ISAC CYTOMETRY Journal of Quantitative Cell Science PART A

MIFlowCyt

Standardization of Flow Cytometric Immunophenotyping for Hematological Malignancies: The FranceFlow Group Experience

Françoise Solly,[1,2†] Fanny Angelot-Delettre,[3,4†] Michel Ticchioni,[5] Franck Geneviève,[6] Hubert Rambaud,[6] Lucile Baseggio,[7] Adriana Plesa,[7] Agathe Debliquis,[8] Francine Garnache-Ottou,[3,4] Anne Roggy,[3] Lydia Campos,[1] Carmen Aanei,[1] Alessandra Rosenthal-Allieri,[5] Marie-Thérèse Georget,[9] Sébastien Lachot,[9] Marie-Christine Jacob,[10] Nelly Robillard,[11] Soraya Wuilleme,[11] Elisabeth Andre-Kerneis,[12] Edouard Cornet,[13] Véronique Salaun,[14] Hind Bennami,[15] Anne-Catherine Lhoumeau,[16] Christine Arnoulet,[17] Hugues Jacqmin,[18] Nicolas Neyman,[18] Véronique Latger-Cannard,[19] Frédéric Massin,[19] Elodie Lainey,[20] Magali Le Garff-Tavernier,[21] Myrto Costopoulos,[21,22] Mickael Roussel,[23] Caroline Mayeur-Rousse,[24] Alice Eischen,[24] Victoria Raggeneau,[25] Coralie Derrieux,[26] Maxime Maurer,[27] Vahid Asnafi,[28,29] Amélie Trinquand,[28,29] Chantal Brouzes,[28,29‡] Ludovic Lhermitte[28,29‡*]

[1]CHU de St Etienne, Laboratoire d'hématologie, Saint-Etienne, France

[2]CHU Vaudois-Lausanne, Laboratoire d'hématologie, Lausanne, Switzerland

[3]Etablissement Francais du Sang Bourgogne Franche-Comte, laboratoire d'hématologie, F-25000 Besançon, France

[4]Université Bourgogne Franche-Comté, INSERM, EFS BFC, UMR1098, Interactions Hôte-Greffon-Tumeur/Ingénierie Cellulaire et Génique, F-25000, Besançon, France

[5]CHU de Nice, Laboratoire d'Immunologie, Nice, France

[6]CHU Angers, Laboratoire d'Hématologie Fédération Hospitalo-Universitaire GOAL, Grand Ouest Against Leukemia, Angers, France

[7]Centre Hospitalier Lyon Sud Hospices civils de Lyon, Laboratoire d'hématologie, Lyon, France

Abstract

Flow cytometry is broadly used for the identification, characterization, and monitoring of hematological malignancies. However, the use of clinical flow cytometry is restricted by its lack of reproducibility across multiple centers. Since 2006, the EuroFlow consortium has been developing a standardized procedure detailing the whole process from instrument settings to data analysis. The FranceFlow group was created in 2010 with the intention to educate participating centers in France about the standardized instrument setting protocol (SOP) developed by the EuroFlow consortium and to organise several rounds of quality controls (QCs) in order to evaluate the feasibility of its application and its results. Here, we report the 5 year experience of the FranceFlow group and the results of the seven QCs of 23 instruments, involving up to 19 centers, in France and in Belgium. The FranceFlow group demonstrates that both the distribution and applicability of the SOP have been successful. Intercenter reproducibility was evaluated using both normal and pathological blood samples. Coefficients of variation (CVs) across the centers were <7% for the percentages of cell subsets and <30% for the median fluorescence intensities (MFIs) of the markers tested. Intracenter reproducibility provided similar results with CVs of <3% for the percentages of the majority of cell subsets, and CVs of <20% for the MFI values for the majority of markers. Altogether, the FranceFlow group show that the 19 participating labs might be considered as one unique laboratory with 23 identical flow cytometers able to reproduce identical results. Therefore, SOP significantly improves reproducibility of clinical flow in hematology and opens new avenues by providing a robust companion diagnostic tool for clinical trials in hematology.

Key terms

immunophenotyping; instrument settings; standardization; quality controls; flow cytometry; hematology; FranceFlow

FLOW cytometric immunophenotyping has proven essential in clinical laboratories for the diagnosis, classification, and monitoring of hematological diseases (1). However, flow cytometry (FC) is often challenged as it is commonly recognized as a highly expert-

ISAC INTERNATIONAL SOCIETY FOR ADVANCEMENT OF CYTOMETRY

[8]Groupe Hospitalier de la région Mulhouse sud Alsace, Laboratoire d'Hématologie, Mulhouse, France

[9]CHRU de Tours, Service d'Hématologie Biologique, Tours, France

[10]CHU de Grenoble, Immunology, Grenoble, France

[11]CHU de Nantes, Laboratoire d'hématologie, Nantes, France

[12]Centre Hospitalier Henri Duffaut, Laboratoire d'Hématologie, Avignon, France

[13]CHU Caen, Laboratoire d'Hématologie; INSERM U1245, Group Biomarkers of lymphoma and solid Tumors Equipe MICAH, Caen, France

[14]Centre Hospitalier Universitaire de Caen, Laboratoire d'hématologie, Caen, France

[15]Institut Curie, Laboratoire d'hématologie, Saint-Cloud, France

[16]CRCM, Aix Marseille Univ, CNRS, Inserm, Marseille, France

[17]Institut Paoli-Calmettes, Département de Biologie du cancer, Institut Paoli-Calmettes, Département de Biologie du cancer, Marseille, France

[18]CHU UCL Namur, Laboratoire d'Hématologie, Namur, Belgium

[19]Centre Hospitalier Universitaire de Nancy, Plateforme de cytometrie en flux, Nancy, France

[20]Hôpital Robert Debré- APHP, Service d'Hématologie Biologique, Hôpital Robert Debré- APHP, UMRS_1131- Institut Universitaire d'Hématologie, Paris, France

[21]Groupe Hospitalier Pitié-Salpêtrière, UF Phénotypage des Hémopathies, Centre d'Ecologie Cellulaire, Service d'hématologie Biologique, Paris, France

[22]INSERM U1138, Programmed Cell Death and Physiopathology of Tumor Cells Centre de recherche des cordeliers, Paris, France

[23]CHU de Rennes, Laboratoire Hématologie, Rennes, France

[24]CHU Strasbourg, Laboratoire d'hématologie, Strasbourg, France

[25]CH Versailles, Service de biologie médicale, Unité d'hématologie et de cytogénétique onco hématologique, Versailles, France

[26]Grand Hôpital de l'Est Francilien, Laboratoire d'Hématologie, Meaux, France

[27]Hôpitaux Civils de Colmar, Laboratoire d'Hématologie, Colmar, France

[28]Université Paris Descartes Sorbonne Cité, Institut Necker-Enfants Malades (INEM), Institut National de recherche Médicale (INSERM) U1151, Paris, France

[29]Hôpital Necker Enfants-Malades, Laboratory of Onco-Hematology, Assistance Publique-Hôpitaux de Paris (AP-HP), Paris, France

Received 27 December 2018; Revised 4 June 2019; Accepted 5 June 2019

Additional Supporting Information may be found in the online version of this article.

*Correspondence to: Ludovic Lhermitte, Université Paris Descartes Sorbonne Cité Institut Necker-Enfants Malades (INEM), Institut National de recherche Médicale (INSERM) U1151, Paris, France. Email: ludovic.lhermitte@aphp.fr

†These authors contributed equally to this work
‡These authors contributed equally to this work
‡*Corresponding author

Published online in Wiley Online Library (wileyonlinelibrary.com)

DOI: 10.1002/cyto.a.23844

dependent technique with little reproducibility across multiple centers (2). Therefore, its use in translational medicine and clinical trials has mainly been hampered because not considered as a robust companion diagnostic tool when used in a multicenter setting.

