The EuroFlow PIDOT external quality assurance scheme: enhancing laboratory performance evaluation in immunophenotyping of rare lymphoid immunodeficiencies

Short title: The EuroFlow PIDOT-EQAS

Neirinck J.[1,2], Buysse M.[2], Brdickova N.[3], Perez-Andres M.[4,5], De Vriendt C.[6], Kerre T. [6], Haerynck F.[7], Bossuyt X.[8,9], van Dongen J.J.M.[4,5], Orfao A.[4,5], Hofmans M.[2], Bonroy C. [1,2], Kalina T.[3]* on behalf of the EuroFlow PID Consortium

1 Department of Diagnostic Sciences, Ghent University, Ghent, Belgium
2 Department of Laboratory Medicine, University Hospital Ghent, Ghent, Belgium
3 CLIP Cytometry, Department of Pediatric Hematology and Oncology, 2nd Faculty of Medicine, Charles University and University Hospital Motol, Prague, Czech Republic
4 Translational and Clinical Research Program, Centro de Investigación del Cáncer and Instituto de Biología Molecular y Celular del Cáncer, Consejo Superior de Investigaciones Científicas (CSIC)-University of Salamanca (USAL), Department of Medicine, IBSAL and CIBERONC, University of Salamanca, Salamanca, Spain
5 Cancer Research Centre (Instituto de Biologıa Molecular y Celular del Cancer (IBMCC), USAL-CSIC; CIBERONC CB16/12/00400), Institute for Biomedical Research of Salamanca (IBSAL), Department of Medicine and Cytometry Service (NUCLEUS Research Support Platform), University of Salamanca (USAL), Salamanca, Spain
6 Department of Haematology, University Hospital Ghent, Ghent, Belgium
7 Department of Pediatric Pulmonology and Immunology and PID Research Laboratory, University Hospital Ghent, Ghent, Belgium
8 Department of Laboratory Medicine, University Hospital Leuven, Leuven, Belgium
9 Department of Microbiology, Immunology and Transplantation, KU Leuven, Leuven, Belgium

Mattias Hofmans, Carolien Bonroy and Tomas Kalina share senior authorship.
*Corresponding author

Word Count: 4162 (abstract 248)
Number of tables: 3
Number of figures: 5
This manuscript contains supplementary material.

Key words: external quality assurance, flow cytometry, EuroFlow, primary immunodeficiencies, standardization

The EuroFlow PIDOT external quality assurance scheme: enhancing laboratory performance evaluation in immunophenotyping of rare lymphoid immunodeficiencies

ABSTRACT

Objectives: The development of External Quality Assessment Schemes (EQAS) for clinical flow cytometry (FCM) is challenging in the context of rare (immunological) diseases. Here, we introduce a novel EQAS monitoring the primary immunodeficiency Orientation Tube (PIDOT), developed by EuroFlow, in both a 'wet' and 'dry' format. This EQAS provides feedback on the quality of individual laboratories (i.e. accuracy, reproducibility and result interpretation), while eliminating the need for sample distribution.

Methods: In the wet format, marker staining intensities (MedFIs) within landmark cell populations in PIDOT analysis performed on locally collected healthy control (HC) samples, were compared to EQAS targets. In the dry format, participants analyzed centrally distributed PIDOT flow cytometry data (n=10).

Results: We report the results of six EQAS rounds across 20 laboratories in 11 countries. The wet format (212 HC samples) demonstrated consistent technical performance among laboratories (median %rCV on MedFIs = 34.5%; average failure rate 17.3%) and showed improvement upon repeated participation. The dry format demonstrated effective proficiency of participants in cell count enumeration (range %rCVs 3.1% - 7.1% for the major lymphoid subsets), and in identifying lymphoid abnormalities (79.3% alignment with reference).

Conclusions: The PIDOT-EQAS allows laboratories, adhering to the standardized EuroFlow approach, to monitor interlaboratory variations without the need for sample distribution, and provides them educational support to recognize rare clinically relevant immunophenotypic patterns of primary immunodeficiencies (PID). This EQAS contributes to quality improvement of PID diagnostics and can serve as an example for future flow cytometry EQAS in the context of rare diseases.

INTRODUCTION

Multi-parameter flow cytometry (FCM) has emerged as an indispensable tool in the diagnostic workup of primary immunodeficiencies (PID) in routine clinical practice. PIDs are a heterogeneous group of rare disorders resulting from various immune system defects [1,2]. The International Union of Immunological Societies (IUIS) has reported approximately 485 different PIDs across ten categories, and their (phenotypical) classifications include, among others, FCM criteria for PID diagnosis [2,3]. Early diagnosis of life-threatening forms of PID, such as severe combined immunodeficiency (SCID), is crucial in facilitating prompt intervention and prevention of complications [4–7]. Additionally, screening for immunophenotypic abnormalities helps risk assessment in common variable immunodeficiency (CVID) [8–10].

FCM standardization across diagnostic laboratories has become an increasing need, aiming to address variability arising from diverse local standard operating procedures (SOPs). Collaborative initiatives in clinical immunology have proven crucial in improving reproducibility and exchanging data among laboratories, especially for rare conditions, such as PID [11–22]. Moreover, standardized FCM procedures have facilitated integration of automated data analysis [23–26]. While several published recommendations for FCM standardization often focus on antibody panel design for disease-specific diagnosis in PID, including SCID, agammaglobulinemia, CVID, hyper-IgM syndrome (HIGM), and autoimmune lymphoproliferative syndrome (ALPS), the EuroFlow Consortium stands out for standardizing the full FCM process [16,18–22]. This consortium provides SOPs for instrument setup and calibration, sample preparation, and reagent panels, contributing greatly to standardized diagnostics also by developing software tools for data analysis, comparison against reference datasets, and disease classification [11,23,24,27–37]. In PID, EuroFlow has introduced and validated an innovative diagnostic algorithm, including different antibody panels, with a primary focus on fast and extensive characterization of lymphocyte cell populations in blood [38–40]. Central in this algorithm, is the PID Orientation Tube (PIDOT), which was shown highly effective in detecting immunophenotypic abnormalities in lymphoid cell counts, especially in severe forms of PID, such as SCID, in combination with the EuroFlow reference database and age-matched reference ranges, aiding in early lymphoid PID diagnostic screening [39,40]. Additionally, PIDOT, able to identify more than 27 (mainly lymphoid) subpopulations, may also be suitable in detection of acquired or secondary immunodeficiencies, as has recently been proven in Good's syndrome [37].

According to the ISO 15189 International Standard for medical laboratories, participation in external quality assessment schemes (EQAS) is required for maintaining and enhancing laboratory performance [41–43]. EQAS encompass evaluating and monitoring the laboratory's operational efficiency, tracking progress in harmonization and standardization efforts, identifying interlaboratory variations, and promptly detecting errors in laboratory procedures, enabling corrective actions [41–43]. Recognizing the importance of EQAS, EuroFlow has developed innovative and easy-to-implement

EQAS, aimed at monitoring the full FCM process for clinical laboratories, which routinely use the EuroFlow protocols [28,44]. These EQAS, consisting of both 'wet' and 'dry' formats, are designed to mimic local routine FCM procedures without the need of complex sample distribution to participants [36,44]. The wet design monitors the entire FCM process, assesses sample preparation and staining, cytometer configuration, and data acquisition, and analysis by evaluating immune staining intensities of predefined 'landmark' cell populations in locally collected samples [28,44,45]. It was previously shown that the FCM variability in these locally collected samples of EuroFlow followers is not larger than in centrally distributed samples [36,44]. The dry design evaluates data analysis, clinical interpretation, and reporting of centrally distributed FCM data of well-defined patients. In lymphoid PID, this dry design is of particular interest, because reporting and interpreting FCM results present significant challenges for laboratory immunologists due to the rarity of lymphoid PID, the diverse immunophenotypic profiles observed amongst specific subtypes and the complexity of analyzing multiparameter FCM data.

