Journal of Immunological Methods 475 (2019) 112348

Contents lists available at ScienceDirect

Journal of Immunological Methods

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

Research paper

Standardization of 8-color flow cytometry across different flow cytometer instruments: A feasibility study in clinical laboratories in Switzerland

Hana Glier[a,*], Ingmar Heijnen[b], Mathieu Hauwel[c], Jan Dirks[d], Stéphane Quarroz[e], Thomas Lehmann[f], Alicia Rovo[d], Kornelius Arn[g], Thomas Matthes[c], Cassandra Hogan[c], Peter Keller[h], Ewa Dudkiewicz[i], Georg Stüssi[j], Paula Fernandez[a], , on behalf of Swiss Cytometry Society

[a] Institute of Laboratory Medicine, Cantonal Hospital Aarau, Aarau, Switzerland
[b] Medical Immunology, University Hospital Basel, Basel, Switzerland
[c] Hematology Laboratory, Geneva University Hospital, Geneva, Switzerland
[d] Diagnostic Hematology, University Hospital Basel, Basel, Switzerland
[e] Hematology Laboratory, Vaudoise University Hospital Center, Lausanne, Switzerland
[f] Clinical Chemistry, Hematology and Immunology Laboratory, Center for Laboratory Medicine, St. Gallen, Switzerland
[g] Center for Laboratory Medicine, Cantonal Hospital, Lucerne, Switzerland
[h] University Clinic of Hematology, Inselspital/University Hospital Bern, University of Bern, Switzerland
[i] Hematology Clinic, University Hospital Zurich, Zurich, Switzerland
[j] Hematology Laboratory, Oncology Institute of Southern Switzerland and EOLAB, Regional Hospital of Bellinzona and Valli, Bellinzona, Switzerland

ARTICLE INFO

Keywords:
Flow cytometry
Cross-platform standardization
Canto
Navios
EuroFlow

ABSTRACT

The EuroFlow Consortium developed a fully standardized flow cytometric approach from instrument settings, through antibody panel, reagents and sample preparation protocols, to data acquisition and analysis. The Swiss Cytometry Society (SCS) promoted a study to evaluate the feasibility of using such standardized measurements of 8-color data across two different flow cytometry platforms – Becton Dickinson (BD) FACSCanto II and Beckman Coulter (BC) Navios, aiming at increasing reproducibility and inter-laboratory comparability of immunophenotypic data in clinical laboratories in Switzerland.

The study was performed in two phases, i.e. a learning phase (round 1) and an analytical phase (rounds 2 and 3) consisting of a total of three rounds. Overall, 10 laboratories using BD FACSCanto II ($n = 6$) or BC Navios ($n = 4$) flow cytometers participated. Each laboratory measured peripheral blood samples from healthy donors stained with a uniform antibody panel of reagents - EuroFlow Lymphoid Screening Tube (LST) – applying the EuroFlow standardized protocols for instrument setup and sample preparation (www.EuroFlow.org). All data files were analyzed centrally and median fluorescence intensity (MedFI) values for individual markers on defined lymphocyte subsets were recorded; variability from reference MedFI values was assessed using performance scores. Data troubleshooting and discussion of the results with the participants followed after each round at SCS meetings.

The results of the learning phase demonstrated that standardized instrument setup and data acquisition are feasible in routine clinical laboratories without previous experience with EuroFlow. During the analytical phase, highly comparable data were obtained at the different laboratories using either BD FACSCanto II or BC Navios. The coefficient of variation of MedFI for 7 of 11 markers performed repeatedly below 30%. In the last study round, 89% of participants scored over 90% MedFI values within the acceptance criteria (P-score), in line with the results of the EuroFlow quality assessment rounds performed by the EuroFlow expert laboratories(Kalina et al., 2015). Central analysis of data allowed identification of deviations from the standardized procedures and technical issues (e.g. failure to perform correct instrument setup and improper compensation).

In summary, here we show that inter-laboratory cross-platform standardization of 8-color flow cytometric measurements in clinical laboratories is feasible and allows for fully comparable MedFI results across BD FACSCanto II and BC Navios instruments. However, adherence to standardized protocols is crucial. Thus,

* Corresponding author at: FACS/Stem cell Laboratory, Institute of Laboratory Medicine, Kantonsspital Aarau, Tellstrasse 25, CH-5001 Aarau, Switzerland.
E-mail address: hana.glier@ksa.ch (H. Glier).

http://dx.doi.org/10.1016/j.jim.2017.07.013
Received 30 April 2017; Received in revised form 26 July 2017; Accepted 26 July 2017
Available online 29 July 2017

training of the laboratory personnel in the EuroFlow standardized procedures is highly recommended to prevent errors in instrument setup and sample preparation.

1. Introduction

Flow cytometry immunophenotyping has become a key tool for diagnosis, classification, staging, and treatment response monitoring of hematological and immunological disorders (Heel et al., 2017; Dongen and Orfao, 2012).

As multicolor antibody panels used in clinical diagnostic laboratories provide increasing amounts of information about single cells resulting in highly complex data sets, extensive expertise for correct interpretation of flow cytometric data is required (Pedreira et al., 2013). However, many clinical laboratories still continue to rely on their own in-house antibody panels and instrument settings which introduces considerable levels of subjectivity in flow cytometric analyses whose results ultimately depend on the experience and knowledge of local experts and the specific (highly variable) reagent panels applied in different laboratories (Dongen and Orfao, 2012). In order to decrease such subjectivity and variability in flow cytometric immunophenotyping measurements, consensus recommendations and guidelines have been recently generated by several expert groups (Johansson et al., 2014; Westers et al., 2012; Feller et al., 2013). Despite this, none of these groups attempted to standardize the whole flow cytometric analysis process which is critical for obtaining reproducible high quality data that are comparable at both the intra- and inter-laboratory levels over time. Such standardized measurements are indispensable e.g. for high-sensitive minimal residual disease monitoring in leukemia and lymphoma/myeloma (Flores-Montero et al., 2017; Theunissen et al., 2016). Moreover, inter-laboratory comparable data is essential for multicentric studies. Thereby standardization of flow cytometric measurements is urgently needed. Since immunophenotyping relies on the assessment of the level of expression of various markers on one or more cell populations as assessed in a (semi)-quantitative manner, achieving comparable median fluorescence intensity (MedFI) measurements might be considered as an endpoint of standardization across different laboratories.

In 2012, the EuroFlow Consortium developed a fully standardized approach for immunophenotyping in hemato-oncology which covers the whole flow cytometric analysis process from instrument setup through antibody panels, reagents, and sample preparation protocols, to innovative software tools and data analysis approaches (Kalina et al., 2012; van Dongen et al., 2012) (www.euroflow.org). Subsequently, a EuroFlow Quality Assessment (QA) program was developed in order to control for the quality of data in individual laboratories that have implemented the EuroFlow standard operating procedures (SOPs) (Kalina et al., 2015).