There was a critical need to overcome this handicap by reducing (harmonization) or eradicating (standardization) variability (3). Major work was made to propose harmonized or standardized procedures for panel design, antibodies and fluorochromes choice, sample preparation, instrument settings, data analysis, and interpretation of the results (2,4–14).

The EuroFlow consortium started in 2006 and aimed to support the development and standardization of fast, accurate, and highly sensitive approaches for diagnosis and classification of hematological malignancies, as well as for the evaluation of treatment effectiveness (http://www.euroflow.org/). The results of their work were published 6 years later, with two major articles regarding the standardization of antibody panels and FC protocols (15,16).

The FranceFlow group was created in October 2010 to share the knowledge resulting from the EuroFlow project with multiple laboratories involved in the immunophenotyping of hematological diseases in France and Belgium. It was initially composed of 15 participating centers and progressively expanded to 32 centers to the date of writing. FranceFlow had two primary objectives: (1) train the participating centers to standardize instrument settings along with the EuroFlow recommendations and (2) evaluate the standardization with several rounds of quality controls (QCs).

To reach these aims, regular meetings were organized and the QCs were set up. To date, the group met 13 times from October 28, 2009 to November 11, 2017. Six rounds (QC1 to QC6) were performed on normal fresh or stabilized peripheral blood samples and stained using a EuroFlow QC tube to compare the FranceFlow results to the EuroFlow data reported by Kalina et al (17). A seventh QC (QC7) was performed on a pathological blood sample and according to a previously validated lymphoid screening tube broadly used among participants (FranceFlow LST, modified from Ref. (16)) to better reflect routine clinical practice. Furthermore, in order to independently measure the benefit from standardization and the impact of reagent variation across centers, the centrally distributed sample was analyzed using both centrally distributed and local reagents.

In this study, we report the FranceFlow experience and achievements over the period of 2009–2015 (Fig. 1A). Our data show that the recommendations from the EuroFlow consortium can be successfully implemented into the practice of numerous clinical laboratories, allowing for standardization of FC results.

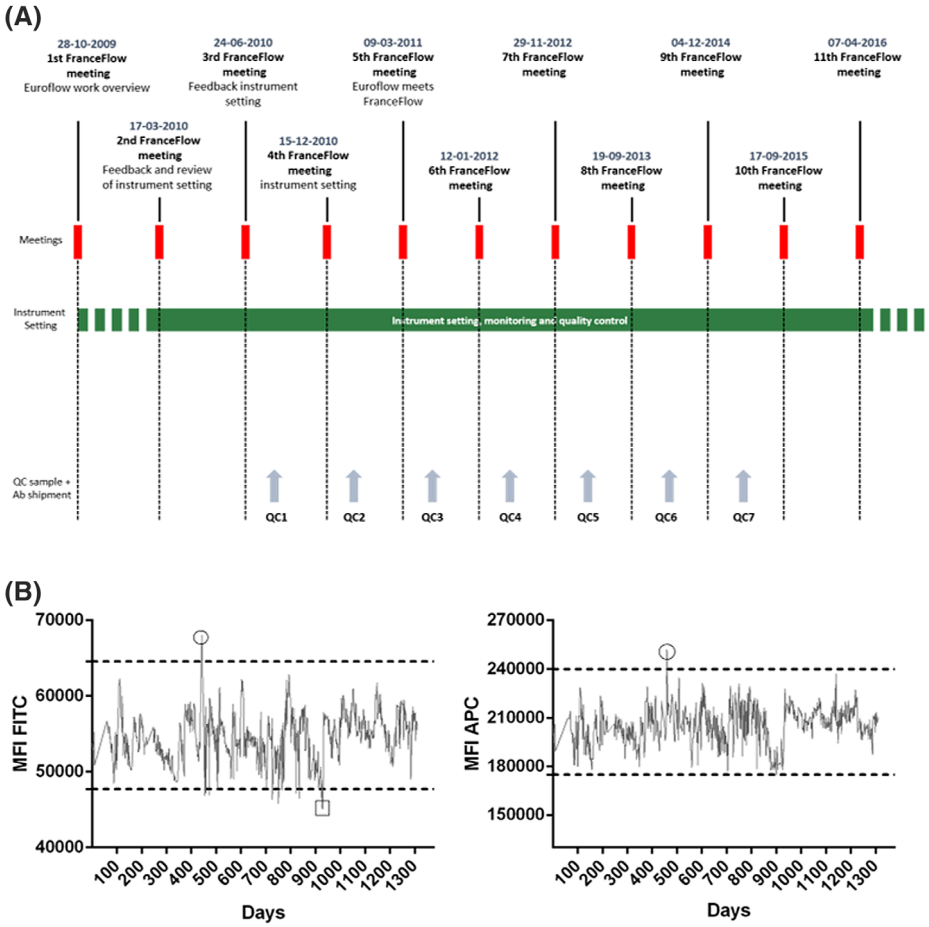

Figure 1. (**A**) Overview of the FranceFlow program from 2009 to 2017. Only results from QC1 to QC7 are described in this article. QC: quality control, Ab: antibody. (**B**) Overtime stability of rainbow beads seventh/eighth peak MFI profile obtained in a single center (n°18) for FITC and APC fluorescence channels, upon 3.5 years (1,300 days) monitoring. Dotted lines represent the acceptable ±15% range around the target MFI. Circles: Examples of one-time values above or below 15% of the target MFI. Square: Out of range MFI values reflecting decreased laser performance and requiring service visit and new setup. MFI: mean fluorescence intensity, FITC: fluorescein isothiocyanate, APC: allophycocyanin. [Color figure can be viewed at wileyonlinelibrary.com]

Table 1. Description of QC rounds

QC ROUND	QC1	QC2	QC3	QC4	QC5	QC6	QC7
Samples (number of)[1]							
Fresh normal peripheral blood	2	2	2	2	1		
Stabilized normal blood					1	1	
Fresh pathological peripheral blood							1
Antibodies							
Distributed by the coordinating center	×	×	×	×	×		× (lyophilized tube[2])
Local						×	× (FranceFlow LST)
Combination[3]							
QC (Table S2A) CD8/CD27/CD4/CD19/CD14/CD3/CD20/CD45	×	×	×	×	×	×	
Lyophilized tube (Table S2B) CD8-lgK/CD56-lgL/CD5/CD19/CD10/CD3/ CD4-CD20/CD45							×
FranceFlow LST (Table S2C) CD8-lgK/CD56-lgλ/CD5/CD19/CD10/CD3/ CD4-CD20/CD45							×

[1] All samples were distributed by the coordinating center.
[2] Except from lgk et lgλ polyclonal antibodies.
[3] Details in the Supporting Information.