Here, we present the design and summary results of the EuroFlow PIDOT external quality assurance scheme (PIDOT-EQAS), spanning across six consecutive EQAS rounds with active participants of 20 laboratories from diverse regions across 11 countries on four continents, using five different FCM instruments from three manufacturers.

METHODS

EQAS set-up and participation

The PIDOT-EQAS, organized by the EuroFlow Consortium and coordinated by the Laboratory of Hematology and Clinical Immunology, University Hospital Ghent (Ghent, Belgium) and the Childhood Leukaemia Investigation Prague (CLIP)-Cytometry Laboratory, 2nd Med School, Charles University (Prague, Czechia), comprises two formats: a wet and dry format. The PIDOT-EQAS underwent beta-testing, involving seven EuroFlow member laboratories (EF members) across two rounds in May and October 2021, with five labs participating in both. The dry format was exclusively tested during the second beta-testing round, including four EF members. Since January 2022, the PIDOT-EQAS has been made public and conducted twice per year, following the scheme presented in Figure 1. Across four EQAS rounds, 20 EF member and Euroflow follower (EF follower, i.e., lab that employs the EuroFlow standardized protocols, but are not EF members) laboratories from 11 countries participated (Table S1), with 85% of the labs participating in multiple rounds. All involved laboratories except one were clinical laboratories, and 12 out of 20 were located in Europe, many of them linked to University Hospitals (9 out of 20). EQAS instructions were made available on the EuroFlow EQAS website at the start of each EQAS round.

Wet format

Blood samples, staining procedures, and data acquisition

Peripheral blood (PB) samples from healthy volunteers (HC) were locally collected in each participating laboratory in EDTA tubes (three HC per EQAS round). As per EQAS instructions, sample processing

including the EuroFlow bulk lysis protocol, followed by the EuroFlow stain-wash protocol with acquisition of at least 106 events after storage for a maximum of 1h at 4°C per sample, was performed [34,36]. BD FACSLyric users were instructed to use the electronic EuroFlow PIDOT assay, available on the EuroFlow website, which contains all information needed for standardized PIDOT acquisition without the need to manually setup the photomultiplier tube (PMT) voltages, except for the compensation [45].

Flow cytometer instruments and reagents

Reported data were acquired on different flow cytometer instruments, including BD FACSCanto II (BD Biosciences, San Jose, CA), BD FACSLyric (BD Biosciences), Omnicyt (Cytognos, S.L., Salamanca, Spain), BC DxFlex (Beckman Coulter, Brea, CA), and BC Navios (Beckman Coulter) (Table S1). Participants were instructed to follow the EuroFlow SOP for instrument setup and compensation [36]. The 8-color/12-marker PIDOT reagents and titers were recommended, previously published by van der Burg M. *et al.* (2019) and available at www.euroflow.org [40].

Dry format

For the dry format, PB samples of well-defined PID cases were analyzed in the reference lab (University Hospital Ghent), using the same sample processing and staining, as described in the EQAS instructions. The samples were acquired on BD FACSCanto II or BD FACSLyric. The generated flow cytometry standard (FCS) files were anonymized and distributed through a secured server to the participants (2 files per EQAS round). Participants were provided with white blood cell counts and clinical information as listed in Table S2.

Flow cytometric data analysis

FCS data analysis of the wet and dry format were performed locally by the participants. Participants were required to use either the Infinicyt Automated Gating & Identification (AG&I-) tool (Cytognos, S.L., Salamanca, Spain) ('AG&I-reports') or to apply manual gating (MG) ('MG-reports') [23]. The EQAS instructions provided an illustrative example of the MG strategy (Figure S1), as well as a link to a digital PIDOT analysis template for the Infinicyt software, which offers guidance on MG.

Data reporting

Participants were required to submit online reports (see Figure S2) per round through a secured server within 30 days from the fixed starting date. The wet format report form included reporting of median fluorescence intensities (MedFIs) from the three locally obtained HC samples per EQAS round stained with PIDOT, following the EQAS instructions (Figure S2). The dry format report form included reporting of absolute cell counts (cells/µL) and 3 different qualitative interpretations (levels 1-3) -1) cell counts interpretation against age-matched reference values as either 'absent', 'decreased', 'normal' or 'increased'; 2) the combined interpretation of the T and B cell maturation patterns; and 3) the most compatible PID subtype based on the immunophenotype and limited clinical information provided-,

with level 2 and 3 being optional (Figure S2). As per EQAS instructions, participants were advised to use Euroflow age-matched reference ranges (published by van der Burg et al. (2019) or available in the AG&I-tool) or their own local reference ranges for cell count interpretation [40]. For the dry format, participants calculated absolute cell counts based on the white blood cell count provided using their local procedures.

Defining EQAS target values

For the wet format, a repository of FCS files was created from 96 HC samples analyzed from 2016-2021 by EF members and during beta-testing round one (i.e., stained with PIDOT, and analyzed following EuroFlow SOPs) on different flow cytometers, and used to compute MedFI values and corresponding reproducibility (%rCV) for each marker subset combination (n = 11). More details on definitions, the stability and %rCV of these MedFIs are presented in Supplementary Methods, Table 1 and Table S3 [46]. Next, the MedFI values were used to define EQAS targets (Table 1) -i.e., the Dmax and the median of MedFIs (qaMedFI)- according to the methods described by Kalina et al. for the Lymphocyte Screening Tube (LST) [44]. See definitions in supplementary Methods.

For the dry format, FCS files of PID cases were collected by the reference lab (University Hospital Ghent) and categorized per IUIS diagnostic category, distinguishing severe T cell defects in IUIS I and IUIS II -two Ataxia Telangiectasia (AT), one CHARGE syndrome, and two SCID cases - one SCID with Omenn syndrome (OS) and one SCID with transplacental maternal engraftment (TME)- and B cell defects in IUIS III -two CVID, one HIGM, and one X-linked agammaglobulinemia (XLA) cases [2]. MG-gating with the age-matched reference ranges published by van der Burg *et al.* (2019) [40] were applied by the reference lab to define EQAS targets for absolute cell enumeration and qualitative data interpretation. The use of HC and patient data was approved by the Ethical Committee of University Hospital Ghent (BC-07300; 2016/1138).

Performance evaluation and definitions

The EQAS supervisors JN, MB, MH, and CB (Ghent), NB and TK (Prague) centrally reviewed all reported data per EQAS round and participating lab, as well as trends across (beta-)EQAS rounds, and participating labs were assessed.

For the wet format, per EQAS round and participating lab, accuracy of reported MedFIs was evaluated using "Performance score (p-score)" metrics and outliers were identified using the EQAS targets (Table 1), as published previously [44]. More details and the p-score and the acceptance criteria can be found in Supplementary Methods. The “wet format overall score” is defined as the percentage of acceptable p-scores for each laboratory across all marker subset combinations per EQAS round. An overall score above 91% (30/33) was considered successful [44]. Scorings between 76% and 91% were considered acceptable. Scorings below 76%, matching the 10^{th} percentile of the well-established EuroFlow LST-EQAS, were considered unsuccessful [28]. The “percentage of failure” is defined as the

percentage of p-scores falling outside the acceptable range per marker subset combination across all EQAS rounds. Reproducibility of reported MedFIs for each EQAS round was assessed.