Of note, full intra- and inter-laboratory reproducibility is also critical to implement the novel EuroFlow software-guided approaches for automated classification of individual cells into specific cell populations, since this concept is based on direct comparison of flow cytometry data files against well-defined databases of data files acquired in different laboratories following standardized protocols (Pedreira et al., 2013). Such automated analysis of locally acquired data files against the EuroFlow databases might facilitate handling of increasingly complex multicolor FCS data files and ease the classification of tumor cell populations into specific disease entities (Costa et al., 2010). Consequently, to take full advantage of the EuroFlow approaches, the measurements of flow cytometric fluorescence and light scatter signals need to be fully standardized.

Although the EuroFlow standardized approach was originally developed for BD FACSCanto II (BD Biosciences, San Jose, CA), it has been successfully applied to other 8-color instruments available in 2006–2009, such as the BD LSRII (BD Biosciences) and Cyan ADP (DakoCytomation) (Kalina et al., 2012). More recently, the EuroFlow Consortium also proposed a standardized instrument setup procedure for the Navios flow cytometer (Beckman Coulter, Miami, FL) (available at www.euroflow.org), thus enabling generation of highly comparable data across different flow cytometric platforms (Novakova et al., J Immunol Methods, this issue).

The Swiss Cytometry Society (SCS) is dedicated to continuously pursuing improvement in quality of flow cytometric measurements in clinical laboratories across Switzerland. In 2014, the SCS started an educational and training program aiming at standardization of flow cytometric measurements to increase reproducibility and inter-laboratory comparability of immunophenotypic data in clinical laboratories in Switzerland. Since the BD FACSCanto II and BC Navios instruments are almost equally represented across clinical laboratories in Switzerland, a cross-platform standardization was considered a prerequisite.

Here, we report on the results of the cross-platform standardization feasibilty study performed in 10 different clinical laboratories across Switzerland using either BD FACSCanto II or BC Navios instruments. In addition, we also report on the specific technical issues encountered and provide information on troubleshooting.

2. Materials and methods

2.1. Study outline

Standardized cytometer setup and sample preparation protocols established by the EuroFlow Consortium were implemented by the 10 SCS clinical laboratories for this study (Kalina et al., 2012). Moreover, the EuroFlow quality assessment (QA) scheme (Kalina et al., 2015) was adopted in order to evaluate quality of the measured data. The only laboratory with previous experience with EuroFlow standardization (Aarau) coordinated the study. Whenever required by the participating laboratories, the coordinating laboratory provided assistance or training in the standardized EuroFlow procedures. The quality of data for the EuroFlow Lymphoid Screening Tube (LST) measurements was tested, using both fresh (< 24 h) peripheral blood from a healthy donor distributed by the coordinating laboratory (round 1) and locally collected fresh (< 24 h) peripheral blood samples from 3 healthy donors (rounds 2 and 3) as a model. Evaluation of the data acquired was based on the deviation of the measured median fluorescence intensity (MedFI) of individual markers on predefined lymphocyte subsets from the expected MedFI established for the EuroFlow QA program and expressed as “Performance score” (P-score) (Kalina et al., 2015). The corresponding coeficients of variation (CV) for MedFI were calculated. In total, three rounds were performed and results of each study round were presented at the SCS meetings, where pitfalls were also discussed. Subsequently, specific troubleshooting information was mailed to the participating laboratories and published on the SCS website (www.cytometry.ch).

2.2. Study rounds

The first round (learning phase) of the study took place in January 2014 and eight clinical laboratories using either BD FACSCanto II (Aarau, Basel Hematology, Basel Immunology, Bellinzona, Bern, Lucerne) or BC Navios instruments (Geneva, St. Gallen) participated. Two independent QA rounds (analytical phase) followed in 2014 and 2015 in which nine clinical laboratories using BD FACSCanto II ($n = 6$) or BC Navios ($n = 3$) participated. One laboratory participating in the

first round with a BC Navios instrument (St. Gallen) was not able to participate in the subsequent study rounds, and two new laboratories using BC Navios (Lausanne, Zurich) joined the study for rounds 2 and 3. One of the two new laboratories required assistance by the coordinating laboratory to perform instrument setup.

2.3. Blood samples

For the first round, an EDTA peripheral blood sample from a healthy donor was sent by the coordinating laboratory to each participating laboratory. Staining of the sample in participating laboratories was performed within 24 h after blood collection. For rounds 2 and 3, EDTA peripheral blood samples from three healthy donors were obtained locally at each participating laboratory and processed within 24 h after blood collection. The blood samples were obtained in accordance with The Swiss Human Research Act and with approval of the local ethical committee.

2.4. Reagents

The LST antibody cocktail – a lyophilized 8-color, 12-antibody premixed reagent combination (CYT-LST) – kindly provided by Cytognos SL (Salamanca, Spain) was used in all three study rounds. The composition of the CYT-LST cocktail is as follows:

CD4 (clone RPA-T4) Pacific Blue, CD20 (clone 2H7) Pacific Blue, CD45 (clone GA90) OC515, CD8 (clone UCHT-4) FITC, polyclonal anti-IgLambda (IgL) FITC, CD56 (clone C5.9) PE, polyclonal anti-IgKappa (IgK) PE, CD5 (clone UCHT-2) PerCP Cy5.5, CD19 (clone 19–1) PE-