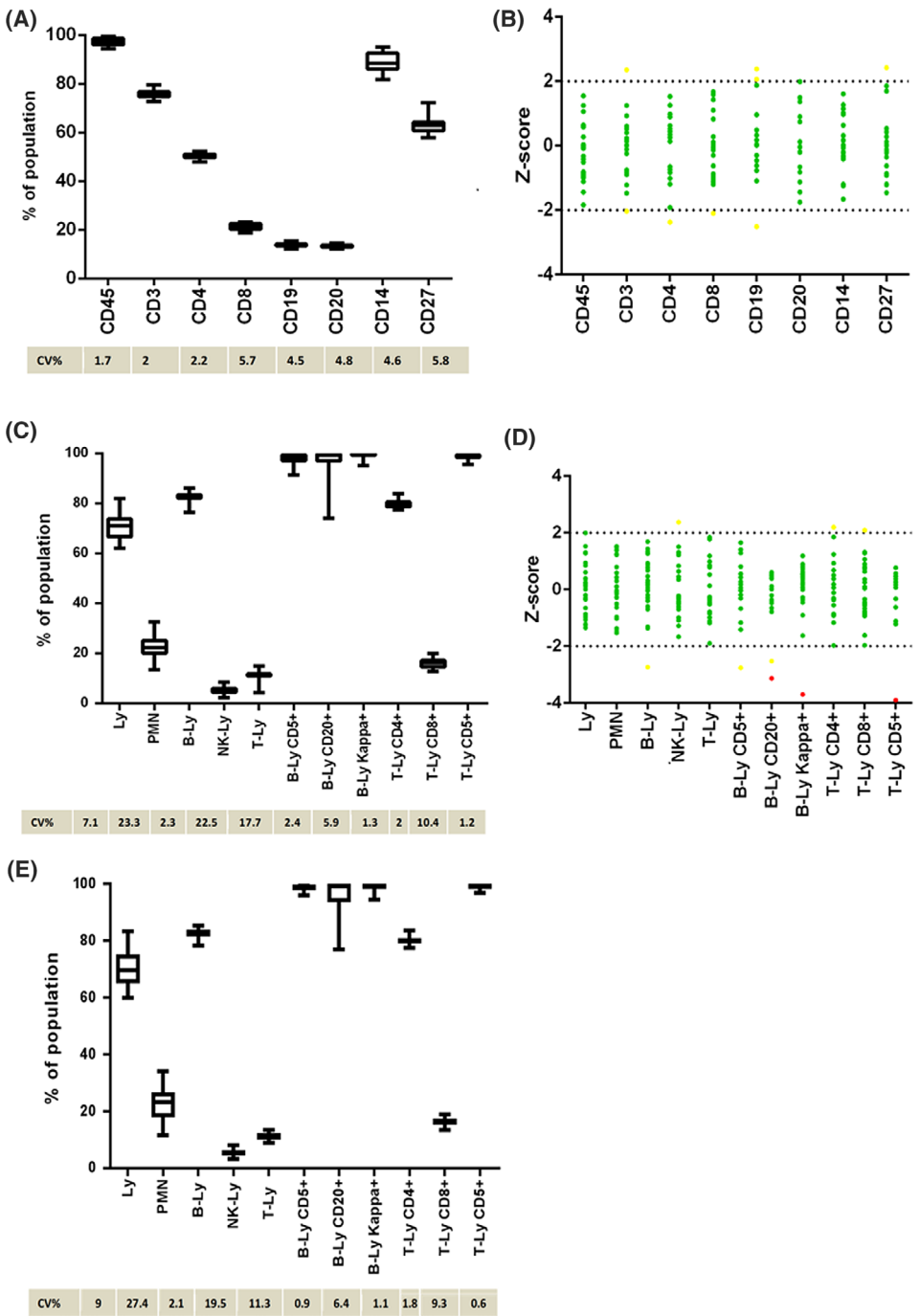

Figure 2. Intercenter comparison of cell subsets percentage measurement. (**A**) Box plot representation of the percentages of cell subsets measured in the Multicheck stabilized blood sample by the 16 centers participating to QC6, representing 22 instruments including 20 BD FACSCantos and 2 Navios. Cell subsets were gated as described in the Material and Methods section: % of CD45 among FSC/SSC lymphocytes, % of CD3, CD19 or CD20 positive cells among FSC/SSC/CD45+ lymphocytes, % of CD4, CD8 or CD27 positive cells among FSC/SSC/CD45+/CD3+ lymphocytes and % of CD14 positive cells among FSC/SSC/CD45+ monocytes. CVs (%) between the centers are indicated below the Box plots. **(B)** Z-scores obtained by each instrument for each indicated cell subset during QC6. Agreement interval is [−2;+2]. Z-score values are represented with a color code: Green, yellow and red for values comprised between −2 and + 2, −3 and − 2 or + 2 and + 3, <−3 or > +3, respectively. A total of 8 of the 176 (4.5%) z-score values were out of the [−2;+2] range: Among 22 participating instruments, they concerned center n°16 and cytometer B of center n°17 for the % of CD3+ cells, center n°5 for the % of CD4+ cells and the % of CD19+ cells, center n°7 for the % of CD8+ and CD19+ cells, center n°15 for the % of CD19+ cells, and the cytometer B of center n°8 for the % of CD27+ cells. (**C**, **E**) box plot representation of the percentages of cell subsets measured in a CLL blood sample by the 19 centers participating to QC7, representing 23 instruments including 22 BD FACSCantos and 1 Navios. Samples were stained with a local FranceFlow LST combination (C) or with a centrally distributed lyophilized tube (E). Cell subsets were gated as described in the Material and Methods section: FSC/SSC/CD45+/single lymphocytes and FSC/SSC/CD45+/CD10+/single neutrophils (PMN) among CD45/

Material and Methods

Participating Centers

Fifteen centers were initially involved in the FranceFlow group, 14 in France and one in Belgium. Most of them were university hospitals. The size of the group reached 19 centers at the time of the last QC reported here, in 2015 (Supporting Information Table S1). Since then, the group has now reached 32 centers to date of writing and is still expanding. In France, these centers represent around 30% of the laboratories using immunophenotyping for the diagnostic and follow-up of hematological malignancies.

Instruments

Supporting Information Table S1 summarizes the number and repartition of the flow cytometers used in this study. BD FACSCanto II (BD Biosciences, San Jose, CA, USA) equipped with three lasers emitting at 405, 488, and 633 nm. Optical filter configurations were identical. Acquisition and analysis were performed using the Diva software (BD Biosciences). Infinicyt (Cytognos, Salamanca, Spain) was used to represent QC results. Two centers were equipped with Navios and a 3-laser 10-color configuration (Beckman Coulter, Miami, FL). They used the CXP software to analyze and report the data.

Reagents

Eight-peak Rainbow bead calibration particles (Spherotech, Lake Forest, IL or BD Biosciences) were used for the instrument setup as previously described (15). In order to reach optimal standardization, the same batches of Rainbow beads were used over all the centers.

The antibodies recommended for establishing compensation settings are detailed in the Supporting Information. Antibody combinations used for the QC tube, the custom lyophilized tube, and the FranceFlow LST combination are described in Supporting Information Table S2.

Antibodies were centrally distributed for QCs 1–5, as well as lyophilized tubes for the QC7 except for Igκ and Igλ polyclonal antibodies (Dako, Les Ulis, France). These custom lyophilized tubes were designed and used for standardization. Local antibodies were also used during the QC6 for the QC combination (same clone and company, various lots) and the QC7 for the FranceFlow LST combination to evaluate the impact of using various clones, lots, and/or volumes on standardization.

Biological Samples

Normal peripheral blood was used for the setup. For the QC1 to the QC4, fresh peripheral blood samples from two healthy donors were centrally distributed. Of the two samples issued for the QC5, one was fresh peripheral blood (QC5a) and the second was a commercial stabilized normal blood (QC5b, Multicheck, reference 340912). For the QC6, only the Multicheck-stabilized blood was used (lot BM113N). The QC7 was realized on a fresh pathological blood sample collected from a patient with chronic lymphoid leukemia (CLL) and distributed by the coordinating center (Table 1). Informed consent was obtained from patients and healthy volunteers following the declaration of Helsinki. This study was performed according to our institution guidelines.

Staining Procedures

See Supporting Information.

Instrument Settings

Instrument setup was performed in each center according to the fundamental principles of the EuroFlow standard operating protocol (SOP) (Kalina et al., section 2),(15) with minor adaptations (see Supporting Information).