For the dry format, per EQAS round and participating lab, accuracy of reported absolute cell counts was evaluated through the calculation of robust z-scores for each cell population and PID case, using the formula: z-score = (R-M)/rSD with R = cell count reported by participant lab, median (M) and rSD of cell counts observed across peers and the reference lab. $|\text{Z-scores}| \leq 3$ were considered acceptable. In addition, accuracy of reported qualitative data was assessed by comparing interpretations against the reference lab and peers. The "dry format overall score" is defined as the percentage of cell count interpretations matching the reference lab per laboratory and per PID case. For cell populations that are expected in very low numbers (at young age), 'normal', 'decreased' and 'absent' were considered acceptable. Reported 'most compatible PID subtype' was defined correct if it matched the reference lab response(s), with 'yes' and 'possible' considered equal. For each PID case, the reproducibility of reported absolute cell counts per population was calculated.

Trends in reproducibility and accuracy for the wet and dry format across EQAS round/PID cases/ were assessed.

Providing feedback to the participants

Upon successful completion of each EQAS round, participants received certificates, together with a summary report of results and frequently encountered issues (see Supplementary Methods). The EQAS supervisors organize educational workshops, during which participants can discuss results.

Statistical methods

Mann-Whitney U tests (MW) or Kruskal Wallis tests (KW) followed by Dunn's test for multiple comparisons were used to compare independent groups. Microsoft Excel for Windows 11 (Microsoft Corporation, Redmond, WA) and Graphpad Prism v10.1.1 (GraphPad Software, Boston, MA) were used for statistical analyses and graphical representation, respectively.

RESULTS

Laboratory performance in the PIDOT-EQAS wet format

A total of 71 reports, including reported data of 212 FCS files from HC samples, were received from 20 laboratories obtained across all EQAS rounds. Most of these FCS files (n=171 in 57 reports) were generated during the four open EQAS rounds, while the others (n=41 in 14 reports) were collected during beta-testing. The reproducibility of reported MedFIs (%rCVs) for each marker subset combination (n=11) per EQAS round and across participating laboratories is presented in Table 2 (more details are provided in Figure S3). Across EQAS rounds, this evaluation showed an average %rCV per marker subset combination ranging from 23% to 58%, with a median %rCV of 35% across all measurements (11 markers in 212 files, n=2332). The highest average %rCVs were observed for TCRγδ-

PE-Cy7 on TCRγδ+ T cells (58%), CD4-PerCP-Cy5.5 on TCD4+ cells (44%), and CD27-BV421 on unswitched B cells/plasma cells (43%).

The accuracy (p-scores) was calculated on reported MedFIs of individual marker subset combinations obtained from a set of three HC reported per participating lab and EQAS round (Figure S4A). Out of all the p-scores (n=2310) across EQAS rounds, 82.7% of the p-scores were acceptable, with median p-values falling within the acceptable range for each marker subset combination (Figure S4B). Individual outlier p-scores represented 6.7% of p-scores. Triplicates of p-scores falling outside the acceptable range were seen in each EQAS round except beta EQAS I and represented 10.6% of all p-scores (Figure S4A). The average percentage of failure of p-scores falling outside the acceptable range per marker across EQAS rounds was 17.3% (range: 11.4% - 28.6%). CD8-FITC on TCD8+ cells (28.6%) and CD4-PerCPCy5.5 on TCD4+ cells (21.9%) showed the highest percentage of failure and a higher frequency of triplicate p-scores falling outside the acceptable range, while the lowest p-scores out of range were documented for CD19-PE-Cy7 on B cells (11.4%). Across all EQAS rounds, 8/11 markers displayed a percentage of failure of 20% or less. Triplicate failure of p-scores might indicate potential issues related to reagents and/or fluorochrome quality, titration or gating mistakes.

The wet format overall scores (n=71), which indicates the percentage of acceptable p-scores per participating lab and EQAS round, and trends throughout the EQAS rounds were assessed in Figure 2A, with a mean and median of 82% and 91%, respectively. Additionally, EF followers showed a trend in improvement in average overall scores with each additional participation round, reaching 82.9% after more than three rounds, comparable to EF members participating in three rounds or more (86.1%) (Figure 2B). Furthermore, no significant differences in overall scores across EQAS rounds were observed between EF members and EF followers, liquid and lyophilized reagents, but lower overall scores were noted for participants using Navios flow cytometers (median = 76%; range: 0% - 79%) (Figures 2C and S5).

Laboratory performance in the PIDOT-EQAS dry format

A total of 59 reports were received from 20 laboratories obtained across all EQAS rounds. The reproducibility of reported absolute cell counts (%rCVs) among participants per cell population and per PID case are presented in Table 3 (Figures S6 and S7). Across PID cases, reported absolute cell counts of most of the cell populations (84.5%, n = 136/161) had %rCVs below 30%, with the absolute cell counts of all major lymphocyte subsets (i.e., total B, total T, TCD4+, TCD8+ and NK cells) showing low average %rCVs of 3.1% to 7.1%. Interestingly, laboratories using the AG&I-tool demonstrated significantly lower cell count variability compared to those applying manual gating (MG) (median %rCV: 2.5% vs 7.6%, p=0.0002, MW). To assess the accuracy of reported absolute cell counts, z-score statistics were used to highlight outliers in the analyses, identifying the laboratories with most outliers (2.6% - 14.8% of all reported values were outliers) (Table S4). Additionally, 87.6% of the reported cell

counts of relevant cell populations for the different PID cases, presented in Figure 3, were considered acceptable (|Z-scores| ≤ 3).

The accuracy to interpret the absolute cell counts in relation to the locally applied reference values per PID case is shown in Figure 4, left panel. Across PID cases, after excluding missing values - 5.7% of all values (n=125/2204)-, 79.3% (n=1649/2079) of the cell count interpretations aligned with the reference lab values. The mean and median of the dry format overall scores (indicating percentage of cell count interpretations matching the reference lab) were 85.3% and 89.0%, respectively (Figure 5A). No differences were seen in the dry format overall scores in AG&I-reports vs MG-reports (89% vs 88%, p=0.2, MW) (Figure 5B). The dry format overall scores were significantly higher for the EF members vs EF followers (94% vs 85%, p=0.002, MW) and varied among PID cases (range: 73% - 100%; CVID vs XLA, p_{adj}=0.0241; XLA vs SCID OS, p_{adj}=0.0355; KW) (Figure 5B).

In addition, laboratories were given the option to identify the most compatible PID subtype in relation to the immunophenotype and minimal clinical information provided (Figure 4, right panel and Table S5). Overall, most of the laboratories that opted to participate in this format (response rate exceeded 80% per case) were able to recognize severely altered immune cell patterns that are compatible with severe PID cases, such as SCID and agammaglobulinemia (Supplementary Results).

DISCUSSION

The EuroFlow PIDOT is a powerful tool for lymphoid PID diagnostic screening [39,40]. In extension, the PIDOT-EQAS scheme was launched for monitoring local PIDOT execution and tracking potential technical and interpretational issues. Unique in this scheme is that evaluation of the performance of the laboratory technique in local settings (wet format) is separated from data interpretation (dry format).

