroFlow consortium relies on reproducible fluorescence intensities over time.

A capture bead system for mouse immunoglobulin light chains was utilized to compare the mean fluorescence intensity of 1323 consecutive antibody lots to the currently used lot of the same monoclonal antibody. In total, 157 different monoclonal antibodies were assessed over seven years. Median relative difference between consecutive lots was 3.8% (range: 0.01% to 164.7%, interquartile range: 1.3% to 10.1%). The relative difference exceeded 20% in 8.8% of all comparisons. FITC labelled monoclonal antibodies (median relative difference: 2.1%) showed a significantly smaller variation between lots than antibodies conjugated to PE (3.5%), PECy7 (3.9%), PerCPCy5.5 (5.8%), APC (5.8%), APCH7 (7.4%), and APCC750 (14.5%). Reagents labelled with Pacific Blue (1.4%), Pacific Orange (2.4%), HV450 (0.7%), and HV500 (1.7%) demonstrated more consistent results compared to conjugates of BV421 (4.1%) and BV510 (16.2%). Additionally, significant differences in lot-to-lot fluorescence stability amongst antibodies labelled with the same fluorochrome were observed between manufacturers.

These observations might guide future quality recommendations for the production and application of fluorescence-labelled monoclonal antibodies in multicolor flow cytometry.

1. Introduction

Modern flow cytometers are able to reliably measure fluorescence intensities. Given the wealth of detailed information provided, the EuroFlow consortium opted to use the intensity of fluorescence signals as principal readout of flow cytometry measurements. Fluorescence intensities assessed simultaneously in eight or more channels using panels that comprise up to seven different tubes create multidimensional spaces that allow the diagnosis, classification, and sensitive follow up of hematological malignancies (Flores-Montero et al., 2017; Theunissen et al., 2017; van Dongen et al., 2012) and detailed analysis of the cellular immune system. Unknown cases or cell populations are classified according to the EuroFlow concept by comparison to multidimensional reference data bases that contain immunophenotypic information on hundreds of previously acquired cases each characterized by a pattern of dozens of continuous fluorescence intensity signals (Pedreira et al., 2013).

Multidimensional fluorescence intensity data allows for diagnosis by flow cytometry with an unprecedented depth. Analysis of this data at the same time also requires novel software tools that utilize technologies to reduce the data representation dimensions, such as principal component analysis (Costa et al., 2010) and other multivariate analytical tools (Pedreira et al., Trends Biotech 2013). Once these tools are

* Correspondence to: S. Böttcher, Rostock University Medical Center, Division of Internal Medicine, Medical Clinic III, Hematology, Oncology and Palliative Medicine, Special Hematology Laboratory, Ernst-Heydemann-Strasse 6, 18057 Rostock, Germany.
** Correspondence to: A. Orfao, Centro de Investigacion del Cancer, Avenida Universidad de Coimbra s/n, Campus Miguel de Unamuno, 37007, Salamanca, Spain.
E-mail addresses: Sebastian.Boettcher@med.uni-rostock.de (S. Böttcher), orfao@usal.es (A. Orfao).

https://doi.org/10.1016/j.jim.2017.03.018
0022-1759/

employed, the end result is dependent on a combination of measurements from all channels. Therefore rigorous standardization of the assessment of all individual fluorescence intensities is an absolute necessity. Diagnosing new cases by comparison to previously acquired cases stored in large data bases obviously requires also the stability of fluorescence intensities over time.

The EuroFlow consortium has developed optimized panels for classification and follow-up of hematological malignancies (Flores-Montero et al., 2016; Theunissen et al., 2017; van Dongen et al., 2012). These panels utilize defined monoclonal antibodies (Mabs) characterized by clone, vendor, label, and titer, which therefore will a priori eliminate major sources of variation in flow cytometry. Moreover, in order to further reduce technical variation, detailed standard operating procedures (SOPs) were developed for instrument set-up (PMT voltages, thresholds, compensation) and staining (Kalina et al., 2012). Importantly, an external quality assessment (QA) scheme comprising biannually centralized ring trials was devised to guarantee the quality of the assessments (Kalina et al., 2015). The QA program utilizes peripheral blood samples of three local healthy donors stained with a modification of the Lymphoid Screening Tube and acquired according to the EuroFlow technical SOPs (Kalina et al., 2012). This program for the first time measures the total technical variation introduced by flow cytometry in a multicenter setting: starting from sample preparation to assessment of fluorescence signals from defined lymphocyte subpopulations. An interlaboratory coefficient of variation (CV) <30% was repeatedly observed for median fluorescence intensities of individual lymphocyte populations over a time period of four years (Kalina et al., 2015). This variation sets the benchmark for the technical precision in flow cytometry aiming at fluorescence intensities as read-out.