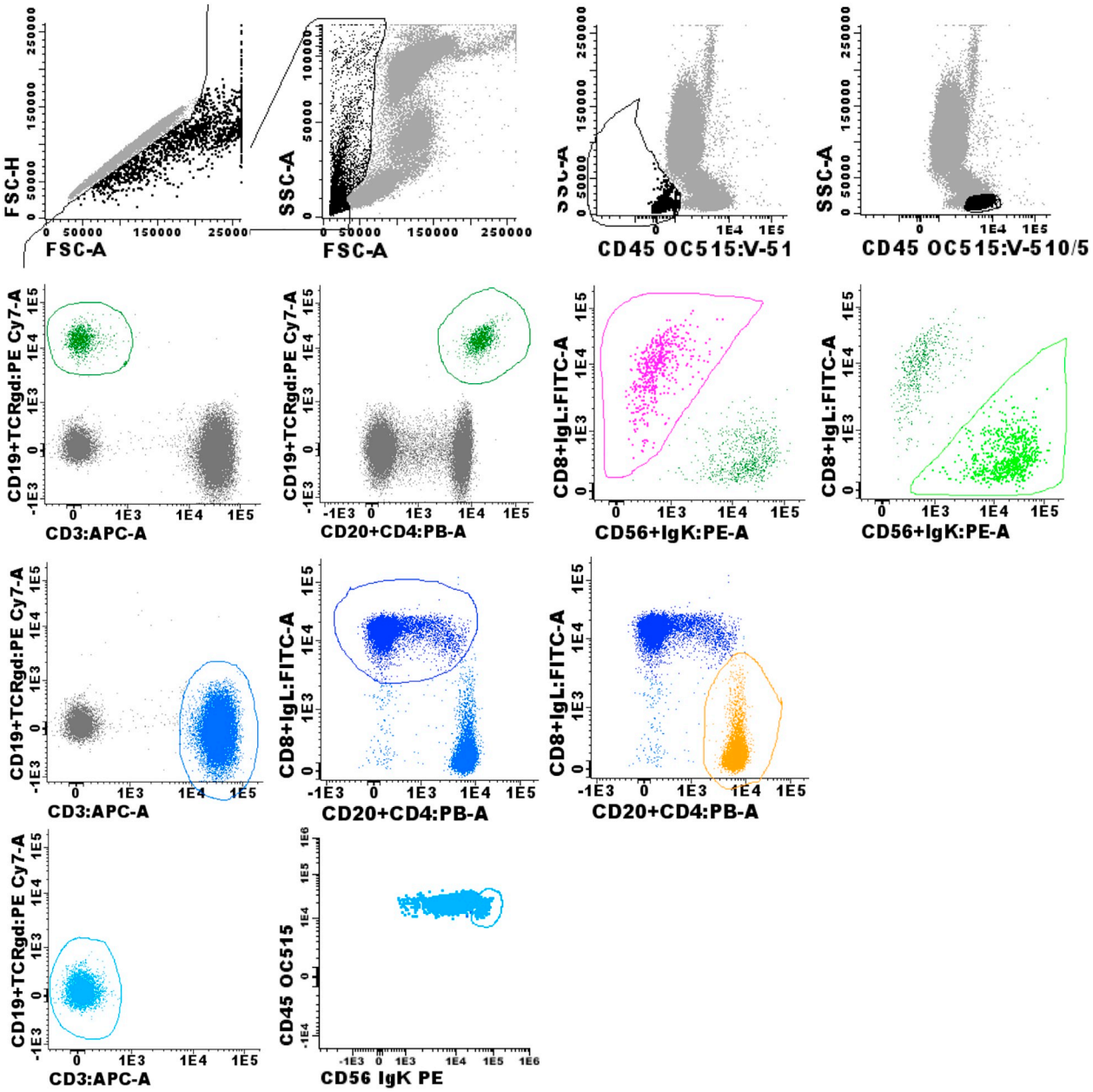

Fig. 1. LST gating strategy.
Lymphocytes were selected upon eliminating doublets, debris, and non-leukocytic (CD45-negative) cells (upper row). Lymphocytes were gated as CD45-positive and SSC^{low} events in CD45 vs. SSC dot plot. Lymphocyte subpopulations were identified using relevant markers: CD19-positive B-cells represented in dark green were divided into IgL-positive (pink) and IgK-positive (light green) subpopulations. CD3-positive T-cells (blue) were divided into CD8-positive (dark blue) and CD4-positive (orange). NK-cells (light blue) were identified as CD3/CD19-double negative and CD56-positive events and the CD56-bright subpopulation of NK-cells was gated for analysis.

Table 1
MedFI variation during study.

Marker	Cell subset	Average coefficients of variation of MedFI [%]		
		Round 1 (n = 8)	Round 2 (n = 27)	Round 3 (n = 27)
CD20 PacB	B-cells	27.8	14.2	14.1
CD4 PacB	CD4pos T-cells	24.5	9.9	7.3
CD45 OC515	T-cells	38.5	23.4	17.1
CD8 FITC	CD8pos T-cells	17.8	13.0	21.3
IgL FITC	IgLpos B-cells	24.8	45.0	38.1
CD56 PE	CD56bright NK-cells	24.5	55.6	34.0
IgK PE	IgKpos B-cells	47.5	80.0	59.2
CD5 PerCP Cy5-5	T-cells	42.4	28.1	21.7
CD19 PE Cy7	B-cells	27.8	21.9	16.6
CD3 APC	T-cells	22.2	22.5	28.9
CD38 APC C750	CD56bright NK-cells	38.0	48.1	47.1

Cyanine7 (PE Cy7), TCRγδ (clone TCR-1) PE Cy7, CD3 (clone 33-2A3) APC, CD38 (clone LD38) APC C750. The EuroFlow-validated lot (EAE01) of Rainbow Calibration Particles, 8 peaks (Spherotech, Lake Forest, IL; referred thereafter in this manuscript to as Rainbow beads) was used in all participating laboratories throughout the study. All the required reagents were centrally acquired and distributed to the participating laboratories by the coordinating laboratory.

2.5. Staining procedure

Specimen staining and erythrocyte lysis (BD FACS Lysing solution, BD Biosciences) were performed using a stain/lyse/wash method as described by the EuroFlow SOP for sample preparation and staining (Part A: Common initial procedure when the EuroFlow antibody panel includes SmIg staining and Part C: staining for surface markers only; www.euroflow.org). All samples were acquired within 1 h after they had been immunostained, using local flow cytometers.

2.6. Instruments

BD FACSCanto II flow cytometers were used in six laboratories and BC Navios flow cytometers in four laboratories. All 3-laser instruments were equipped with their standard optical configuration for a blue laser emitting at 488 nm, a red laser emitting at 633 nm (BD FACSCanto II) or 638 nm (BC Navios), and a violet laser emitting at 405 nm. Of note, BC Navios instruments are equipped with two additional photomultiplier tubes (PMT) allowing detection of signals emitted by other fluorochromes. Although the optical configuration of the two cytometers is generally similar, BC Navios differs substantially from BD FACSCanto II flow cytometers in collection of the OC515 signal (BP550/40 band pass filter vs. BP510/50 on BD FACSCanto II).

2.7. Instrument setup and compensation

Standardized instrument setup of the BD FACSCanto II and BC Navios instruments was performed according to the EuroFlow SOP for instrument setup and compensation (www.euroflow.org). Briefly, the voltage of each PMT was set to reach the target mean fluorescence intensity (MFI) for the 7th peak of the EuroFlow-validated Rainbow beads lot (see www.euroflow.org). Fluorescence compensation was calculated using Diva v6 (BD Biosciences) or Kaluza v1.2 or later (Beckman Coulter) software, based on a defined set of single stained peripheral blood compensation tubes as proposed by EuroFlow (see www.euroflow.org).