Building on experience gained from the EuroFlow LST-EQAS, the PIDOT-EQAS applied a similar set-up for the wet format using locally collected and processed samples from healthy volunteers [28,44]. This solution circumvents the need for a fixed cellular sample distribution, which would be impossible in rare disease settings (e.g., with low sample cellularity due to frequent lymphopenia and/or neutropenia) and thus reaches a much larger number of participants worldwide. Moreover, it was already demonstrated that the variability of MedFI values of specific PB subsets is not larger in this approach compared to centrally collected and distributed samples [36,44]. The PIDOT-EQAS demonstrated a median MedFI %rCV of 34.5%, in line with the LST-EQAS's median MedFI %CV of 32.5%, and demonstrated similar reproducibility in markers assessed in both PIDOT and LST-EQAS (CD8-FITC, CD19-PE, and CD3-APC) [44]. Assessing laboratory performances using wet format overall scores demonstrated improvement with repeated participation similar to the trends observed in LST-EQAS [44,45]. Interestingly, no significant differences were detected in overall scores obtained after liquid *vs* lyophilized staining, and across most device types, in line with previous EuroFlow studies [33,35,47].

An important difference with other EQAS/proficiency testing programs for immunodeficiency diagnostics (e.g., INSTAND EQAS, UK-NEQAS), is that the PIDOT-EQAS does not distribute (stabilized) whole blood samples amongst the participants but is built on the locally collected samples (wet format) and electronically distributed FCS filed (dry format) [44]. For example, UKNEQAS distributes robust stabilized lymphocyte suspensions, while INSTAND EQAS delivers fresh, unstabilized blood without refrigeration, primarily in the German-speaking region. By eliminating sample distribution, PIDOT-EQAS avoids the risks of sample deterioration by fixation and/or by lengthy transport and ensures that the evaluated samples represent 'real-life' sample matrices [36,44]. Also, it circumvents the need to collect a large sample volume from a pediatric patient for distribution [44]. Additionally, the number of cell populations assessed varies across programs. For instance, the UKNEQAS evaluates a limited set (i.e., total T, TCD4+, TCD8+, B and NK cells), while INSTAND EQAS and PIDOT-EQAS conduct more comprehensive evaluations, including additional T and B cell maturation subsets [48–50]. The low variability in absolute cell count enumeration observed for the major lymphocyte subsets in the PIDOT-EQAS aligns with observations in other studies [15,49]. Furthermore, PIDOT-EQAS demonstrate a significant reduction in cell count variability when using the AG&I-tool compared to manual gating, consistent with the findings of our previous study [23].

The PIDOT-EQAS provides a major advantage by not only evaluating the technical quality in FCM measurements and absolute cell count enumeration following data analysis among clinical labs, but also assessing their capabilities to interpret and report complex FCM data of PID cases. In daily diagnostic PID care, it is of utmost important that both specialized and non-specialized labs have the proficiency to descriptively identify relevant lymphoid cell abnormalities through the analysis and interpretation of FCM data against age-matched reference ranges. This study demonstrated that the vast majority of laboratories successfully identified relevant lymphoid abnormalities in all cases analysed. Of note, discordances in cell count interpretations for individual cell types observed during these EQAS rounds could be attributed to variations in the age-matched reference ranges and/or data analysis strategies (i.e., MG or AG&I) locally applied (Figure S8). To improve the results, we advocate the use of EuroFlow standardized FCM procedures and the regularly updated age-matched reference values, available on the EuroFlow website after online registration (https://euroflow.org/protocols/), aiming for increased standardization.

Moreover, the PIDOT-EQAS can be regarded as an accessible training tool for laboratories to enhance their skills in recognizing the clinically important but rare abnormal immunophenotypic patterns that urgently need further clinical management action (particularly SCID). However, it should be noted that clinical interpretation of flow cytometric immune profiles is not a task of the clinical (immunology) lab specialists in every country. Due to the variability in tasks assigned to the lab specialists worldwide, this part of the dry format was made optional. Our data demonstrate that almost all labs (that opted to answer this part of the dry format) successfully identified the SCID cases based

on the immunophenotype and limited clinical information provided, and centres experiencing challenges in this regard were identified. Therefore, the PIDOT-EQAS also holds an important instructive value, given the fact that even in specialized diagnostic labs such cases only occur 2-3 times a year. For cases diagnosed with CID, ranging from 'CID less profound than SCID' to 'CID with associated and/or syndromic features', as well as cases with primary B cell defects, which exhibit a broader spectrum of T and/or B cell abnormalities, the immunophenotype might be linked with various PID subtypes and is therefore not specific, aligning with the corresponding diagnostic criteria, as also captured by the diversity in the responses obtained in the PIDOT-EQAS. Effective communication between the flow laboratory and clinical departments is therefore critical to integrate flow data with the clinical context and determine the need for further functional and genetic analyses.

The PIDOT-EQAS has some limitations. First, the PIDOT antibody cocktail, and thus also the EQAS only cover a limited repertoire of the (lymphoid) PIDs that can be detected by FCM. Indeed, the antibody selection as well as the PIDOT format (no intracellular staining nor stimulation) does not allow for identification of the full spectrum of lymphoid PIDs (i.e. STAT3 DN, ZNF341, defects of BTK, CD21, CD132, CD40 ligand, ...), but is mostly focussed on identifying absent populations and/or deviations in the maturation patterns. Secondly, this EQAS focusses solely on the PIDOT assay performed according to the EuroFlow principles, and thus cannot replace established EQAS/proficiency testing (PT) studies in which more populations and more analytical variables are evaluated. Thirdly, the wet format of the PIDOT-EQAS provides EuroFlow adopters with an answer to the question: "Can I obtain data in my own setting comparable with the reference EuroFlow data set ?". Non-adopters can still gain educational support from the software analysis of the provided PID cases and their interpretation. Fourthly, our results indicated that further group efforts are needed to derive optimized PIDOT-specific target values for specific instrument hardware, such as Navios devices [33]. Finally, suboptimal results can arise from several sources -e.g. reagents volumes and/or quality, bulk lysis methods or insufficient washing, compensation or gating issues; or use of other age-matched reference ranges or gating. The QA supervisors cannot pinpoint the exact issue causing suboptimal results for individual participants; it is the responsibility of the participating lab to trace back their potential issues.

In summary, PIDOT-EQAS effectively monitors the quality of PIDOT immunophenotyping analysis, and reporting procedures in local settings, which is enabled by EuroFlow standardization. Furthermore, PIDOT-EQAS offers a unique educational support platform to learn from very severe cases that urgently require further clinical care and that are rarely encountered even in the setting of national diagnostic reference centres. The PIDOT-EQAS leads to improved quality flow cytometry testing in PID screening, serving also as an example for future immunophenotyping EQAS in the context of rare diseases.

ACKNOWLEDGEMENTS

All members of the EuroFlow consortium and participating laboratories are acknowledged for their contribution to the network and for fruitful discussions.

RESEARCH ETHICS

This study was approved by University Hospital Ghent Ethics Committee, Ghent, Belgium (BC-07300; 2016/1138) and was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

INFORMED CONSENT

Informed consent was obtained from all individuals included in this study, or their legal guardians or wards.