Despite the above, the EuroFlow QA program only assesses limited numbers of Mabs. Whether or not the lot-to-lot stability of Mabs would impact on the stability of the assessments of the fluorescence intensity was therefore unknown for the majority of antigens used in the EuroFlow panels.

Preliminary data in the literature suggests that lot-to-lot variation might impact on the fluorescence intensities measured, even for reagents specifically aiming at a equimolar fluorochrome-to-antibody-ratios (Wang et al., 2011). However, comprehensive data sets on the lot-to-lot stability of a large number of fluorescence labelled Mabs tested with the same technology are not publicly available. Vendors typically provide such information for in vitro diagnostic (IVD) labelled products only and without a clear description of the methods used to generate the data. In order to get first insights into the lot-to-lot variation of fluorescently labelled Mabs one of the EuroFlow laboratories (University of Schleswig-Holstein, Kiel, Germany) embarked upon a systematic assessment of the variability of all consecutive Mab lots used for routine and research applications.

At present, there is no clear consensus on the technology to be used to compare the lot-to-lot variability of fluorescence-labelled Mabs. For example, the current guidelines on the use of multicolor flow cytometry in the diagnosis of hematological neoplasms by the British Committee for Standards in Haematology recommend assessing lot-to-lot variability but neither elaborate on methods nor on acceptance criteria (Johansson et al., 2014). Therefore, we devised a simple test based on the specific binding of mouse immunoglobulins bearing the kappa light chain to microsphere capture beads and the relative comparison of mean fluorescence intensities of consecutive lots. The test system is inspired by the Simply Cellular system produced by Bangs Laboratories (Bangs, Fishers, IN) intended for the quality control of fluorochrome-labelled Mabs and for quantitative flow cytometry (Zenger et al., 1998), but for cost-effectiveness uses BD Biosciences (BD, Heidelberg, Germany) CompBeads Anti-Mouse Ig. CompBeads Anti-Mouse Ig specifically bind any mouse κ light chain-bearing immunoglobulin. The system compares the effective fluorochrome-to-antibody ratio of consecutive Mab lots to each other.

We herein report our seven years' experience from this pilot trial.

2. Material and methods

2.1. Staining procedure with CompBeads Anti-Mouse Ig κ and flow cytometry data acquisition

BD CompBeads Anti-Mouse Ig, κ (catalogue number 552843) were diluted 1:3 with phosphate buffered saline (PBS, Sigma, Schnelldorf, Germany). The Mab volume routinely used to stain 500,000 leukocytes was incubated with 25 μl of the prediluted CompBead mixture for 15 min at room temperature. Mabs included into testing were purchased from the following companies: Beckman-Coulter (Krefeld, Germany), BD, Cytognos (Salamanca, Spain), BioLegend (San Diego, CA), Dako (Glostrup, Denmark), eBioscience (San Diego, CA), Invitrogen (Carlsbad, CA), Miltenyi (Bergisch Gladbach, Germany), R&D (Minneapolis, MN), and Southern Biotech (Birmingham, AL). Companies were randomly assigned letters A through J for the purpose of this manuscript. For each comparison a separate tube was prepared for the old and new Mab lots, respectively. The bead suspension was washed once with 3 ml PBS and resuspended with 400 μl PBS. 10,000 beads per lot were acquired using a BD FACSCanto II flow cytometer equipped with violet, blue and red lasers (excitation wavelengths 405 nm, 488 nm, and 633 nm respectively) after standard EuroFlow instrument set-up (Kalina et al., 2012). Beads stained with consecutive lots of the same Mab were acquired directly one after another at the same instrument. Fluorescence (Table 1) was measured with compensation disabled using photomultiplier tube (PMT) 1 (emission wavelength filter 450/50 nm) for Pacific Blue (PacB), Brilliant Violet 421 (BV421), as well as Horizon V450 (HV450) labelled Mabs and PMT2 (510/50 nm) for Pacific Orange (PacO), BV510 as well as HV500 labelled Mabs. Fluorescence from fluorescein isothiocyanate (FITC)-, phycoerythrin (PE)-, peridinin chlorophyll protein cyanine 5.5 (PerCPCy5.5)- and phycoerythrin cyanine 7 (PECy7)- labelled Mabs were acquired by PMTs 3 through 6 (emission filters: 530/30 nm, 585/42 nm, 670 nm long pass (LP), and 780/60 nm, respectively). The allophycocyanine (APC) signal was detected by PMT 7 (660/20 nm), whereas APC Hilite7 (APCH7) and APCC750 signals were acquired at PMT 8 (780/60 nm).