Standardization Monitoring

Instrument performance was monitored on a daily or every other day basis in each center according to EuroFlow recommendations (15). Whenever the instrument performance check failed, cleaning, de-gassing the flow cell, and laser delay verification were performed. Whenever the performance was not restored, the manufacturer's intervention was required. Of note, daily start-up was performed according to the manufacturer's recommendations (BD Biosciences or Beckman Coulter) prior to the EuroFlow protocol.

Quality Controls

Seven external QCs were centrally distributed to each participant of the group between September 2010 and April 2015 (Table 1). Acquisition of these distributed samples was synchronized in all centers, meaning that these samples were stained only once they had been received by all group members to avoid biological variations related to sample age before being processed. A minimum of 50,000 total events was recorded.

Data collection was centralized for the QC1 to the QC5, whereas it was performed locally for the QC6 and the QC7 as described in the Supporting Information. Processing of the data was centrally performed for all the QCs.

Statistical Analysis

For the QC6 and the QC7, accuracy was assessed by the SDI (Z-score) calculated as follows: SDI = (m-M)/SD, where m is the result of the instrument of each cell subset or of each median fluorescence intensity (MFI) measurement, M is the mean of the peer group, and SD is the standard deviation of the peer group.

SSC leucocytes, CD19+ CD3− B lymphocytes, CD19- CD3- CD5- CD56+ NK cells and CD19− CD3+ T lymphocytes among FSC/SSC/CD45/single total lymphocytes, B lymphocytes positive for CD5, CD20 or Igκ light chain among total B lymphocytes, and finally T lymphocytes expressing CD4 + CD8−, CD4-CD8+ or CD5+. CVs for Igλ + B lymphocytes, CD56+ and CD10+ T lymphocytes were not interpreted because these cell subsets were barely present. CVs (%) between the centers are indicated below the Box plot. (**D**) Z-scores obtained by each instrument for each indicated cell subset during QC7. Lyophilized tubes were distributed to each center for staining. For a total of 23 participating instruments and 11 cellular subsets evaluated, 9/253 (3.6%) Z-score values were out of the [−2;+2] range. [Color figure can be viewed at wileyonlinelibrary.com]

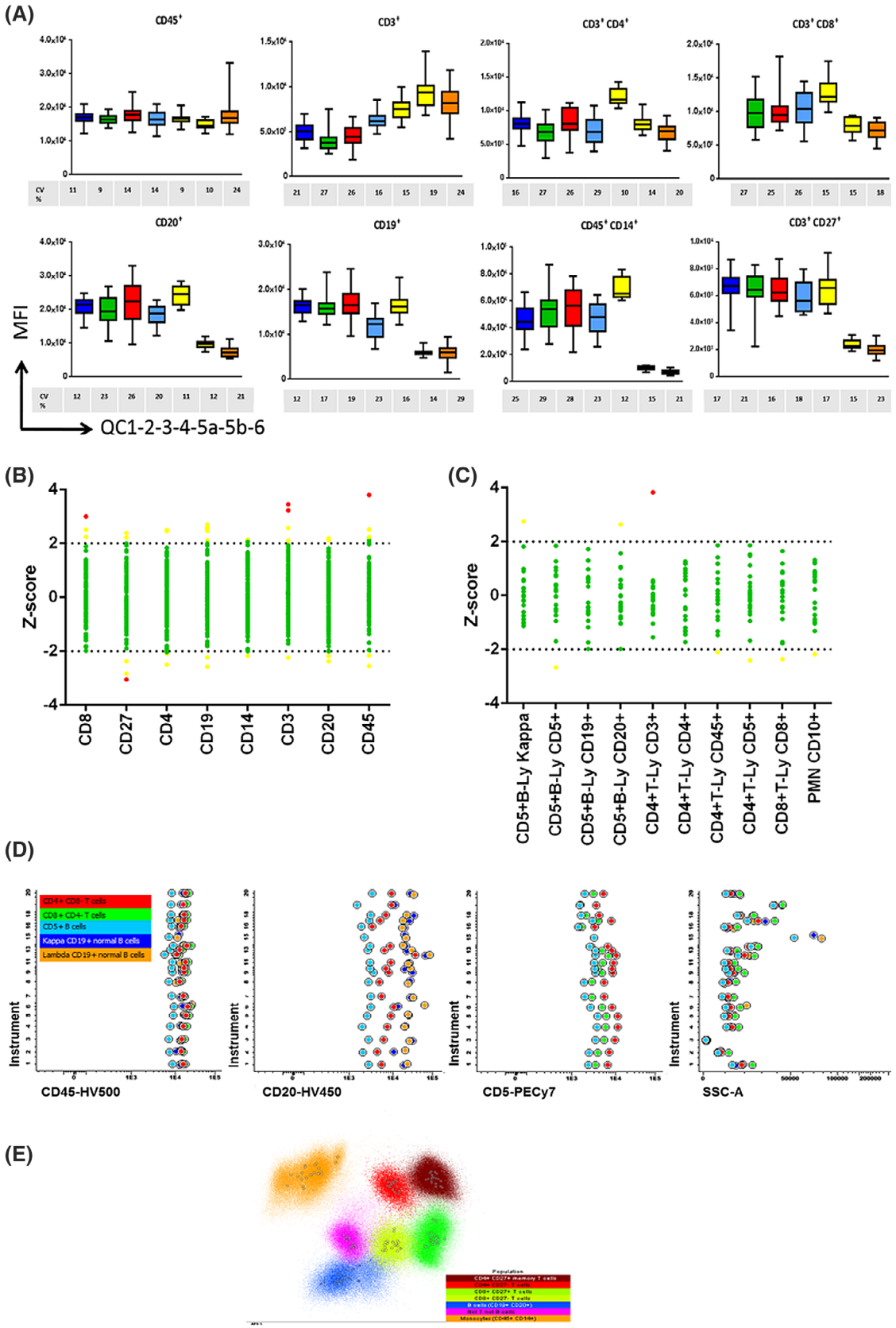

Figure 3. Legend on next page.

The SDI indicates how many SDs a result is different from the mean of all participating centers results. Z-score values within the range of [−2;+2]recapitulate the results found by the other participating centers. The [−3;−2] or [+2;+3] intervals represent a narrow divergence from the participating centers. Values under −3 or above +3 reflect outliers.

MFI span was calculated as subtraction of logarithmically transformed fifth percentile MFI from 95th percentile MFI to illustrate the impact of variability on the graphical interpretation of the data (15).

Results

Instrument Setting and Monitoring of Instrument Performances

Standardization of flow cytometer instrument settings aimed to improve intra- and inter-center reproducibility of FC results over time. Two practical sessions were organized to train participants for the SOP designed by EuroFlow (15). Daily monitoring of the instruments' performance according to the SOP was easily adopted in all the centers and enabled the evaluation of the level of standardization.

We tested the intra-center reproducibility of Rainbow beads MFI values. These Rainbow beads emit a given and stable level of fluorescence. Each participant measured Rainbow beads MFI values on a daily basis (or every other day) for a period of 12–27 months. The CV representing the overall intra-center (or instrument) variation of the fluorescence measured was ≤10%, apart from one center for the PE-Cy7 channel (Supporting Information Table S3). As illustrated in Figure 1B, FITC and APC channels, MFI values for the brightest Rainbow beads peak mostly remained within the ±15% range as permitted by the SOP over a follow-up of more than 4 years (1,300 days). On rare occasions, MFI values above or below 15% of the target MFI as a result of minor instrument failures were resolved by simple measures (clean, de-gas of flow cell) (Fig. 1B, circles). In contrast, a continuous decrease of MFI, possibly leading to out of range MFI values, reflected alterations of the laser performance and were corrected after a requested service visit (Fig. 1B, squares). The FranceFlow protocol appeared to be very sensitive to instrument deviations. Therefore, we were more stringent than the manufacturer's recommendations. Consequently, less technical support and new setups were required. Compensation setup (including voltages) had to be repeated 11 times on average within a period of 4 years over all the centers (range 3–20), that is, 2.8 times a year, and 13 service visits were requested (range 6–27).