AUTHOR CONTRIBUTIONS

ToK, CB, MH, AO and JJMvD contributed to the conception and design of the study. For the wet format, CB, MH, XB, MPA, AO, and JJMvD provided flow cytometric data for reference data set. For the dry format, the team of the University Hospital Ghent (CB, MH, MB, and JN) was assigned as the reference lab and was responsible for the pre-round preparation, including PID case selection, data acquisition, quality control of FCS files, data analysis and interpretation. TeK, CDV and FH (University Hospital Ghent) provided the clinical data of the PID cases. In the post-round phase, ToK, and NB were responsible for the collection of the participants' results via the EuroFlow EQAS website and the generation of the EQAS certificates. JN and MB were assigned as second reviewer of the EQAS certificates and generated the summary reports under the supervision of MH, CB and ToK. JN drafted the original draft manuscript. MB, MH, CB, XB, MPA, AO, NB, and ToK revised and edited the original draft of the manuscript. All authors contributed to the manuscript revision, read and approved the submitted version.

COMPETING INTERESTS

All other authors state no conflict of interest.

RESEARCH FUNDING

This study has been supported by an FWO TBM grant (Research Foundation – Flanders. T000119N). JN is a PhD fellow of the Research Foundation Flanders (FWO 11L2822N). TK was supported by the project NU23-07-00170 of the Czech Republic Ministry of Health, Charles University Research Centre program No. UNCE/24/MED/003 and the European Union – Next Generation EU – program No. LX22NPO5102. MPA received funding from the Instituto de Salud Carlos III (ISCIII) through the project "PI20/01712" and co-founded by the European Union. The coordination and innovation processes of this study were supported by the EuroFlow Consortium.

DATA AVAILABILITY

The raw data can be obtained on request from the corresponding author.

411 **REFERENCES**

[1] Bousfiha A, Jeddane L, Picard C, Al-Herz W, Ailal F, Chatila T, et al. Human Inborn Errors of Immunity: 2019 Update of the IUIS Phenotypical Classification. J Clin Immunol 2020;40:66–81. https://doi.org/10.1007/s10875-020-00758-x.

[2] Tangye SG, Al-Herz W, Bousfiha A, Cunningham-Rundles C, Franco JL, Holland SM, et al. Human Inborn Errors of Immunity: 2022 Update on the Classification from the International Union of Immunological Societies Expert Committee. J Clin Immunol 2022;42:1473–507. https://doi.org/10.1007/s10875-022-01289-3.

[3] Ma CS, Tangye SG. Flow Cytometric-Based Analysis of Defects in Lymphocyte Differentiation and Function Due to Inborn Errors of Immunity. Front Immunol 2019;10:2108. https://doi.org/10.3389/fimmu.2019.02108.

[4] Kalina T, Bakardjieva M, Blom M, Perez-Andres M, Barendregt B, Kanderová V, et al. EuroFlow Standardized Approach to Diagnostic Immunopheneotyping of Severe PID in Newborns and Young Children. Front Immunol 2020;11:371. https://doi.org/10.3389/fimmu.2020.00371.

[5] Shai S, Perez-Becker R, Andres O, Bakhtiar S, Bauman U, von Bernuth H, et al. Incidence of SCID in Germany from 2014 to 2015 an ESPED* Survey on Behalf of the API*** Erhebungseinheit für Seltene Pädiatrische Erkrankungen in Deutschland (German Paediatric Surveillance Unit) ** Arbeitsgemeinschaft Pädiatrische Immunologie. J Clin Immunol 2020;40:708–17. https://doi.org/10.1007/s10875-020-00782-x.

[6] Kwan A, Abraham RS, Currier R, Brower A, Andruszewski K, Abbott JK, et al. Newborn screening for severe combined immunodeficiency in 11 screening programs in the United States. JAMA 2014;312:729–38. https://doi.org/10.1001/jama.2014.9132.

[7] Ameratunga R, Woon ST, Gillis D, Koopmans W, Steele R. New diagnostic criteria for common variable immune deficiency (CVID), which may assist with decisions to treat with intravenous or subcutaneous immunoglobulin. Clin Exp Immunol 2013;174:203–11. https://doi.org/10.1111/cei.12178.

[8] Wehr C, Kivioja T, Schmitt C, Ferry B, Witte T, Eren E, et al. The EUROclass trial: defining subgroups in common variable immunodeficiency. Blood 2008;111:77–85. https://doi.org/10.1182/blood-2007-06-091744.

[9] Warnatz K, Denz A, Dräger R, Braun M, Groth C, Wolff-Vorbeck G, et al. Severe deficiency of switched memory B cells (CD27(+)IgM(-)IgD(-)) in subgroups of patients with common variable immunodeficiency: a new approach to classify a heterogeneous disease. Blood 2002;99:1544–51. https://doi.org/10.1182/blood.v99.5.1544.

[10] Blanco E, Perez-Andres M, Arriba-Mendez S, Serrano C, Criado I, Del Pino-Molina L, et al. Defects in memory B-cell and plasma cell subsets expressing different immunoglobulin-subclasses in patients with CVID and immunoglobulin subclass deficiencies. J Allergy Clin Immunol 2019;144:809–24. https://doi.org/10.1016/j.jaci.2019.02.017.

[11] Kalina T. Reproducibility of Flow Cytometry Through Standardization: Opportunities and Challenges. Cytometry A 2020;97:137–47. https://doi.org/10.1002/cyto.a.23901.

[12] Maecker HT, McCoy JP, Nussenblatt R. Standardizing immunophenotyping for the Human Immunology Project. Nat Rev Immunol 2012;12:191–200. https://doi.org/10.1038/nri3158.

[13] Streitz M, Miloud T, Kapinsky M, Reed MR, Magari R, Geissler EK, et al. Standardization of whole blood immune phenotype monitoring for clinical trials: panels and methods from the ONE study. Transplant Res 2013;2:17. https://doi.org/10.1186/2047-1440-2-17.

[14] Ivison S, Malek M, Garcia RV, Broady R, Halpin A, Richaud M, et al. A standardized immune phenotyping and automated data analysis platform for multicenter biomarker studies. JCI Insight 2018;3. https://doi.org/10.1172/jci.insight.121867.

[15] Jamin C, Le Lann L, Alvarez-Errico D, Barbarroja N, Cantaert T, Ducreux J, et al. Multi-center harmonization of flow cytometers in the context of the European “PRECISESADS” project. Autoimmun Rev 2016;15:1038–45. https://doi.org/10.1016/j.autrev.2016.07.034.

[16] Boldt A, Borte S, Fricke S, Kentouche K, Emmrich F, Borte M, et al. Eight-color immunophenotyping of T-, B-, and NK-cell subpopulations for characterization of chronic immunodeficiencies. Cytometry B Clin Cytom 2014;86:191–206. https://doi.org/10.1002/cyto.b.21162.

[17] Biancotto A, Fuchs JC, Williams A, Dagur PK, McCoy JP Jr. High dimensional flow cytometry for comprehensive leukocyte immunophenotyping (CLIP) in translational research. J Immunol Methods 2011;363:245–61. https://doi.org/10.1016/j.jim.2010.06.010.

[18] Abraham RS, Aubert G. Flow Cytometry, a Versatile Tool for Diagnosis and Monitoring of Primary Immunodeficiencies. Clin Vaccine Immunol 2016;23:254–71. https://doi.org/10.1128/cvi.00001-16.

[19] O’Gorman MRG. Flow Cytometry Assays in Primary Immunodeficiency Diseases. Methods Mol Biol 2018;1678:321–45. https://doi.org/10.1007/978-1-4939-7346-0_14.