2.2. Gating strategy

Within a forward scatter/side scatter (FSC/SSC) density plot the major bead population was marked by a polygonal gate (Fig. 1). Events meeting that first gate were depicted in a histogram. The main fluorescence distribution was marked by an interval gate. Events included into both gates were used for assessing the mean fluorescence intensity (MFI) of the stained bead population.

2.3. Statistical methods

For each comparison of consecutive lots of the same Mab the relative difference in MFI was calculated as follows: absolute value $((MFI_{old} - MFI_{new}) / (MFI_{old} + MFI_{new}) * 2)$ and reported as percentage. Mann-Whitney and Kruskal-Wallis tests were used for comparisons of relative differences in MFI between two and more groups, respectively. The nonparametric correlation coefficient r was calculated according to Spearman. The interquartile range (IQR) was calculated as 25th to 75th percentiles. Significance was set at 0.05. Statistical analyses were performed using GraphPad Prism v4 (San Diego, CA).

3. Results

3.1. Overall lot-to-lot variation

From October 21, 2009 until October 31, 2016 the MFI of a total of 1323 lots of directly conjugated Mabs were compared to their directly following lots. The testing comprised all Mabs used by Second

Table 1
Relative difference in MFI between consecutive Mab lots by fluorochrome and optical setup of the flow cytometer. Emission filter characteristics per photomultiplier tube (PMT) are provided (LP: long pass). Significance levels: *p < 0.05, **p < 0.001, ***p < 0.0001 compared to HV450. Comparisons between fluorochromes measured by the same PMT: #p < 0.01. The percentage of lot-to-lot comparisons failing the acceptance criteria (<40% for APCH7 and APCC750, <30% for APC and PerCPCy5.5, <20% for all other conjugates) are given as failure rate.

PMT	Emission filter [nm]	Fluorochrome	n	Minimum	25th Percentile	Median	75th percentile	Maximum	Failure rate [%]
1	450/50	PacB#	8	0.0	1.0	*1.4*	3.9	15.0	0.0
		BV421** #	6	0.9	2.1	*4.1*	24.9	30.2	16.7
		HV450#	54	0.1	0.3	*0.7*	2.1	16.6	0.0
2	510/50	PacO* #	14	0.3	0.5	*2.4*	12.8	20.5	7.1
		BV510*** #	16	1.9	8.0	*16.2*	23.7	31.1	37.5
		HV500 #	4	0.6	1.0	*1.7*	3.4	4.9	0.0
3	530/30	FITC***	325	0.0	0.7	*2.1*	6.1	46.2	4.0
4	585/42	PE***	311	0.0	1.3	*3.5*	8.9	164.7	4.2
5	670 LP	PerCPCy5.5***	214	0.0	2.1	*5.8*	15.0	97.0	7.9
6	780/60	PECy7***	50	0.1	1.2	*3.9*	12.3	32.3	4.0
7	660/20	APC***	247	0.0	2.5	*5.8*	13.7	83.8	6.9
8	780/60	APCH7***	52	0.8	2.8	*7.4*	14.2	59.4	5.8
		APCC750***	22	0.1	1.8	*14.5*	23.2	59.6	9.1

Department of Medicine, University of Schleswig-Holstein, Kiel, Germany during that time period for routine and research applications.

We report on the comparative MFI analyses of consecutive lots of 157 different Mabs. At median 6 pairs of lots per Mab were compared (range: 1 to 45, IQR: 2 to 10). Lots were tested according to the order they were delivered by the manufacturer. Median relative difference between consecutive lots was 3.8% (range: 0.01% to 164.7%, IQR: 1.3% to 10.1%). Old and new lots differed by a maximum of 20% in 91.2% of all comparisons and by a maximum of 30% in 95.9% of all cases (Fig. 2).

3.2. Lot-to-lot variation by fluorochrome

We observed a marked association between the fluorochrome and the obtained lot-to-lot stability of the directly labelled Mabs (Table 1 and Fig. 3). The lowest median variation was recorded for HV450 (0.7%), whereas that value reached a median of 16.2% for BV510. In comparison to the most stable fluorochrome label, lot-to-lot variation was significantly higher for BV421 (median 4.1%) as well as for PacO (2.4%) and, particularly, BV510 (16.2%). Moreover, the variability between lots also significantly exceeded the variability of the most stable label HV450 for Mabs labelled to FITC (median 2.1%), PE (3.5%), PerCPCy5.5 (5.8%), PECy7 (3.9%), APC (5.8%) and both fluorochrome conjugates assessed using PMT 8 (medians 7.4% and 14.5% for APC-H7 and APCC750, respectively).