2.8. Data analysis

Central analysis of all data files obtained on BD FACSCanto II (FCS files) and BC Navios (LMD files) was performed by the coordinating laboratory on "merged" data files (data files were combined into one metafile for analysis and graphical presentation) using the Infinicyt software v1.7 (Cytognos SL). Prior to merging FCS and LMD data files in Infinicyt, the data were automatically rescaled by the software to a common scale allowing simultaneous analysis of the digitalized signals from distinct instruments. For data analysis, a predefined gating strategy (Fig. 1) was used. Briefly, after excluding doublets and cell debris, peripheral blood lymphocytes were gated in the side scatter (SSC^{low}) and $CD45^{high}$ area. Within lymphocytes, B-cells were gated as CD19/CD20-double positive events and subsequently divided into IgK-positive and IgL-positive subpopulations, T-cells were defined as CD3-positive events and further gated into CD4-positive and CD8-positive T-cells, and NK-cells were identified as CD3/CD19-double negative events expressing CD56. Only $CD56^{bright}$ NK-cells were considered for analysis in order to reduce heterogeneity in CD56 and CD38 expression by these cells. For each of the above-defined lymphocyte subsets, MedFI values for specific markers expressed by the gated lymphocyte subpopulations were recorded and MedFI CVs were calculated (Table 1). Evaluation of TCRγδ staining on T cells was not performed, as no P-score criteria were defined for this subset in EuroFlow QA program (Kalina et al., 2015). All gated lymphocyte subsets were displayed against the file numbers for all fluorescence channels and data was visually inspected for any potentially altered staining pattern. The presence of MedFI outlier values for individual samples was identified following P-score criteria and the problems causing these inconsistencies were tracked down.

2.9. Statistical methods

The coefficients of variation (CV) of MedFI values recorded for each marker analyzed on pre-defined lymphocyte subsets in individual FCS and LMD files were calculated as:

CV = (standard deviation) / (average of MedFI) × 100 [%].

To quantify deviations from reference values the "Performance score" (P-score) metrics was used as adopted by Kalina et al. (2015) for EuroFlow QA program:

P-score = ($\log_{10}$MedFI-$\log_{10}$qaMedFI)/Dmax.

Briefly, P-score values were calculated as the difference of the logarithmic transformation of an actual MedFI value from the logarithmic transformation of all MedFI values from all QA rounds (qaMedFI) and divided by Dmax. Dmax is the maximal allowed difference calculated as the 95th percentile of all absolute values of difference of actual MedFI from qaMedFI. To follow less stable values (IgL FITC, and IgK PE; CV repeatedly > 60%), 90th percentile values were used instead. Thus, P-score represents acceptance criteria for the MedFI variation of each marker evaluated on its corresponding cell population. A two-tailed Student's t-test ($\alpha = 0.05$) was used to assess differences in MedFI values between BD FACSCanto II and BC Navios instruments.

3. Results

3.1. Round 1 (learning phase)

Median fluorescence intensity (MedFI) values for individual markers on well-defined lymphocyte subsets showed an average CV of 30.3% (range: 17.8%–47.5%) among all measurements. The CV of MedFI for 7/11 markers was below 30% (Table 1). P-scores allowed for rapid identification of MedFI outliers. Over 90% of all MedFI values in 6/8 laboratories (75% of measured samples/files) performed within the P-score acceptance criteria (Table 2). Only one laboratory scored under 70% of MedFI values within P-score ranges not acceptable for QA, due to failure to perform correct instrument setup (see 3.3 Pitfalls Section,

Table 2
Summary results of all study rounds showing number of MedFI values within range for the acceptable variation (P-score).

Correct MedFI values within P-score criteria per laboratory			
Laboratory	Round 1	Round 2	Round 3
1	11/11 (100.0%)	32/33 (97.0%)	33/33 (100.0%)
2	10/11 (90.9%)	31/33 (93.9%)	31/33 (93.9%)
3	9/11 (81.8%)	32/33 (97.0%)	31/33 (93.9%)
4	11/11 (100.0%)	33/33 (100.0%)	33/33 (100.0%)
5	7/11 (63.6%)	29/33 (87.9%)	31/33 (93.9%)
6	11/11 (100.0%)	30/33 (90.9%)	33/33 (100.0%)
7	11/11 (100.0%)	32/33 (97.0%)	33/33 (100.0%)
8	11/11 (100.0%)	N/A	N/A
9	N/A	32/33 (97.0%)	29/33 (87.9%)
10	N/A	33/33 (100.0%)	33/33 (100.0%)

and Fig. 2). No significant differences in MedFI between the data files measured on BD FACSCanto II and BC Navios instruments were found (see Table 3 for MedFI summary). Compensation errors were detected in data files from 5 laboratories (see Fig. 2B and 3.3 Pitfalls Section for details).

3.2. Rounds 2 and 3 (analytical phase)

The results of rounds 2 and 3 showed that highly comparable data can be obtained across different laboratories using two different flow cytometry platforms. However, inconsistent MedFI values in individual laboratories were identified with P-score in both rounds. Symmetrical shifts for MedFI values in triplicate samples from an individual laboratory would typically suggest systematic problems in that particular laboratory, whereas individual outliers are rather related to the individual sample/donor. In each of the two rounds, 8/9 laboratories (89%) scored above 90% values correct (within allowed P-score ranges) and one laboratory scored 88% values correct (Table 2). No significant differences in MedFI values for the analyzed lymphocyte subsets were found among data files measured on BD FACSCanto II and BC Navios instruments (see Table 3 for MedFI summary). Incorrect compensation was observed in both, round 2 and round 3 (see 3.3 Pitfalls Section for details).

In round 2, the MedFI values for markers on gated lymphocyte subsets were distributed with an average CV of 32.9% (range: 9.9%–80.0%). Higher CVs were observed for IgK and IgL light chains (80% and 45%, respectively) and for CD56 and CD38 on CD56^{bright} NK-cells (55.6% and 48.1%, respectively). The CV of MedFI for 7/11 markers systematically performed below 30%. In total, 12 individual MedFI outliers were identified, one for each evaluated antigen and corresponding lymphocyte subpopulation, with exception for CD4 (Pacific Blue channel) on T cells where 2 MedFI outliers in data files from two different laboratories were identified. Further, decreased MedFI values for IgK and IgL on B cells were observed in data files from one laboratory (see 3.3 Pitfalls Section for more details).

In round 3, the MedFI values for individual markers on gated lymphocyte subsets were distributed with an average CV of 27.8% (range: 7.3%–59.2%). The highest CV was observed for IgK light chains (59.2%) on B-cells followed by CD38 on CD56^{bright} NK-cells (47.1%), IgL light chains on B-cells (38.1%) and CD56 on CD56^{bright} NK-cells (34%). These results were comparable to those achieved by experienced EuroFlow laboratories (Kalina et al., 2015). In total, 10 MedFI outliers were identified, one for each evaluated antigen and corresponding lymphocyte subpopulation, with exception for CD56 (PE channel) on CD56^{bright} NK-cells where no MedFI outlier was detected.

Importantly, an overall clear improvement in the quality of data was achieved in rounds 2 and 3 vs. round 1, as documented by the overall performance of the evaluated MedFI values for each laboratory (Table 2).