[20] Takashima T, Okamura M, Yeh TW, Okano T, Yamashita M, Tanaka K, et al. Multicolor Flow Cytometry for the Diagnosis of Primary Immunodeficiency Diseases. J Clin Immunol 2017;37:486–95. https://doi.org/10.1007/s10875-017-0405-7.
[21] Oliveira JB, Notarangelo LD, Fleisher TA. Applications of flow cytometry for the study of primary immune deficiencies. Curr Opin Allergy Clin Immunol 2008;8:499–509. https://doi.org/10.1097/ACI.0b013e328312c790.
[22] Kanegane H, Hoshino A, Okano T, Yasumi T, Wada T, Takada H, et al. Flow cytometry-based diagnosis of primary immunodeficiency diseases. Allergol Int 2018;67:43–54. https://doi.org/10.1016/j.alit.2017.06.003.
[23] Linskens E, Diks AM, Neirinck J, Perez-Andres M, De Maertelaere E, Berkowska MA, et al. Improved Standardization of Flow Cytometry Diagnostic Screening of Primary Immunodeficiency by Software-Based Automated Gating. Front Immunol 2020;11:584646. https://doi.org/10.3389/fimmu.2020.584646.
[24] Pedreira CE, Costa ESD, Lecrevise Q, Grigore G, Fluxa R, Verde J, et al. From big flow cytometry datasets to smart diagnostic strategies: The EuroFlow approach. J Immunol Methods 2019:112631. https://doi.org/10.1016/j.jim.2019.07.003.
[25] Van Gassen S, Callebaut B, Van Helden MJ, Lambrecht BN, Demeester P, Dhaene T, et al. FlowSOM: Using self-organizing maps for visualization and interpretation of cytometry data. Cytometry A 2015;87:636–45. https://doi.org/10.1002/cyto.a.22625.
[26] Emmaneel A, Quintelier K, Sichien D, Rybakowska P, Marañón C, Alarcón-Riquelme ME, et al. PeacoQC: Peak-based selection of high quality cytometry data. Cytometry A 2022;101:325–38. https://doi.org/10.1002/cyto.a.24501.
[27] van Dongen JJ, Lhermitte L, Böttcher S, Almeida J, van der Velden VH, Flores-Montero J, et al. EuroFlow antibody panels for standardized n-dimensional flow cytometric immunophenotyping of normal, reactive and malignant leukocytes. Leukemia 2012;26:1908–75. https://doi.org/10.1038/leu.2012.120.
[28] Kalina T, Brdickova N, Glier H, Fernandez P, Bitter M, Flores-Montero J, et al. Frequent issues and lessons learned from EuroFlow QA. J Immunol Methods 2019;475:112520. https://doi.org/10.1016/j.jim.2018.09.008.
[29] van Dongen JJ, Orfao A. EuroFlow: Resetting leukemia and lymphoma immunophenotyping. Basis for companion diagnostics and personalized medicine. Leukemia 2012;26:1899–907. https://doi.org/10.1038/leu.2012.121.
[30] Lhermitte L, Mejstrikova E, van der Sluijs-Gelling AJ, Grigore GE, Sedek L, Bras AE, et al. Automated database-guided expert-supervised orientation for immunophenotypic diagnosis and classification of acute leukemia. Leukemia 2018;32:874–81. https://doi.org/10.1038/leu.2017.313.
[31] Flores-Montero J, Kalina T, Corral-Mateos A, Sanoja-Flores L, Perez-Andres M, Martin-Ayuso M, et al. Fluorochrome choices for multi-color flow cytometry. J Immunol Methods 2019;475:112618. https://doi.org/10.1016/j.jim.2019.06.009.
[32] Flores-Montero J, Grigore G, Fluxá R, Hernández J, Fernandez P, Almeida J, et al. EuroFlow Lymphoid Screening Tube (LST) data base for automated identification of blood lymphocyte subsets. J Immunol Methods 2019;475:112662. https://doi.org/10.1016/j.jim.2019.112662.
[33] Novakova M, Glier H, Brdickova N, Vlkova M, Santos AH, Lima M, et al. How to make usage of the standardized EuroFlow 8-color protocols possible for instruments of different manufacturers. J Immunol Methods 2017;475:112388. https://doi.org/10.1016/j.jim.2017.11.007.
[34] Diks AM, Bonroy C, Teodosio C, Groenland RJ, de Mooij B, de Maertelaere E, et al. Impact of blood storage and sample handling on quality of high dimensional flow cytometric data in multicenter clinical research. J Immunol Methods 2019;475:112616. https://doi.org/10.1016/j.jim.2019.06.007.
[35] van der Velden VH, Flores-Montero J, Perez-Andres M, Martin-Ayuso M, Crespo O, Blanco E, et al. Optimization and testing of dried antibody tube: The EuroFlow LST and PIDOT tubes as examples. J Immunol Methods 2017;475:112287. https://doi.org/10.1016/j.jim.2017.03.011.
[36] Kalina T, Flores-Montero J, van der Velden VH, Martin-Ayuso M, Bottcher S, Ritgen M, et al. EuroFlow standardization of flow cytometer instrument settings and immunophenotyping protocols. Leukemia 2012;26:1986–2010. https://doi.org/10.1038/leu.2012.122.
[37] Torres-Valle A, Aragon L, Silva SL, Serrano C, Marcos M, Melero J, et al. In-depth blood immune profiling of Good syndrome patients. Frontiers in Immunology 2023;14.
[38] van Dongen JJM, van der Burg M, Kalina T, Perez-Andres M, Mejstrikova E, Vlkova M, et al. EuroFlow-Based Flowcytometric Diagnostic Screening and Classification of Primary Immunodeficiencies of the Lymphoid System. Front Immunol 2019;10:1271. https://doi.org/10.3389/fimmu.2019.01271.
[39] Neirinck J, Emmaneel A, Buysse M, Philippé J, Van Gassen S, Saeys Y, et al. The Euroflow PID Orientation Tube in the diagnostic workup of primary immunodeficiency: Daily practice performance in a tertiary university hospital. Frontiers in Immunology 2022;13. https://doi.org/10.3389/fimmu.2022.937738.