Differences in lot-to-lot stability likely represent features of the fluorochrome conjugation process and its quality control, not the detector. This is particularly evident from the observation that alternative fluorochrome labels detectable by the same PMT varied significantly amongst each other for PMT 1 (PacB, BV421, and HV450; $p < 0.01$) and PMT2 (PacO, BV510 and HV500; $p < 0.01$).

The performance characteristics of the test described herein were initially evaluated until April 12, 2010. In order to reflect the relationship between lot-to-lot stability and fluorochrome, we defined the minimum acceptance criteria per PMT based on that initial set of data (data not shown). We were thus balancing the need of stable MFI assessments with the performance of commercially available Mabs per detector channel. Mabs labelled to conjugates detectable by PMT 8 (APCH7, APCC750) had to demonstrate a lot-to-lot difference of <40%, APC and PerCPCy5.5 labelled Mabs were required to differ <30% between lots and stability was considered acceptable for lot-to-lot differences below 20% for all other Mabs. Depending on the fluorochrome, these acceptance criteria were met in over 90% of Mabs for 11 of the 13 fluorochromes evaluated. Only for BV421 and BV510 the acceptance criteria were met in a lower percentage: 83.3% and 62.5% of pair-wise comparisons, respectively (Table 1).

3.3. Lot-to-lot variation by manufacturer

We next investigated whether the labelling process utilized by different manufacturers impacted on the lot-to-lot stability of the Mab conjugates (Table 2). To safeguard against random results, manufacturers for which fewer than 10 comparisons were performed were not included into this analysis. Test results from different manufacturers were available for Mabs labelled to FITC, PE, and APC, respectively. Manufacturer B produced the least variable FITC conjugates. In comparison to FITC labelled Mabs from this company, Mabs purchased from vendors D and E were significantly less stable. According to our test system manufacturer A's PE labelled Mabs demonstrated the highest stability over time, whereas PE Mabs from manufacturers B and E showed significantly greater variation. We did not detect significant differences in variation of fluorescence intensities between APC labelled Mabs obtained from the two companies (A and B) evaluated.

3.4. Reproducibility of the test

Whenever new lots did not meet the acceptance criteria as outlined in Section 3.2, we performed the same lot-to-lot comparison again. The repeat comparison (58 repeated experiments, Fig. 4) in general yielded similar results as the first assessment. We computed a statistically significant correlation of first and second assessments of the MFI of consecutive Mab lots ($r = 0.71$, $p < 0.0001$). Seven out of 58 repeat assessments (12.1%) yielded a relative difference in MFI between lots that met the acceptance criteria for the second assessment. The repeat comparison in those cases generally resulted in a variation slightly superior than the variation measured in the first assessment.

3.5. Availability and performance of alternative lots

After the establishment of acceptance criteria per fluorochrome label (see Section 3.2) manufacturers were systematically requested to provide alternative lots for Mabs whenever a relative difference exceeding those criteria was observed. On 12 occasions alternative lots of the same Mabs were available. These alternative reagent lots showed more similar MFIs compared to the previous lot in five cases (41.7%), now complying with the requested performance criteria. For the remaining seven alternative reagent lots, the new lot again exceeded the predefined acceptance criteria.

4. Discussion

The current paper reports on the lot-to-lot variability observed between consecutive lots of fluorochrome-labelled Mabs in a