3.3. Pitfalls

During central data analysis, Rainbow bead files acquired in individual laboratories during instrument setup were merged and analyzed prior to the LST data files. In the learning phase (round 1), lower than expected MedFI values for the 7th peak of Rainbow beads were observed in all fluorescence channels within one data file (Fig. 2A, laboratory 5). When closely evaluated, all 8 peaks of the Rainbow beads appeared to be shifted towards lower fluorescence intensities in all fluorescence channels, and the MedFI of the 8th peak of the Rainbow beads seemed to be aligned with the 7th peak of the Rainbow beads in the files measured in other laboratories (Fig. 2A: laboratory 5 vs. laboratories 1–4 and 6–8). Indeed, in that particular laboratory, the target mean fluorescence intensity values for the 7th peak of the Rainbow beads were applied to the 8th peak of the beads instead during instrument setup, resulting in incorrect photomultiplier (PMT) voltage settings. These erroneous instrument settings where then used to acquire the LST-stained peripheral blood sample which resulted in lower MedFI values for all evaluated markers on gated lymphocyte subpopulations (Fig. 2B; data file 5) as documented by rather poor performance results of laboratory 5 in round 1 (Table 2).

Further, three laboratories (all BD FACSCanto II users) failed to collect data on the height values for the light scatter signal (FSC-H) that allow for gating of singlets and discrimination of doublets, respectively (data not shown).

Most importantly, compensation errors were detected in all three study rounds. Overcompensated data could be readily identified since the MedFI values of those cell populations not expressing a given marker typically shift towards negative values and pile on the axis when a logarithmic scale is used to display MedFI of the data against file number in Infinicyt software. In round 1, various markers in data files from 5 different laboratories appeared to be overcompensated (see Fig. 2B, black arrows). In one center, overcompensation for APC spillover in the APC C750 channel resulted in a diminished CD38 signal on NK-cells (data file 3 in Fig. 2B). During rounds 2 and 3, overcompensated data could be spotted as symmetrical shifts in MedFI for all three samples acquired in one laboratory. In round 2, data files acquired in 6 participating laboratories exhibited improper compensation for spillover in one or more of the following fluorescence channels: FITC, PE, and PE Cy7 (data not shown). In round 3, data files acquired in one participating laboratory showed improper compensation for spillover in both FITC and PerCP Cy5.5 channels (see Fig. 3, data files 16–18), data files in two other laboratories exhibited erroneous compensation for spillover in PE Cy7 channel (Fig. 3, data files 1–3 and 10–12).

Moreover, deviations from the EuroFlow SOP for sample preparation and staining could be identified, most strikingly insufficient washing of the plasma from peripheral blood before staining with antibodies against surface IgK and IgL light chains. This resulted in decreased labeling of IgK and IgL light chains on the surface of B cells and subsequently, in lower MedFI signals measured due to antibody competition with the abundant immunoglobulin molecules in the remaining plasma, as illustrated in Fig. 4 where a data file with poor performance score obtained in round 2 is displayed against a reference data file.

In addition, two BD FACSCanto II laboratories submitted data for central analysis in a wrong format (“Diva Experiment” files instead of FCS files). “Diva Experiment” files do not contain complete information about the data as part of the information is saved separately within an XML file which is exported together with “Diva Experiment” upon data acquisition in Diva software. When third party software is used for data analysis, the piece of information saved within the XML file cannot be directly read by the software and the data might not be displayed correctly. In this particular case we observed swapped FSC-W vs. FSC-H and SSC-W vs. SSC-H parameters. Although the “Diva Experiment” files bear an .fcs extension, they can be discriminated from FCS files by the file name reading a consecutive number assigned to the files in Diva

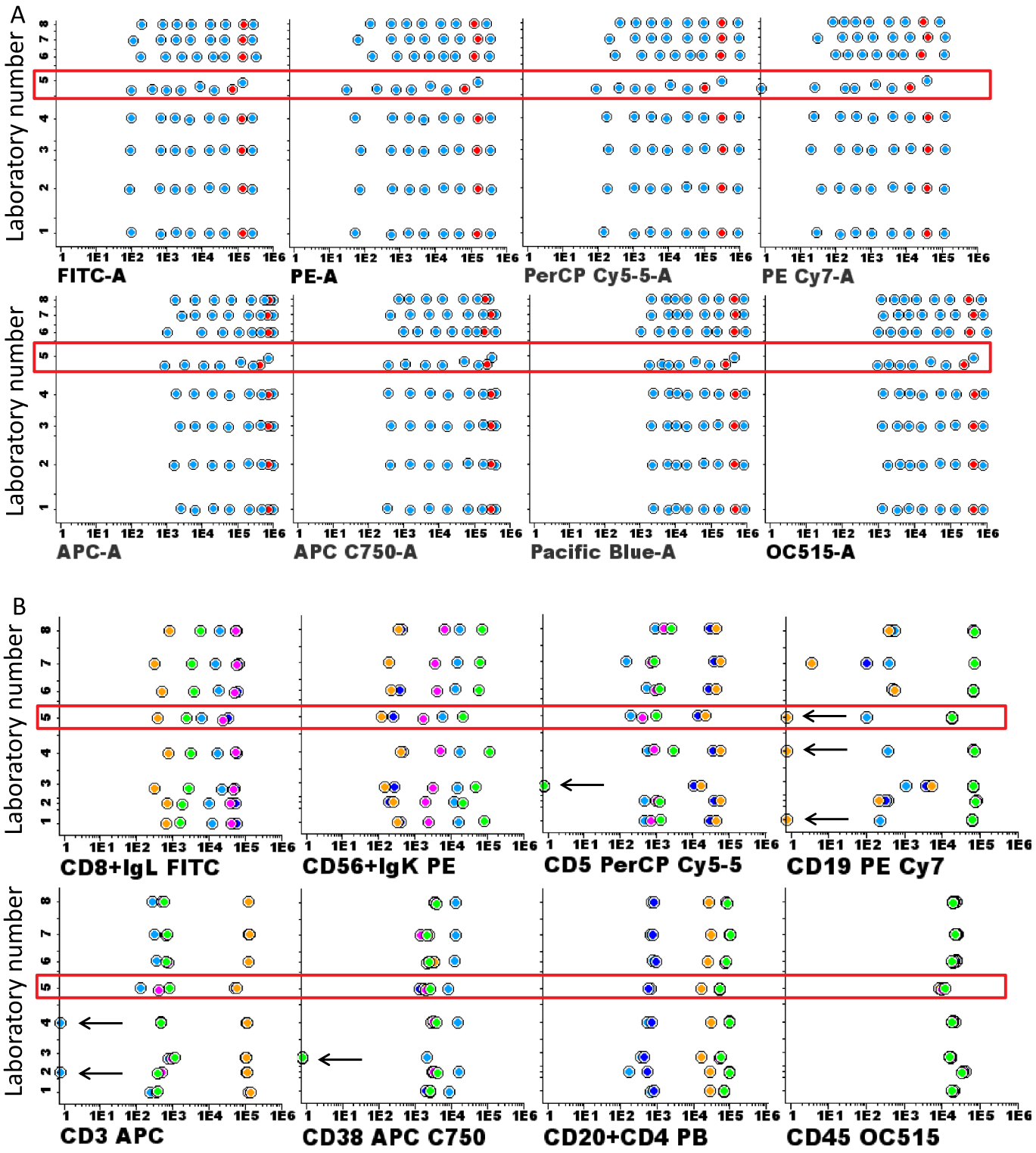

Fig. 2. Incorrect instrument setup during round 1.