[40] van der Burg M, Kalina T, Perez-Andres M, Vlkova M, Lopez-Granados E, Blanco E, et al. The EuroFlow PID Orientation Tube for Flow Cytometric Diagnostic Screening of Primary Immunodeficiencies of the Lymphoid System. Front Immunol 2019;10:246. https://doi.org/10.3389/fimmu.2019.00246.
[41] Blasutig IM, Wheeler SE, Bais R, Dabla PK, Lin J, Perret-Liaudet A, et al. External quality assessment practices in medical laboratories: an IFCC global survey of member societies. Clinical Chemistry and Laboratory Medicine (CCLM) 2023. https://doi.org/10.1515/cclm-2023-0057.
[42] Sciacovelli L, Secchiero S, Padoan A, Plebani M. External quality assessment programs in the context of ISO 15189 accreditation. Clinical Chemistry and Laboratory Medicine (CCLM) 2018;56:1644–54. https://doi.org/10.1515/cclm-2017-1179.
[43] Keppens C, Boone E, Gameiro P, Tack V, Moreau E, Hodges E, et al. Evaluation of a worldwide EQA scheme for complex clonality analysis of clinical lymphoproliferative cases demonstrates a learning effect. Virchows Arch 2021;479:365–76. https://doi.org/10.1007/s00428-021-03046-0.
[44] Kalina T, Flores-Montero J, Lecrevisse Q, Pedreira CE, van der Velden VH, Novakova M, et al. Quality assessment program for EuroFlow protocols: summary results of four-year (2010-2013) quality assurance rounds. Cytometry A 2015;87:145–56. https://doi.org/10.1002/cyto.a.22581.
[45] Glier H, Heijnen I, Hauwel M, Dirks J, Quarroz S, Lehmann T, et al. Standardization of 8-color flow cytometry across different flow cytometer instruments: A feasibility study in clinical laboratories in Switzerland. J Immunol Methods 2019;475:112348. https://doi.org/10.1016/j.jim.2017.07.013.
[46] Wan X, Wang W, Liu J, Tong T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. BMC Medical Research Methodology 2014;14:135. https://doi.org/10.1186/1471-2288-14-135.
[47] Glier H, Novakova M, Te Marvelde J, Bijkerk A, Morf D, Thurner D, et al. Comments on EuroFlow standard operating procedures for instrument setup and compensation for BD FACS Canto II, Navios and BD FACS Lyric instruments. J Immunol Methods 2019;475:112680. https://doi.org/10.1016/j.jim.2019.112680.
[48] Immunophenotyping - Cellular Immunodeficiency Diagnostic - INSTAND e.V. n.d. https://www.instand-ev.de/en/instand-eqas/eqa-program/offer/immunophenotyping-cellular-immunodeficiency-diagnostic/ (accessed January 11, 2024).
[49] Reilly JT, Barnett D. UK NEQAS for leucocyte immunophenotyping: the first 10 years. J Clin Pathol 2001;54:508–11. https://doi.org/10.1136/jcp.54.7.508.
[50] Whitby L, Granger V, Storie I, Goodfellow K, Sawle A, Reilly JT, et al. Quality control of CD4+ T-lymphocyte enumeration: results from the last 9 years of the United Kingdom National External Quality Assessment Scheme for Immune Monitoring (1993-2001). Cytometry 2002;50:102–10. https://doi.org/10.1002/cyto.10094.

TABLES

Table 1 Reproducibility of median fluorescence intensity values (%rCV) of the reference dataset (n=96) and the EQAS target values (Dmax and qaMedFI) of the PIDOT-EQAS wet format

Marker	Subset	%rCV of MedFI[a]	Dmax[b]	Median MedFI [c] (qaMedFI)
IgD:FITC	On pre-germinal-center B cells	25%	0.2122	5734
IgM:PerCP-Cy5-5	On unswitched B cells + plasma cells	26%	0.2624	30017
CD27:BV421	On unswitched B cells + plasma cells	32%	0.2894	2557
CD4:PerCP-Cy5-5	On TCD4+ cells	28%	0.2309	7153
CD8:FITC	On TCD8+ cells	17%	0.1507	13542
CD19:PE-Cy7	On B cells	21%	0.1976	13553
CD45RA:BV510	On B cells	19%	0.2889	33664
CD3:APC	On T cells	23%	0.1871	37691
CD45:APCH7 or APC-C750	On T cells	30%	0.2473	22243
TCRγδ:PE-Cy7	On TCRγδ+ T cells	57%	0.3960	6991
CD16+CD56:PE	On NK cells	30%	0.3263	36218

[a]The reproducibility of MedFIs per marker subset combination expressed as robust coefficient of variations (%rCV) with formula: %rCV = [robust Standard Deviation (rSD)]/(median) X 100 (%) with rSD calculated using the interquartile range and sample size [46] bThe Dmax is defined as the maximum allowable difference, calculated as the 95th percentile of all absolute differences between individual MedFIs and the qaMedFI per marker subset combination. [c]The qaMedFI is calculated as the median MedFI per marker subset combination.

Table 2 Reproducibility of reported median fluorescence intensity values (%rCV) in the PIDOT-EQAS wet format

Marker - Subset	Beta (n=41)	EQAS I (n=45)	EQAS II (n=39)	EQAS III (n=42)	EQAS IV (n=45)	Average
IgD:FITC on PreGC B cells	34%	39%	31%	47%	31%	36%
IgM:PerCP-Cy5.5 on unswitched B cells + plasma cells	27%	39%	39%	77%	30%	42%
CD27:BV421 on unswitched B cells + plasma cells	37%	45%	40%	46%	50%	43%
CD4:PerCP-Cy5.5 on TCD4+ cells	26%	37%	35%	48%	76%	44%
CD8:FITC on TCD8+ cells	23%	43%	33%	40%	20%	32%
CD19:PE-Cy7 on B cells	18%	34%	27%	24%	28%	26%
CD45RA:BV510 on B cells	19%	34%	26%	35%	24%	27%
CD3:APC on T cells	30%	26%	12%	21%	23%	23%
CD45:APC H7 or APC-C750 on T cells	27%	32%	26%	42%	61%	38%
TCRγδ:PE-Cy7 on TCRγδ+ T cells	71%	77%	45%	37%	62%	58%
CD16+CD56:PE on NK cells	37%	52%	29%	39%	41%	40%
Average	32%	42%	31%	41%	41%	37%

Beta = beta-testing dataset; EQAS I to IV = open EQAS round datasets; PreGC: pre-germinal-center. NK: natural killer

Table 3 Reproducibility of reported absolute cell counts (%rCV) in the PIDOT-EQAS dry format.