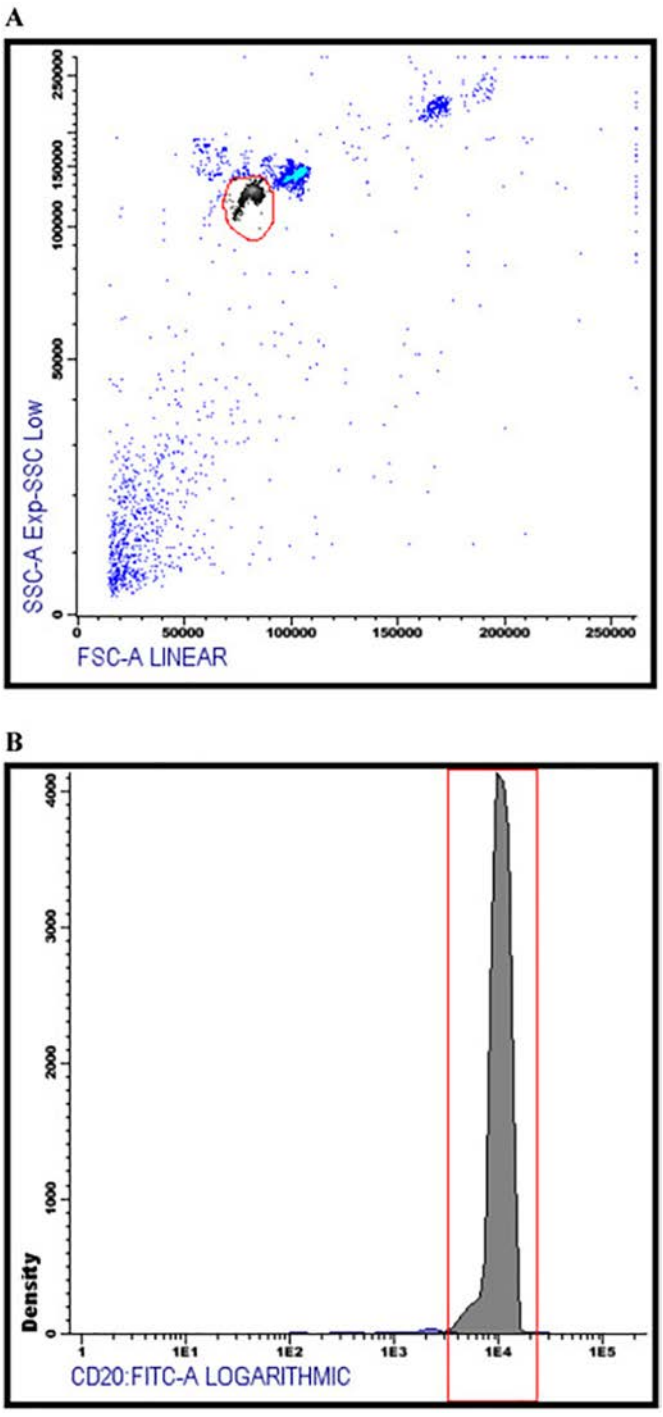

Fig. 1. Gating strategy used for the assessment of MFI of consecutive Mab lots. (A) The major bead population is marked by the polygon gate. (B) The fluorescence of this major bead population is shown as a single parameter histogram. An interval gate is used to exclude outlier events. MFI of the events within the interval gate is used to compare a given lot to its direct predecessor.

comprehensive survey over seven years in a single EuroFlow laboratory. The data presented herein might guide the consortium and other researchers in the field on feasible quality criteria for fluorochrome labelled Mabs. At the same time it provides a possible tool to evaluate comparability and identify potential differences between consecutive lots of Mab reagents.

The lot-to-lot difference in fluorescence intensities between consecutive lots of the same Mabs did not exceed 20% in >90% of all

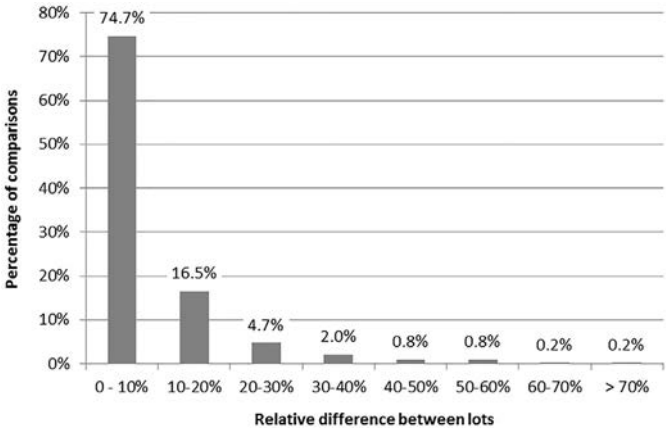

Fig. 2. Frequency distribution of relative differences between lots (n = 1323).

comparisons. Given the observed overall technical variability measured in the EuroFlow QA program (Kalina et al., 2015) (CV < 30% for stable antigens), this difference can be considered optimal. Greater differences of up to 30% might be considered acceptable, but they clearly have the potential to negatively impact on the stability of the classification systems as developed by EuroFlow.

In our system the lot-to-lot variation was dependent on the fluorochrome label. Concerning the canonical fluorochromes the smallest variation was observed for FITC, while PE-labelled Mabs were less stable between lots. The higher variation for PE compared to FITC is somewhat surprising as PE can be conjugated at specific ratios to Mabs and is therefore considered particularly well suited for quantitative flow cytometry (Davis et al., 1998). Even more variation was seen for Mabs labelled to PerCPCy5.5 and APC. Due to the more complicated labelling process (Roederer et al., 1996) the greater variation detected for the tandem dyes PECy7, APCH7, and APCC750 was on the other hand expected. The observed grading in lot-to-lot stability between fluorochromes might guide future investigators in choosing labels for antigens for which stable quantification over extended periods of time is particularly important.