Inter-center reproducibility of Rainbow beads MFI values enabled participants to work with comparable signals over space. This was evaluated by daily measurements of the eighth Rainbow beads peak MFIs (seventh for violet laser channels) over a year or more, which represents a total of 5,399 measurements per channel for all instruments. The CV of the MFI values did not exceed 7% between all 12 centers and 13 instruments (Supporting Information Table S3).

Inter-Center Reproducibility of Cell Subsets Percentages and MFI Values Measurements

Inter-center reproducibility was evaluated on both (normal) stabilized and (normal or pathological) fresh blood samples (Table 1).

Considering the measurement of cell subset percentages, sample variability was reduced by using a stabilized blood, with excellent reproducibility with CVs <6% between the 22 cytometers from the 16 participating centers to the QC6. The highest CVs (4–6%) were observed for CD27, CD8, CD20, CD14, and CD19-positive subsets, while CVs below 2.2% were observed for CD45, CD3, and CD4-positive subsets (Fig. 2A). CD45+ lymphocytes, CD20+ lymphocytes, and CD14+ monocyte percentages had Z-scores within the range of [−2;+2]. For the other subsets, a maximum of 3 of the 22 centers had Z-scores between ([−2;−3] and [+2;+3]) highlighting a narrow divergence. No center had a Z-score in the critical zone below −3 or upper +3. In total, 4.5% (8/176) of the Z-score values were considered slightly divergent for this QC (Fig. 2B). Of note, different lot numbers of the same clones of antibodies were used in the different centers with no specific instructions for gating strategy. Despite this, no significant differences in the calculated CVs and Z-scores were observed (data not shown).

To better reflect daily practice, we ran the QC7 using a fresh pathological blood sample (CLL) and the lymphoid screening antibody panel commonly used by the participating

Figure 3. Assessment of interlaboratory mean fluorescence intensity (MFI) standardization. **(A)** Box plot representations of MFI values observed overtime (QC1 to QC6) for eight antigens evaluated in subsets of peripheral lymphocytes and monocytes gated as described in the Material and Methods section. Data were obtained from two distributed healthy donor samples (except QC5a: 1 donor) in 13,13,14, 10, and 16 centers for QC1 to QC5a respectively, and from a stabilized blood Multicheck sample in 17 and 15 centers (21 instruments) for QC5b and QC6. *N* = 26, 26, 28, 20, 16, 17, and 21, respectively. CD8 was not retained for analysis of QC1 because the clone used did not work well in the QC tube. **(B)** Z-scores obtained for MFI values on each tested sample in all participating centers from QC1 to QC6. Green: Z-scores in the [−2;+2] concordance interval; yellow: Z-scores in the [−3;-2] or [+2;+3] alarm zone; red: Z-scores in the <−3 or > +3 discordance zone. For example, for CD27 MFI and from QC1 to QC6, 148/154 samples tested in all participating centers received a Z-score comprised between −2 and + 2, 5/154 a [−3;-2] or [+2;+3] Z-score and 1/154 a < −3 or > 3 Z-score. **(C)** Z-scores obtained for MFI values during QC7. A sample from a patient with chronic lymphocytic leukemia (CLL) was distributed, stained in 18 centers (22 instruments) using a distributed lyophilized tube and analyzed locally. Cellular subsets were gated as described in Material and Methods. The number of samples receiving the corresponding Z-score is indicated for each marker. Overall 8 of the 220 (3.6%) Z-score values were outside the [−2;+2] interval. **(D)** Illustration on a logarithmic scale of CD45, CD20 and CD5 MFI and SSC values obtained by each instrument for several cellular subsets during QC7. The same CLL sample was stained in 17 centers (20 instruments) using their local lymphoid screening tube (FranceFlow LST) and analyzed locally. CVs between the instruments on CD5+ clonal B-cells (light blue) were 11% for CD45, 20% for CD20 and 19% for CD5. For comparison, the non-standardized SSC parameter was also represented. **(E)** Assessment of interlaboratory standardization at the time of QC4 using APS view. The selected subpopulations are color-coded as listed. Each circle represents the median value of a gated subset for one sample and center. Two normal blood samples were distributed to 11 centers. APS1 view was configured with all parameters except FSC/SSC. [Color figure can be viewed at wileyonlinelibrary.com]