Panel 2A shows results of cytometer setup using Rainbow beads. Dots represent MedFI of each of the 8 peaks of the Rainbow beads in their respective fluorescence channels. 7th peak (reference) is highlighted in red. Laboratory 5 performed erroneous cytometer setup by applying the target values for the 7th peak of the Rainbow beads to the 8th peak instead (highlighted in the red rectangle).

Panel 2B shows results of LST stained peripheral blood. Dots represent MedFI of each evaluated antigen/lymphocyte subpopulation pair. Each laboratory acquired one peripheral blood sample. Erroneous instrument setup performed in laboratory 5 was used for acquisition of LST tube resulting in suboptimal MedFI values (highlighted in the red rectangle). Overcompensated data can be identified piling on the vertical axes (indicated by black arrows). Color codes: IgL-positive B-cells (pink), IgK-positive B-cells (light green), CD8-positive T-cells (dark blue), CD4-positive T-cells (orange), NK-cells (light blue). Laboratories 1–5 and 7 used BD FACSCanto II and laboratories 6 and 8 used BC Navios instruments.

Table 3
Summary of MedFI values obtained for analyzed lymphocyte subset and antigen pairs on Canto vs. Navios instruments.

Marker	Cell subset	Average of MedFI Canto vs. Navios					
		Round 1 (n = 8)		Round 2 (n = 27)		Round 3 (n = 27)	
		Canto (n = 6)	Navios (n = 2)	Canto (n = 6)	Navios (n = 3)	Canto (n = 6)	Navios (n = 3)
CD20 PacB	B-cells	19,827.10	20,450.23	20,538.95	18,296.60	21,331.20	20,611.69
CD4 PacB	CD4pos T-cells	6345.70	6455.06	6275.69	5765.78	5981.54	5692.93
CD45 OC515	T-cells	5385.50	5651.04	5165.63	5033.88	5504.96	6424.42
CD8 FITC	CD8pos T-cells	13,740.04	15,484.07	16,045.21	17,214.08	16,022.69	17,449.69
IgL FITC	IgLpos B-cells	11,201.20	13,290.08	12,097.85	13,635.10	13,793.70	14,177.94
CD56 PE	CD56bright NK-cells	13,362.45	15,379.33	3814.47	2721.99	3766.45	4464.17
IgK PE	IgKpos B-cells	14,527.12	16,834.99	16,386.37	13,221.11	19,151.13	15,104.61
CD5 PerCP Cy5–5	T-cells	8403.47	8390.18	12,557.83	9449.73	9800.40	7661.82
CD19 PE Cy7	B-cells	15,079.81	16,930.35	14,657.23	15,163.43	18,174.25	16,382.86
CD3 APC	T-cells	26,350.01	30,497.32	35,157.85	37,768.38	42,762.92	42,769.42
CD38 APC C750	CD56bright NK-cells	2723.66	3300.06	1649.62	2639.38	3647.42	3557.81

software. Data files from these laboratories were included in the central analysis after re-submission in the correct FCS file format.

4. Discussion

In this study, we tested the feasibility of standardized flow cytometric measurement of 8-color data in clinical laboratories, following the EuroFlow SOP, using two different flow cytometry platforms: BD FACSCanto II and BC Navios.

For this purpose, we chose the LST antibody combination, as it represents one of the most used ones in hematooncology and it is readily available as EuroFlow-approved pre-mixed cocktail (www.cytognos.