Mabs purchased from different vendors showed significant differences in lot-to-lot stability, likely reflecting differences in the production and/or QA criteria employed by different manufacturers. This observation should caution against generalization of recent

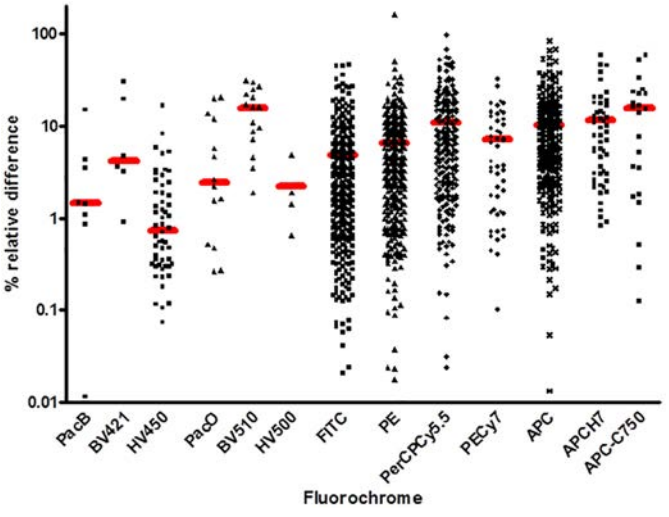

Fig. 3. Relative difference in MFI of consecutive Mab lots grouped by fluorochrome. Each symbol represents a comparative assessment, red bars symbolize medians per fluorochrome.

Table 2
Relative difference in MFI between consecutive FITC, PE, and APC Mab lots by manufacturer. For FITC compared to manufacturer B: $^{**}p < 0.01$, for PE compared to manufacturer A $^{###}p < 0.001$, all other comparisons within the same fluorochrome by manufacturer not significant. Failure rates designate lot-to-lot differences exceeding 20% (FITC, PE) and 30% (APC), respectively.

Fluorochrome	Manufacturer	N	Minimum	25th Percentile	Median	75th percentile	Maximum	Failure rate [%]
FITC	A	38	0.1	1.2	*2.6*	6.1	16.5	0.0
	B	229	0.0	0.6	*1.6*	5.3	27.1	3.1
	D**	17	0.2	2.5	*4.1*	10.5	28.5	5.9
	E**	37	0.1	1.2	*4.4*	8.1	46.2	13.5
PE	A	61	0.0	0.8	*2.2*	3.7	13.9	0.0
	B###	216	0.0	1.5	*4.0*	9.7	50.9	3.7
	E###	11	1.5	5.6	*10.8*	29.1	34.4	27.3
	G	12	0.0	0.8	*4.7*	8.8	28.9	8.3
APC	A	33	0.0	3.0	*5.0*	13.1	35.9	3.0
	B	183	0.1	2.4	*5.8*	12.9	66.6	4.4

investigations suggesting that CD20 quantification might be possible by quantitative flow cytometry using labels such as PerCPCy5.5 or APC in combination with CD4 labelled to the same fluorochrome as reference standard (Degheidy et al., 2016). This approach appears restricted to well-characterized Mabs obtained from one single manufacturer. Furthermore, given the increasing importance of reliable fluorescence quantification manufacturers are urged to revise their internal quality control processes in a way that results in comparable Mab-conjugates over time. We have observed that certain manufacturers would consider all novel Mab lots acceptable if they at least achieve the same MFI values as the previous one, even if the new lot would yield significantly higher fluorescence intensities. Mabs are suitable for the EuroFlow approach only if they show stable effective fluorescence emission per single Mab over many years, whereas both brighter and dimmer lots compromise the diagnostic precision.