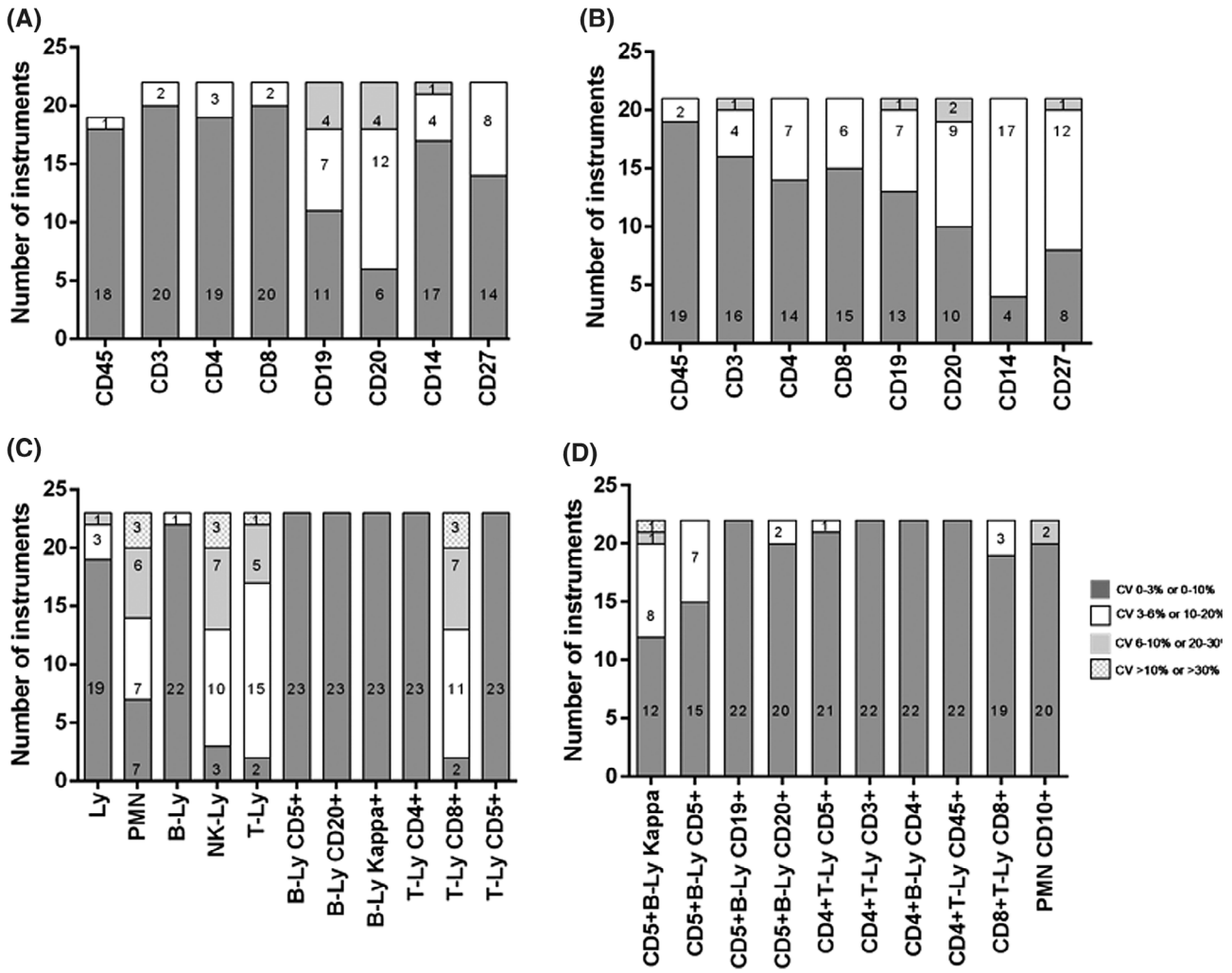

Figure 4. Assessment of intralaboratory standardization. **(A, B)** CVs of cell subsets for percentage (A) and MFI (B) obtained on a Multicheck stabilized blood stained with local antibodies and analyzed 15 times in 16 (A) or 15 (B) participating centers, representing a total of 22 (A) or 21 (B) instruments (QC6). **(C, D)** CVs obtained on a CLL blood sample stained using five lyophilized tubes and analyzed in 19 (C) or 18 (D) participating centers, representing a total of 23 (C) or 22 (D) instruments (QC7). For cell subsets percentage measurement, CVs comprised between 0–3%, 3–6%, 6–10%, and > 10% are represented in dark gray, white, light gray and grid pattern, respectively (A and C). For MFI measurement, CVs comprised between 0–10%, 10–20%, 20–30%, and > 30% are represented in dark gray, white, light gray and grid pattern, respectively (B and D). The number of instruments is indicated for each marker.

laboratories (FranceFlow LST). Using lyophilized tubes to decrease technical variation related to reagents, CVs were higher across the 23 cytometers of the 19 participating centers compared to those obtained during the QC6 with a stabilized sample and the QC combination, up to 20% for lymphoid subpopulations and 27% for CD10+ neutrophils (PMN).

Nevertheless, clinically relevant subsets for CLL analysis showed CVs below 6.4% (Fig. 2E). All cytometers (23/23) obtained a Z-score within the range of [−2;+2] for 3 of the 11 cell subsets (FSC/SSC/CD45+/single lymphocytes, FSC/SSC/CD45+/CD10+/single neutrophils, and CD19− CD3+ lymphocytes). All but one cytometer (22/23) generated results within the same range for 7 of the 11 cellular subsets (CD19+ CD3− lymphocytes, B lymphocytes positive for CD20, Igκ light chain or CD5, CD19− CD3− CD5− CD56+ cells and T lymphocytes expressing CD4+CD8− or CD4-CD8+). Twenty-one out of 23 cytometers had Z-scores within this range for the CD20+ B-lymphocyte subset. Overall, at least 21 of the 23 instruments achieved Z-scores within the range of [−2;+2] for all subsets and only 3.6% (9/253) of all Z-score values were outside the range [−2;+2] (Fig. 2D). Using local reagents, CVs were in the same range (Fig. 2C). Only 5% (14/276) of z-score values were outside the range [−2;+2] (data not shown).

Concerning MFI measurements of normal fresh or stabilized blood samples, CVs across the participating centers were <30% (Fig. 3A). Only 3.5% (41/1185) of z-score values from all markers, centers and the QCs (QC1 to QC6) were outside the [−2;+2] interval, with no significant difference between the markers (Fig. 3B). The fluctuation of the MFI around the median, called MFI span (10), spanned <0.3 logscale for all the QC markers (CD20, CD4, CD45, CD8, CD19, CD3, CD27, and CD14) for normal fresh (QC1 to 5)

and stabilized (QC5 and QC6) blood samples. This demonstrates the relative stability of MFI values over QC rounds from September 2010 to February 2013 on different normal blood samples (data not shown).

The QC6 was run using local reagents in contrast to the QC1-5. The CVs were also <30% between the 21 cytometers of all 15 participating centers, including 2 Navios, and the highest CV (29%) was observed for CD19 positive subsets (Fig. 3A, orange). Z-scores were all within the range of [−2;+2] except for two cytometers and three markers, representing 2.4% (4/168) of Z-score values outside the [−2;+2] interval (data not shown).

In the QC7, lyophilized tubes were first used by all 18 participating centers (22 instruments) to reduce reagent-related variation. MFI CVs for positive markers in all channels except PE (not evaluable) ranged between 9% (CD45 on CD4+ T cells) and 33% (CD10 on PMN) for monoclonal antibodies. These CVs were < 20% for all the relevant CLL markers (except for the polyclonal antibody kappa).

Z-scores were in the range of [−2;+2] for all the centers for CD19 on B cells and CD4 on T-cells, and for all but one center for 8 of the 10 markers. Only 3.6% (8/220) of z-score values were found outside the [−2;+2] interval (Fig. 3C). To reflect daily practice, we subsequently stained the same sample using local reagents. These CVs were slightly superior but remained below 30% for CLL markers, except for the polyclonal anti-kappa marker, which had a CV of 67% (Fig. 3D).

Z-scores were within the range of [−2;+2] for all the centers for CD19 on B cells and CD5 on T-cells, for 17/18 centers for 6/9 markers and for 14/18 cytometers for CD8 on T-cells. Only 5.6% (10/180) of Z-score values were outside the [−2;+2] interval (data not shown). In contrast to the standardized parameters, non-standardized parameters such as the Side Scatter gave very variable results between the centers (Fig. 3D).

Finally, the automated population separator (APS) was able to clearly distinguish clusters of lymphocyte subsets and monocytes in the two normal blood samples distributed for the QC4, suggesting that the technical variability between the instruments and the sample variability were minor compared to the biological differences between these cellular subsets revealed by the antibodies used (Fig. 3E).

Table 2. CVs of MFI values obtained after standardization

		CD20	CD45	CD8	CD27	CD4	CD19	CD14	CD3	N	N1	N2	COMBINATION
EF, Kalina,	Stabilized[1]	15,2[2]	13,95[3]	11,4[4]	32,9[5]	24,7[6]	11,1[7]	43,8[8]	38,7[9]	8	8	1	QC
Leukemia 2012	Fresh[10]	16,9[2]	15,55[3]	16,9[4]	28[5]	28,4[6]	15,4[7]	22,7[8]	48,4[9]	30	8	3–4	QC
EF, Kalina,	2010[10]	32,6[2]	28,85[3]	15,5[4]		19,3[6]	32,4[7]		29,5[8]	30	10	3	LST
Cytometry,	2011[10]	25,8[2]	37,35[3]	14,3[4]		15,3[6]	24,5[7]		38,5[8]	33	11	3	LST
2015	2012[10]	25,6[2]	19,15[3]	39,2[4]		16,9[6]	26,6[7]		40,9[8]	33	11	3	LST
	2013[10]	29,3[2]	28,45[3]	31,2[4]		18,6[6]	14[7]		22,5[8]	15	5	3	LST
	2013 Mix[1,10]	21 6[2]	11,55[3]	23,2[4]		10,9[6]	16[7]		35,7[8]	27	9	3	lyo LST
	MFI span	0,42[2]	0,315[3]	0,43[4]		0,27[6]	0,51[7]		0,54[8]	123			
FF, Solly, 2019	QC1[11]	12[12]	11[13]		17[5]	16[6]	12[7]	25[8]	21[9]	26	13	2	QC
	QC2[11]	23[12]	9[13]	27[4]	21[5]	27[6]	17[7]	29[8]	27[9]	26	13	2	QC
	QC3[11]	26[12]	14[13]	25[4]	16[5]	26[6]	19[7]	28[8]	26[9]	28	14	2	QC
	QC4[11]	20[12]	14[13]	26[4]	18[5]	29[6]	23[7]	23[8]	16[9]	20	10	2	QC
	QC5a[11]	11[12]	9[13]	15[4]	17[5]	10[6]	16[7]	12[8]	15[9]	16	16	1	QC
	MFI span	0,21[12]	0,12[13]	0,22[4]	0,14[5]	0,26[6]	0,20[7]	0,2[8]	0,23[9]	116			
	QC5b[1]	12[12]	10[3]	15[4]	15[5]	14[6]	14[7]	15[8]	19[9]	17	17	1	QC
	QC6[1]	21[12]	24[13]	18[4]	23[5]	20[6]	29[7]	21[8]	24[9]	21	21	1	QC
	QC7[14]	20,2[12]	14,6[13]	36[4]		28,5[12]	18,8[7]		21,7[9]	20	20	1	FF LST
	QC7[14]	16,5[12]	9[13]	33,7[4]		9,1[12]	14,2[7]		24,4[9]	22	22	1	Lyophilized tube