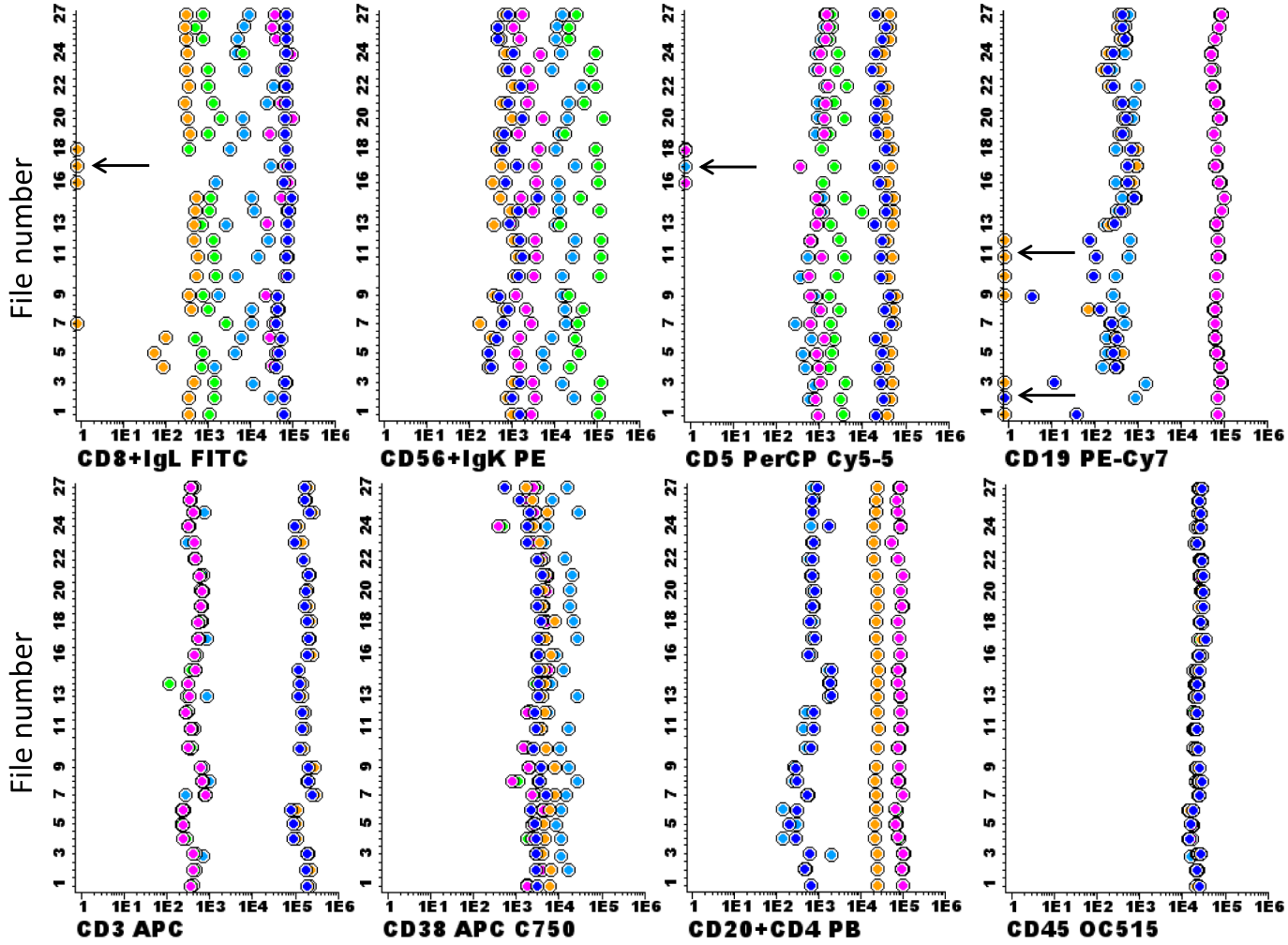

Fig. 3. Results of round 3.
Dots represent MedFI of each evaluated antigen/lymphocyte subpopulation pair. Each laboratory acquired three peripheral blood samples stained with LST tube. Highly comparable data could be obtained using BD FACSCanto II (files 1–18) and BC Navios (files 19–27) cytometers. Black arrows indicate overcompensated data. Color codes: IgL-positive B-cells (pink), IgK-positive B-cells (light green), CD8-positive T-cells (dark blue), CD4-positive T-cells (orange), NK-cells (light blue).

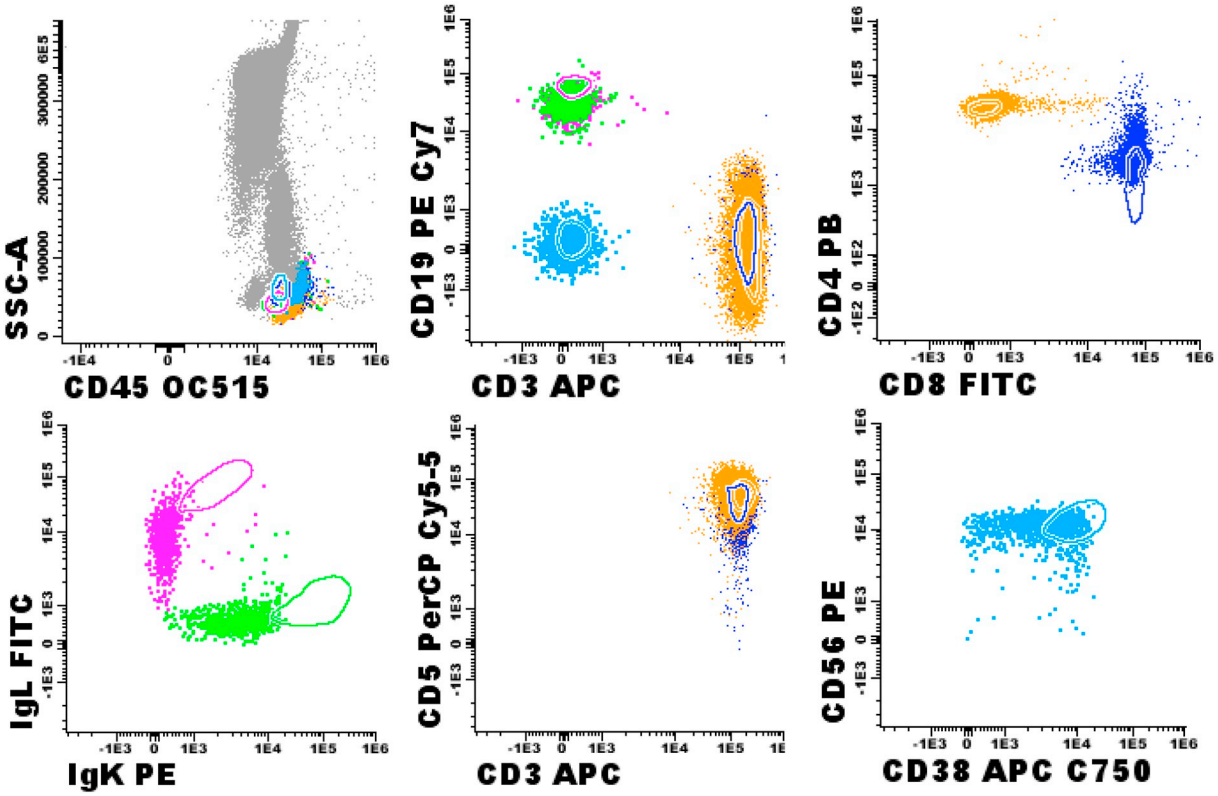

Fig. 4. An example of a data file with a performance score of 64% compared to a reference data file.
Example of a data file with a poor MedFI performance score (results represented as dots) obtained in round 1 is displayed against the expected expression patterns (shown as reference image contours). CD19, IgK and IgL on B-cells show lower fluorescence signal intensities than expected, as well as CD38 on NK-cells, whereas CD45 on T-cells shows higher signal intensity than expected. Color codes: IgL-positive B-cells (pink), IgK-positive B-cells (light green), CD8-positive T-cells (dark blue), CD4-positive T-cells (orange), NK-cells (light blue).

com), as well as dried tube (van der Velden et al., 2017). In order to eliminate the influence of lot-to-lot antibody variability on the MedFI readouts (Böttcher et al., 2017), we used the same lot of a lyophilized LST antibody mixture which was distributed to all participating laboratories by the coordinating laboratory. Moreover, during the learning phase (round 1) we opted to use aliquots of one healthy donor peripheral blood sample which was sent to participating laboratories, too. This allowed us to directly assess technical issues and non-adherence to SOP by the participants, as the impact of biological variability of the material as well as the variability of the reagents used could be neglected. Discussion of the results of the learning phase (round 1) with the participating laboratories allowed identification of critical steps in the standardized EuroFlow protocols, and also the most common errors in their execution.

In rounds 2 and 3 the MedFI values obtained for individual markers on gated lymphocyte subsets were distributed with an average CV of 32.9% and 27.8%, respectively. The CV of MedFI for 7/11 markers performed repeatedly under 30% in line with results of the EuroFlow expert laboratories (Kalina et al., 2015). Higher CVs were observed for IgK and IgL staining, as well as for CD38 and CD56 staining on $CD56^{bright}$ NK-cells, which might also translate a greater biological variability among B-cells and NK-cells from different donors for these specific markers, respectively (Kalina et al., 2015). However, higher CVs observed for IgK and IgL could also be, at least partially, caused by non-adherence to standard sample preparation procedures, as discussed hereafter.