As this is a first pilot study on the lot-to-lot stability of Mabs, it is hampered by obvious technical limitations. We assessed differences in fluorochrome-to-antibody ratios but are unable to detect differences in affinity to the target epitope that might also contribute to different fluorescence intensities emitted by different lots of Mabs. Our bead-based test system has advantages in practicability and results in narrow fluorescence distributions (CV typically 15%, Fig. 1B) but cannot detect the target epitope binding on cells. Other known factors with a possible impact on fluorescence intensities (e.g. pH, washing procedures, clones used) are unlikely to affect the performance of Mabs when used according to EuroFlow SOPs, as these factors are kept stable. By design the system used here is unable to compare lots of Mabs bearing immunoglobuline lambda light chains, polyclonal reagents, and Mabs generated in species other than mice. However, such reagents are used only to a very limited extent in EuroFlow panels and generally in clinical laboratory diagnostics. The test system described herein is unsuitable for off-the-shelf Mab mixtures, as the beads cannot discriminate between different Mabs present in such mixtures. Our test approach works relative to the lot of a particular Mab currently in use, so that shifts in Mab quality that develop gradually over several lots cannot be detected. This is due to the fact that we use different lots of capture beads that might bind different amounts of antibodies and hence yield different fluorescence intensities. The measurements described herein are not affected by variations in capture bead lots, as we normalize our results relative to results obtained with the preceding Mab lot using exactly the same lot of capture beads. It might be possible to utilize beads with specified antibody binding capacity in future investigations which would in turn yield absolute MFI results and hence could even detect gradual shifts of fluorescence signals over time. Our current system works fairly reliable with < 10% clearly different results on repeat testing. Nevertheless, due to the possibility of pipetting errors a repeat assessment is always indicated when expected target performance values are not met.

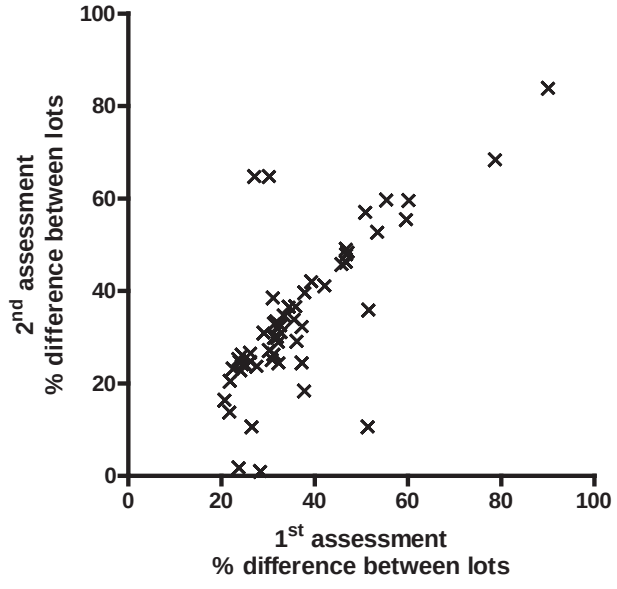

Fig. 4. Retest reliability of relative differences in MFI. Each cross symbolizes a repeat assessment of the same pair of consecutive Mab lots (n = 58 pairs of lot-to-lot comparisons; r = 0.71, p < 0.0001).

The complexity of pipetting eight or more Mabs to a single sample aliquot can be significantly reduced by preparing in-house Mab premixes. In keeping e.g. with recommendations from the British Committee for Standards in Haematology (Johansson et al., 2014) we recommend preparing such premixes. We detected no significant impact of premixing on the performance of the Mabs when compared to single Mab pipetting (data not shown). Any testing of lot-to-lot stability as described herein should be performed prior to preparing Mab premixes. Moreover, we recommend assessing the first sample stained with a new premix particularly careful (e.g. independently by two technicians) for the expected reactivity on normal leukocytes in order to minimize the possibility of missing or wrong Mabs in the premix. In our experience premixes are stable for 28 days as long as the expiry dates of the individual Mabs are taken into account. Premixes should not include Mabs labelled with the tandem dyes PECy7, APCC750, or APCH7, as these reagents according to the manufacturers require special stabilization solutions in defined concentrations in order to maintain stability. Moreover, the EuroFlow consortium recommends staining back-bone markers of multi-tube panels prior to the additional markers from the panel, so that these back-bone markers are not included into premixes either. However, in an attempt to increase the applicability of premixes, the consortium is currently evaluating the feasibility of the inclusion of tandem dyes into premixes (cf. van der Velden et al., this issue of the Journal of Immunological Methods). Moreover, efforts to extent the stability period of premixes beyond the currently accepted 28 days limit are also underway.

In summary, information presented herein proves the feasibility of Mab lot-to-lot testing over extensive periods of time. Our preliminary

data suggest that for the vast majority of Mabs a lot-to-lot variation of the effective fluorochrome-to-antibody ratio of <20% is achievable, in line with the technical variation from other causes observed by the consortium. A relative difference in fluorescence intensity of <20% between lots might therefore be considered a desirable target. We observed an association of the lot-to-lot variability to fluorescent dye and manufacturer.

Acknowledgements

We are grateful to Elke Harbst, Jamileh Hanani, Linda Falck, Steffen Bachmann, Lada Henseleit, Julia Böge, Claudia Stepien, Kristin Schröder, and Maike Ipsen for excellent technical support.