QC: quality control tube, LST and Lyo LST: cf Kalina et al, FF LST and lyophilized tube: cf Table 1; EF: Euroflow; FF: FranceFlow; N: number of samples; N1: number of instruments; N2: number of samples per instrument.
[1]Stabilized blood.
[2]Pacific Blue.
[3]Pacific Orange.
[4]FITC.
[5]PE.
[6]PerCP Cy 5.5.
[7]PECy7.
[8]APC.
[9]APC H7.
[10]Local, fresh normal peripheral blood.
[11]Distributed by the coordinating center, fresh normal peripheral blood.
[12]Horizon V450.
[13]Horizon V500.
[14]Distributed by the coordinating center, fresh pathological peripheral blood.

Intra-Center Reproducibility of Cell Subset Percentages and MFI Values Measurements

Intra-center reproducibility was evaluated during the QC6 using a stabilized Multicheck blood sample that was distributed by the coordinating center. The blood sample was stained, measured, and analyzed once per day for 15 days by 22 instruments (16 participating centers) (Supporting Information Fig. S3). During the QC7, a fresh CLL blood sample distributed by the coordinating center, as well as five lyophilized tubes were measured and analyzed by 23 instruments (19 centers).

Using the Multicheck stabilized blood, intra-center CVs of the percentage of cell subsets varied between 0.34 and 9.5% depending on cell subsets and centers. CD45+ and CD3+ lymphocytes had the lowest CVs and CD20 positive lymphocytes had the highest CVs in the majority of the centers (data not shown). CVs were <3% for the majority of cell subsets and centers. They ranged from 3 to 6% for CD20+ lymphocytes. For a couple of instruments, CVs ranged from 6 to 10% for CD19+ or CD20+ lymphocytes (n = 4/22) or CD14+ monocytes (n = 1/22) (Fig. 4A). Using a fresh CLL blood sample, CVs were <10% for the majority of cell subsets and centers. Importantly, they were <3% for all centers for major lymphoid subsets: CD5+, CD20+ or kappa+ B lymphocytes, CD3+, CD5+ or CD4 + CD8- T lymphocytes (Fig. 4C).

Intra-center CVs of MFI values varied between 1.87 and 22.63% depending on the markers and the centers. CD45 gave the lowest CV and CD20 and CD14 gave the highest CV in the majority of the centers (data not shown). CVs were <20% for the majority of markers and instruments. They were usually <10% for CD45 (19/21 instruments), CD3 (16/21), CD4 (14/21), CD8 (15/21), and CD19 (13/21). They were higher but almost always <20% for CD20 (19/21 instruments), CD14 (21/21), and CD27 (20/21) (Fig. 4B). Using a fresh CLL blood sample, CVs were <10% for the majority of markers and centers (Fig. 4D), except for lambda and CD56 which were expressed by a minority of cells in this sample (data not shown).

Discussion

Here, we report 5 years of experience within FranceFlow in which we aimed to translate the standardized instrument setting protocol designed by EuroFlow into clinical practice in France and Belgium. This was a two-step process based on (1) a training step with theoretical conferences and practical training and (2) a validation step in which the seven QCs were run to evaluate the impact of standardization. These steps were carried out in parallel so that the QC results could be reported and commented on during meetings to promote a positive feedback loop.

The design of the QCs was changed over time to adapt to the evolution within FranceFlow. First, the number of participants to the QC increased from 13 to 19 as the FranceFlow group expanded. Second, the original aim of the QC was to demonstrate the reproducibility of the results published by EuroFlow (15). As such, we used the tube configuration applied in EuroFlow to be able to compare our results. Once this was validated, we aimed to assess the effect of standardization in daily practice. To do so, we replaced the original QC tube with the FranceFlow LST tube adapted from the EuroFlow LST (16). By analyzing a centrally distributed biological sample, using both common and local reagents, we could show that the instrument procedure gave highly reproducible data across 23 instruments.

Supporting these statements, CVs were <30% in the QC1-6 for most of the MFI values on positive cells (Fig. 3). These results were reproducible both with fresh (the QC1 to the QC5) and stabilized (QC5 and QC6) normal blood samples, using central or local reagents (QC6) and when analysis was centralized (QC1 to 5) or local (QC6). EuroFlow data from eight laboratories using a similar QC protocol resulted in CVs of the MFI values of <44% (<17% for 4/8 markers) (15) (Table 2). From the QC1 to the QC6, 96.5% of the FranceFlow centers obtained Z-scores within the range [−2; +2], suggesting that outliers were most likely related to mistakes in instrument settings in a few centers rather than real instrument performance variation. Finally, testing of a pathological (CLL) sample stained with the FranceFlow LST also gave CVs of <30% for relevant markers (QC7). They were slightly higher for CD8 on T cells and CD10 on PMN, for light chains (Igκ polyclonal antibodies) as already reported (17,18), and for rare subsets which were excluded from the analysis (Igλ, CD56 and C10 on T cells) (18). Once again, these results are consistent with those obtained with normal blood samples during four rounds of QCs involving 9–11 EuroFlow laboratories between 2010 and 2013 (17) (Table 2). Of note, in contrast to the French group QC program, samples or reagents were not centrally distributed in the Euroflow quality assessment program, except for a lyophilized mix during the 2013 QC (17). Although the utilization of lyophilized tubes did not significantly affect CV ranges and Z-scores, they tended to perform slightly better than local reagents. Interestingly, 10 Swiss laboratories, including 6 BD FACSCanto II and 4 Navios, also reported an average CV of 30% between the centers for MFI values in normal blood lymphocyte subsets, using a lyophilized FranceFlow LST antibody mixture (18). Importantly, ≥95% of the FranceFlow centers participating in the QC7 obtained z-scores within the acceptable range. To the best of our knowledge, no such QC on a pathological sample has been previously reported in the literature. Beyond the real MFI values, interpretation of clinical flow data mainly relies on the measurement of percentages of the expression of individual markers. As expected, we also observed constant CVs of percentages of relevant populations of <7% across all the QCs. In keeping with this, Z-scores were within the acceptable range for over 95% of the centers if not all regardless of the reagents used and the stabilized/fresh or the normal/pathological type of sample considered. Altogether, our data demonstrate that the translation of the EuroFlow SOP into practice within the FranceFlow group was successful with comparable efficiency. Interestingly, when using the FranceFlow LST, which is used in general practice, the EuroFlow protocol produced highly homogeneous data.

The aforementioned inter-center reproducibility further extends to intra-center reproducibility with CVs <3% for percentages of cell subsets and < 20% for the MFI of most

markers. This demonstrates that data acquisition is very stable within a center in terms of MFI over time. Altogether, the 19 FranceFlow laboratories can be considered as one unique laboratory comprising of 23 identical instruments that are stable over time. Practically, this is of high relevance to minimal residual disease (MRD) analysis, which relies on the interpretation of signal intensities that require high stability over time between the analysis of the diagnosis and the follow-ups (12).