Importantly, no significant differencies between BD FACSCanto II and BC Navios were observed. In fact, almost identical MedFI values were obtained on BD FACSCanto II and BC Navios instruments in different laboratories following the EuroFlow SOPs and the LST staining of peripheral blood. In the GEIL study, Solly et al. (2013) reported superimposable data between BD FACSCanto II and BC Navios instruments when using 8-peak Rainbow beads and single-stained CD16-positive granulocytes, however, they observed histogram shifts in the PE Cy7 and APC Cy7 channels. Inter-laboratory studies published by other groups did not attempt to fully standardize the measurements and to evaluate the measured fluorescence signals (Feller et al., 2013; Westers et al., 2012; Johansson et al., 2014). However, as the analysis and interpretation of flow cytometric immunophenotyping data is based on assessment of the level of expression of various markers on different cell populations in a (semi)-quantitative manner, the flow cytometric measurements need to be standardized and measured signal expressed as MedFI may be used to evaluate the data quality, as proved feasible here, for the first time in clinical laboratories in Switzerland.

The EuroFlow QA program (Kalina et al., 2015) was created to assess quality of flow cytometric data using MedFI values as readout, allowing for inter-laboratory comparison of standardized flow cytometry measurements. We adopted the EuroFlow QA program and performance score assessment, which allowed us to identify inconsistent MedFI values in individual participating laboratories. All inconsistencies were tracked down and underlying technical problems and/or non-adherence to the EuroFlow SOPs were identified. In study rounds 2 and 3, 89% of participants scored 90% (P-score) values within the acceptance criteria, in line with the results of the EuroFlow QA rounds performed among EuroFlow expert laboratories (Kalina et al., 2015). Interestingly, our results show that even laboratories without previous experience with the EuroFlow approach can reach highly comparable data by following the standardized EuroFlow protocols.

Despite this, several technical issues were identified during the learning phase (round 1). The most striking was failure to perform correct instrument setup when a wrong peak of the Rainbow beads was taken as a reference for target mean fluorescence intensity setup. This was, at least in part, due to the lack of previous experience with Rainbow beads in that particular laboratory. This erroneous cytometer

setup led to poor QA performance for MedFI values of this laboratory in round 1 (< 70% acceptable P-score values), but not in subsequent study rounds. No instrument setup problems were encountered in other participating laboratories including those without previous experience with the EuroFlow approach.

Non-adherence to standard sample preparation procedures could also be identified in some laboratories, namely insufficient removal of the plasma from peripheral blood before staining for surface IgK and IgL light chains. This resulted in low labeling of IgK and IgL light chains on the surface of B cells, due to interference of (abundant) plasma immunoglobulins. However, the adherence to the SOPs was improved upon discussion of the results with the participating laboratories, and variability decreased e.g. in round 3.

Another issue noted in several participating laboratories was failure to activate acquisition of the light scatter characteristics that would allow doublet discrimination during data analysis. However, this failure did not result in unacceptable QA performance.

Although the EuroFlow QA program adopted for the study was not designed to assess fluorescence compensation, significant deviations from the expected values could be readily detected. Thus incorrect compensation was frequently found through both study phases and all study rounds, consistent with those findings reported for the EuroFlow QA rounds performed in EuroFlow expert laboratories (Kalina et al., 2015).

A face-to-face participant meeting with discussion of the results and troubleshooting was organized after each study round as a part of the SCS meetings. All participating laboratories were provided with specific troubleshooting notifications addressing the problems encountered, their possible reasons and problem solutions. On-site (re)training was performed by the coordinating laboratory when considered necessary by the participating laboratories. Improvement of the QA results during the study rounds 2 and 3 compared to round 1 documents the importance of education, proper training and troubleshooting.

In summary, here we show that inter-laboratory cross-platform standardization of 8-color flow cytometry measurements is feasible in clinical laboratories without previous experience by using the EuroFlow standardization protocols and SOPs. Highly comparable (almost identical) data between BD FACSCanto II and BC Navios instruments was obtained following the EuroFlow standardized procedures for sample preparation and instrument setup. However, proper training of laboratory personnel and adherence to the SOPs are essential for obtaining the desired standardization outcome.

Acknowledgments

We acknowledge Cytognos SL (Salamanca, Spain) for providing the LST antibody cocktail (CYT-LST) free of charge for all three study rounds. We thank Dr. Tomas Kalina, Associate Professor, CLIP – Childhood Leukemia Investigation Prague, Department of Paediatric Hematology and Oncology, Second Faculty of Medicine, Charles University and University Hospital Motol, Prague, Czech Republic, for his valuable advice and input.

References

Böttcher, S., et al., 2017. Lot-to-lot stability of antibody reagents for flow cytometry. J. Immunol. Methods. http://dx.doi.org/10.1016/j.jim.2017.03.018.

Costa, E.S., 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. Leuk. Off. J. Leuk. Soc. Am. Leuk. Res. Fund UK 24, 1927–1933.

Dongen, J.J.M. Van, Orfao, A., 2012. EuroFlow: resetting leukemia and lymphoma immunophenotyping. In: Basis for companion diagnostics and personalized medicine. 26. pp. 1899–1907.

van Dongen, J.J.M., et al., 2012. EuroFlow antibody panels for standardized n-dimensional flow cytometric immunophenotyping of normal, reactive and malignant leukocytes. Leukemia 26, 1908–1975.

Feller, N., et al., 2013. Defining consensus leukemia-associated immunophenotypes for detection of minimal residual disease in acute myeloid leukemia in a multicenter setting. Blood Cancer J. 3, e129.

Flores-Montero, J., et al., 2017. Next generation flow (NGF) for highly sensitive and standardized detection of minimal residual disease in multiple myeloma. Leukemia. http://dx.doi.org/10.1038/leu.2017.29.

Heel, K., et al., 2017. Blood reviews developments in the immunophenotypic analysis of haematological malignancies. In: YBLRE. 27. pp. 193–207.

Johansson, U., et al., 2014. Guidelines on the use of multicolour flow cytometry in the diagnosis of haematological neoplasms. Br. J. Haematol. 165, 455–488.

Kalina, T., et al., 2012. EuroFlow standardization of flow cytometer instrument settings and immunophenotyping protocols. Leukemia 26, 1986–2010.

Kalina, T., et al., 2015. Quality assessment program for EuroFlow protocols: summary results of four-year (2010 − 2013) quality assurance rounds. Cytometry. Part A 87, 145–156.

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

Solly, F., et al., 2013. Comparable flow cytometry data can be obtained with two types of instruments, Canto II, and Navios. A GEIL study. Cytometry. Part A 83, 1066–1072.

Theunissen, P., et al., 2016. Standardized flow cytometry for highly sensitive MRD measurements in B-cell acute lymphoblastic leukemia. Blood 129, 347–358.

van der Velden, V.H.J., et al., 2017. Optimization and testing of dried antibody tube: the EuroFlow LST and PIDOT tubes as examples. J. Immunol. Methods. http://dx.doi.org/10.1016/j.jim.2017.03.011.

Westers, T.M., et al., 2012. Standardization of flow cytometry in myelodysplastic syndromes: a report from an international consortium and the European LeukemiaNet Working Group. Leukemia 26, 1730–1741.