References

Costa, E.S., Pedreira, C.E., Barrena, S., Lecrevisse, Q., Flores, J., Quijano, S., Almeida, J., del Carmen Garcia-Macias, M., Bottcher, S., Van Dongen, J.J., et al., 2010. Automated pattern-guided principal component analysis vs expert-based immunophenotypic classification of B-cell chronic lymphoproliferative disorders: a step forward in the standardization of clinical immunophenotyping. Leukemia 24, 1927–1933.

Davis, K.A., Abrams, B., Iyer, S.B., Hoffman, R.A., Bishop, J.E., 1998. Determination of CD4 antigen density on cells: role of antibody valency, avidity, clones, and conjugation. Cytometry 33, 197–205.

Degheidy, H., Abbasi, F., Mostowski, H., Gaigalas, A.K., Marti, G., Bauer, S., Wang, L., 2016. Consistent, multi-instrument single tube quantification of CD20 in antibody bound per cell based on CD4 reference. Cytometry B Clin. Cytom. 90, 159–167.

Flores-Montero, J., de Tute, R., Paiva, B., Perez, J.J., Bottcher, S., Wind, H., Sanoja, L., Puig, N., Lecrevisse, Q., Vidriales, M.B., et al., 2016. Immunophenotype of normal vs. myeloma plasma cells: toward antibody panel specifications for MRD detection in multiple myeloma. Cytometry B Clin. Cytom. 90, 61–72.

Flores-Montero, J., Flores, L.S., Paiva, B., Puig, N., García-Sánchez, O., Böttcher, S., van der Velden, V.H.J., Pérez-Morán, J., Vidriales, M.B., García-Sanz, R., et al., 2017. Next generation flow (NGF) for highly sensitive and standardized detection of minimal residual disease in multiple myeloma. Leukemia (epub ahead of print).

Johansson, U., Bloxham, D., Couzens, S., Jesson, J., Morilla, R., Erber, W., Macey, M., 2014. Guidelines on the use of multicolour flow cytometry in the diagnosis of haematological neoplasms. British Committee for Standards in Haematology. Br. J. Haematol. 165, 455–488.

Kalina, T., Flores-Montero, J., Lecrevisse, Q., Pedreira, C.E., van der Velden, V.H., Novakova, M., Mejstrikova, E., Hrusak, O., Bottcher, S., Karsch, D., et al., 2015. Quality assessment program for EuroFlow protocols: summary results of four-year (2010−2013) quality assurance rounds. Cytometry A 87, 145–156.

Kalina, T., Flores-Montero, J., van der Velden, V.H., Martin-Ayuso, M., Bottcher, S., Ritgen, M., Almeida, J., Lhermitte, L., Asnafi, V., Mendonca, A., et al., 2012. EuroFlow standardization of flow cytometer instrument settings and immunophenotyping protocols. Leukemia 26, 1986–2010.

Pedreira, C.E., Costa, E.S., Lecrevisse, Q., van Dongen, J.J., Orfao, A., 2013. Overview of clinical flow cytometry data analysis: recent advances and future challenges. Trends Biotechnol. 31, 415–425.

Roederer, M., Kantor, A.B., Parks, D.R., Herzenberg, L.A., 1996. Cy7PE and Cy7APC: bright new probes for immunofluorescence. Cytometry 24, 191–197.

Theunissen, P., Mejstrikova, E., Sedek, L., van der Sluijs-Gelling, A.J., Gaipa, G., Bartels, M., Sobral da Costa, E., Kotrova, M., Novakova, M., Sonneveld, E., et al., 2017. Standardized flow cytometry for highly sensitive MRD measurements in B-cell acute lymphoblastic leukemia. Blood 129 (3), 347–357.

van Dongen, J.J., Lhermitte, L., Bottcher, S., Almeida, J., van der Velden, V.H., Flores-Montero, J., Rawstron, A., Asnafi, V., Lecrevisse, Q., Lucio, P., et al., 2012. EuroFlow antibody panels for standardized n-dimensional flow cytometric immunophenotyping of normal, reactive and malignant leukocytes. Leukemia 26, 1908–1975.

Wang, L., Abbasi, F., Jasper, G.A., Kreitman, R.J., Liewehr, D.J., Marti, G.E., Stetler-Stevenson, M., 2011. Variables in the quantification of CD4 in normals and hairy cell leukemia patients. Cytometry B Clin. Cytom. 80, 51–56.

Zenger, V.E., Vogt, R., Mandy, F., Schwartz, A., Marti, G.E., 1998. Quantitative flow cytometry: inter-laboratory variation. Cytometry 33, 138–145.