Likewise, some flow diagnoses rely again on MFI measurement: for example, the Matutes scoring system for CLL diagnosis (19). A variability of 30% appears barely visible on a log scale so that analysis of the same sample would lead to the same data and interpretation wherever it is analyzed. Therefore, such level of standardization significantly enhances the robustness of clinical flow. One would expect that such protocol could reshape the clinical FC practice and facilitate its use in clinical trials as a companion diagnostic test. As an example, the measurement of CD20 expression on B-cell precursor acute lymphoblastic leukemia (BCP-ALL) is not standardized. Using our platform associated with a unified clone, fluorochrome and titer of CD20 antibody would enable a robust assessment of CD20 expression in BCP-ALL.

Altogether, we demonstrate herein the broad applicability of the EuroFlow SOP for instrument setup. The protocol was successfully implemented in 23 instruments, including Navios machines, proving a tremendous cost-to-benefit ratio in terms of time spent to set up the instrument and resulting performances. The protocol could considerably improve the reproducibility of flow in hematology, opening new avenues in both clinical diagnostics and translational medicine. This will be achieved by performing regular QCs to monitor the standardization and organizing regular meetings to maintain competence in the field. This is even more pertinent as the FranceFlow group is currently expanding to 32 centers in French-speaking European countries (Belgium, Switzerland, France) which will require constant training/monitoring; therefore, the beneficial aspects of flow standardization will be far-reaching.

Acknowledgments

We acknowledge Charlotte Smith and Guillaume Andrieu for their support in editing this article.

Literature Cited

1. Craig FE, Foon KA. Flow cytometric immunophenotyping for hematologic neoplasms. Blood 2008;111(8):3941–3967.
2. Preijers FW, Huys E, Favre C, Moshaver B. Establishment of harmonization in immunophenotyping: A comparative study of a standardized one-tube lymphocyte-screening panel. Cytometry B Clin Cytom 2014;86(6):418–425.
3. Fuertes I. Towards harmonization or standardization in governmental accounting? The international public sector accounting standards board experience. J Comp Policy Anal 2008;10(4):327–345.
4. Bene MC, Nebe T, Bettelheim P, Buldini B, Bumbea H, Kern W, Lacombe F, Lemez P, Marinov I, Matutes E, et al. Immunophenotyping of acute leukemia and lymphoproliferative disorders: aconsensus proposal of the European Leukemia Net Work Package 10. Leukemia 2011;25(4):567–574.
5. Costa ES, Pedreira CE, Barrena S, Lecrevisse Q, Flores J, Quijano S, Almeida J, del Carmen Garcia-Macias M, Bottcher S, Van Dongen JJ, et al. Automated pattern-guided principal component analysis vsexpert-based immunophenotypic classification of B-cell chronic lymphoproliferative disorders: a step forward in the standardization of clinical immunophenotyping. Leukemia 2010;24(11):1927–1933.
6. Feller N, van der Velden VH, Brooimans RA, Boeckx N, Preijers F, Kelder A, de Greef I, Westra G, Te Marvelde JG, Aerts P, et al. Defining consensus leukemia-associated immunophenotypes for detection of minimal residual disease in acute myeloid leukemia in a multicenter setting. Blood Cancer J 2013;3:e129.
7. Flores-Montero J, Sanoja-Flores L, Paiva B, Puig N, García-Sánchez O, Böttcher S, van der Velden VHJ, Pérez-Morán JJ, Vidriales MB, García-Sanz R, et al. Next generation flow for highly sensitive and standardized detection of minimal residual disease in multiple myeloma. Leukemia 2017;31(10):2094–2103.
8. Johansson U, Bloxham D, Couzens S, Jesson J, Morilla R, Erber W, Macey M, British Committee for Standards in Haematology. Guidelines on the use of multicolour flow cytometry in the diagnosis of haematological neoplasms. British Committee for Standards in haematology. Br J Haematol 2014;165(4):455–488.
9. Novakova M, Glier H, Brdickova N, Vlkova M, Santos AH, Lima M, Roussel M, Flores-Montero J, Szczepanski T, Bottcher S, et al. How to make usage of the standardized EuroFlow 8-color protocols possible for instruments of different manufacturers. J Immunol Methods 2017; pii: S0022-1759(17)30139-4.
10. Preijers FW, van der Velden VH, Preijers T, Brooimans RA, Marijt E, Homburg C, van Montfort K, Gratama JW. Fifteen years of external quality assessment in leukemia/lymphoma immunophenotyping in The Netherlands and Belgium: A way forward. Cytometry B Clin Cytom 2016;90(3):267–278.
11. Solly F, Rigollet L, Baseggio L, Guy J, Borgeot J, Guérin E, Debliquis A, Drenou B, Campos L, Lacombe F, et al. Comparable flow cytometry data can be obtained with two types of instruments, Canto II, and Navios. A GEIL study. Cytometry A 2013;83(12):1066–1072.
12. Theunissen P, Mejstrikova E, Sedek L, van der Sluijs-Gelling A, Gaipa G, Bartels M, Sobral da Costa E, Kotrová M, Novakova M, Sonneveld E, et al. Standardized flow cytometry for highly sensitive MRD measurements in B-cell acute lymphoblastic leukemia. Blood 2017;129(3):347–357.
13. Westers TM, Ireland R, Kern W, Alhan C, Balleisen JS, Bettelheim P, Burbury K, Cullen M, Cutler JA, Della Porta MG, et al. Standardization of flow cytometry in myelodysplastic syndromes: A report from an international consortium and the European LeukemiaNet working group. Leukemia 2012;26(7):1730–1741.
14. Wood BL, Arroz M, Barnett D, DiGiuseppe J, Greig B, Kussick SJ, Oldaker T, Shenkin M, Stone E, Wallace P. 2006 Bethesda International Consensus recommendations on the immunophenotypic analysis of hematolymphoid neoplasia by flow cytometry: optimal reagents and reporting for the flow cytometric diagnosis of hematopoietic neoplasia. Cytometry B ClinCytom 2007;72(Suppl 1):S14–S22.
15. Kalina T, Flores-Montero J, van der Velden VH, Martin-Ayuso M, Bottcher S, Ritgen M, Almeida J, Lhermitte L, Asnafi V, Mendonca A, et al. EuroFlow standardization of flow cytometer instrument settings and immunophenotyping protocols. Leukemia 2012;26(9):1986-2010.
16. van Dongen JJ, Lhermitte L, Bottcher S, Almeida J, van der Velden VH, Flores-Montero J, Rawstron A, Asnafi V, Lecrevisse Q, Lucio P, et al. EuroFlow antibody panels for standardized n-dimensionalflow cytometric immunophenotyping of normal, reactive and malignant leukocytes. Leukemia 2012;26(9):1908-1975.
17. Kalina T, Flores-Montero J, Lecrevisse Q, Pedreira CE, van der Velden VHJ, Novakova M, Mejstrikova E, Hrusak O, Böttcher S, Karsch D, et al. Quality assessment program for EuroFlow protocols: Summary results of four-year (2010-2013) quality assurance rounds. Cytometry A 2015;87(2):145–156.
18. Glier H, Heijnen I, Hauwel M, Dirks J, Quarroz S, Lehmann T, Rovo A, Arn K, Matthes T, Hogan C, et al. Standardization of 8-color flow cytometry across different flow cytometer instruments: A feasibility study in clinical laboratories in Switzerland. J Immunol Methods 2017; pii:S0022-1759(17)30205-3.
19. Moreau EJ, Matutes E, A'Hern RP, Morilla AM, Morilla RM, Owusu-Ankomah KA, Seon BK, Catovsky D. Improvement of the chronic lymphocytic leukemia scoring system with the monoclonal antibody SN8 (CD79b). Am J Clin Pathol 1997;108(4):378–382.