		B cells	Pre-GC B cells	Unswitched B cells+PC	Switched B cells+PC	T cells	TCD4$^+$ cells	Naive TCD4$^+$ cells	CM TCD4$^+$ cells	EM TCD4$^+$ cells	Effector TD TCD4$^+$ cells	TCD8$^+$ cells	Naive TCD8$^+$ cells	CM TCD8$^+$ cells	EM TCD8$^+$ cells	Effector TD TCD27$^+$CD8$^+$ cells	Effector TD TCD8$^+$ cells	DN TCR$\gamma\delta^+$ T cells	TCR$\gamma\delta^+$ T cells	NK cells
CVID	Median (/µL)	109.0	94.0	13.7	1.3	1176.0	558.0	194.0	296.5	54.9	7.7	525.8	141.5	214.1	34.4	21.6	122.0	8.7	77.7	141.0
	%rCV	0.6%	0.6%	15.5%	7.8%	0.3%	1.4%	4.8%	4.1%	12.2%	70.5%	1.0%	4.1%	3.7%	37.7%	30.9%	9.9%	12.6%	0.6%	4.0%
AT	Median (/µL)	75.2	30.9	28.7	19.3	401.0	204.9	35.5	153.9	16.7	0.6	136.5	44.1	88.4	5.6	3.3	3.0	1.6	57.6	376.5
	%rCV	2.9%	14.2%	19.4%	11.6%	5.9%	6.6%	10.3%	6.8%	14.5%	-*	6.1%	13.1%	17.7%	22.5%	32.9%	14.6%	172%	9.5%	6.8%
XLA	Median (/µL)	1.2	1.2	0.0	0.0	3540	1920	1671	213	20.7	1.1	1125.5	869.0	179.9	1.9	59.6	12.4	23.0	475.5	124.0
	%rCV	45.1%	61.3%	-*	-*	1.8%	1.7%	5.0%	18.3%	7.7%	125%	1.3%	2.6%	6.3%	58.3%	1.8%	6.9%	8.6%	2.5%	2.3%
SCID OS	Median (/µL)	0.0	0.0	0.0	0.0	2835.0	2559	1.7	1418	1114	13.7	96.6	0.1	80.0	15.7	0.0	0.9	23.9	146.5	1105
	%rCV	-*	-*	-*	-*	12.8%	13.0%	222%	9.7%	17.0%	95.3%	19.8%	-*	21.2%	9.4%	-*	-*	11.4%	12.8%	14.9%
ALPS/	Median (/µL)	359.0	235.0	117.0	4.1	1745	1039	559.2	473.0	8.0	0.1	544.5	403.5	91.7	7.4	21.4	20.6	90.8	65.8	77.7
	%rCV	4.8%	2.4%	11.2%	20.0%	2.0%	2.0%	5.2%	4.1%	4.9%	-*	1.6%	1.6%	7.7%	30.5%	35.4%	11.6%	3.8%	5.7%	3.5%
CID	Median (/µL)	15.0	13.6	0.6	0.0	400.5	196.5	96.7	81.8	13.8	0.4	164.0	51.8	72.7	5.5	1.5	32.0	5.1	34.4	70.9
	%rCV	1.7%	14.0%	-*	-*	0.4%	2.1%	11.3%	14.7%	15.5%	-*	0.7%	3.1%	5.8%	6.0%	71.9%	1.7%	19.1%	2.1%	1.6%
SCID	Median (/µL)	473.5	472.5	0.3	0.0	28.0	23.6	0.0	20.5	3.0	0.0	0.9	0.0	0.9	0.0	0.0	0.0	2.3	1.1	33.7
	%rCV	1.3%	1.0%	-*	-*	1.1%	1.0%	-*	2.2%	2.1%	-*	-*	-*	-*	-*	-*	-*	7.2%	11.3%	9.0%
HIGM	Median (/µL)	186.0	86.2	99.3	0.1	1321.1	434.4	67.5	266.6	94.2	6.0	832.5	79.7	342.5	36.5	206.9	156.8	10.7	43.9	50.2
	%rCV	3.6%	2.8%	9.6%	-*	2.1%	2.4%	14.0%	4.8%	3.2%	31.4%	1.8%	9.4%	26.9%	37.8%	34.7%	4.9%	7.9%	5.4%	7.9%
Charge	Median (/µL)	1939.0	1787	109.4	35.5	33052	5550	0.1	2145	3401	0.8	23486	0.3	18673	4318	0.0	10.0	3748	66.2	745.0
	%rCV	1.4%	4.6%	54.6%	28.7%	2.1%	1.2%	-*	2.1%	2.2%	-*	2.8%	-*	4.1%	8.4%	-*	60.1%	21.6%	5.3%	7.3%
AT/	Median (/µL)	104.0	54.8	42.3	4.0	2435.5	483.0	34.3	334.5	101.0	17.1	1618.5	22.4	1135.1	220.0	16.7	192.0	8.2	318.8	684.4
	%rCV	2.5%	40.5%	60.7%	93.3%	2.5%	1.8%	42.6%	3.7%	7.4%	37.2%	3.1%	22.7%	2.8%	25.6%	68.3%	25.8%	16.7%	1.8%	5.1%
Average % rCV		7.1%	15.7%	28.5%	35.3%	3.1%	3.3%	39.4%	7.1%	8.7%	71.9%	4.2%	8.1%	10.7%	26.2%	39.4%	16.9%	28.1%	5.7%	6.2%
Median %rCV		2.5%	4.6%	17.5%	20.0%	2.1%	1.9%	10.8%	4.5%	7.6%	70.5%	1.8%	4.1%	6.3%	25.6%	34.7%	10.8%	12.0%	5.4%	6.0%

* "-": the rCV% is not interpretable due to near absent absolute cell counts (cells/µL). CM: central memory; DN: double negative; EM: effector memory; PC: plasma cells; Pre-GC B cells: pre-germinal center B cells; NK: natural killer; TD: terminally differentiated; ALPS: auto-immune lymphoproliferative syndrome; AT: ataxia-telangiectasia; CID: combined immunodeficiency; CSR: class-switched recombination defect; CVID: common variable immunodeficiency; HIGM: hyper-IgM syndrome; ME: maternal engraftment; OS: Omenn syndrome; SCID: severe combined immunodeficiency; XLA: X-linked agammaglobulinemia

FIGURE LEGENDS

Figure 1 Overview of the EuroFlow PIDOT-EQAS. The phases of each EQAS round encompassing the preparation of the EQAS rounds by the EQAS supervisors (indicated in blue), the analysis and reporting of the results by the participants (indicated in yellow), and evaluation of the reported results and preparation of the EQAS certificates by the EQAS supervisors (indicated in green).

Figure 2 Accuracy of median fluorescence intensity values (wet format overall score (%)) in the PIDOT-EQAS wet format and trends. The "wet format overall score" is defined as the percentage of acceptable p-scores for each laboratory across all marker subset combinations per EQAS round. **A** Graphical representation of the wet format overall scores per participating lab across EQAS rounds. Overall scores above 91% were considered successful. Scorings between 76% and 91% were considered acceptable, and scorings below 76% were considered unsuccessful. The grey dotted lines represent the thresholds of 76% and 91%. **B** The wet format overall score in relation to the number of participations for both EF members and EF followers **C** Comparison of the wet format overall scores across EQAS rounds, for EuroFlow (EF) member *vs* EF followers, distinct devices and reagents. Medians are indicated with horizontal lines. *$p<0.05$, Mann Whitney U Test.

Figure 3 Accuracy of reported absolute cell counts (z-scores) in the PIDOT-EQAS dry format. Dot plots show reported absolute cell counts of the most relevant cell populations per PID case. Outlier absolute cell counts with z-scores > 3 and z-scores ≤ -3 are indicated in green and red, respectively. The median cell count is indicated by a black horizontal line. The rows of the heatmaps indicate the calculated z-scores -i.e. grey indicate acceptable z-scores ($|z\text{-score}| \leq 3$), red and green colors indicate z-scores above 3 and below 3, respectively- of participating labs (A-T with "REF" = reference lab) organized per distributed PID case. The AT and CVID case, included in the beta-testing round, involved only four participating labs. Noninterpretable z-scores due to (near) absent absolute cell counts (cells/µL) are left blank. Missing values are indicated by a cross through the cell. Abbreviations: CM: central memory; DN: double negative; EM: effector memory; PC: plasma cells; Pre-GC B cells: pre-germinal center B cells; NK: natural killer; TD: terminally differentiated. ALPS: auto-immune lymphoproliferative syndrome; AT: ataxia-telangiectasia; CID: combined immunodeficiency; CSR: class-switched recombination defect; CVID: common variable immunodeficiency; HIGM: hyper-IgM syndrome; TME: transplacental maternal engraftment; OS: Omenn syndrome; SCID: severe combined immunodeficiency; XLA: X-linked agammaglobulinemia

Figure 4 Accuracy of reported qualitative data in the PIDOT-EQAS dry format. The qualitative data included the cell count interpretation of reported absolute cell counts (left) and identification of the most compatible PID subtype (right). The PID cases are organized according to the IUIS classification ((**A)** IUIS I, (**B)** IUIS II, and (**C)** IUIS III). The rows of the heatmaps indicate the reports of participating labs (A-T with "REF" = reference lab) organized per distributed PID case with the reference diagnosis

depicted in the grey boxes at the left side of the heatmaps, and the columns indicate the evaluated cell populations (left) and most compatible PID subtypes as defined in the checklist provided in the report form (right). Abbreviations: CM: central memory; DN: double negative; EM: effector memory; PC: plasma cells; Pre-GC B cells: pre-germinal center B cells; NK: natural killer; TD: terminally differentiated. ALPS: auto-immune lymphoproliferative syndrome; AT: ataxia-telangiectasia; CID: combined immunodeficiency; CSR: class-switched recombination defect; CVID: common variable immunodeficiency; HIGM: hyper-IgM syndrome; TME: transplacental maternal engraftment; OS: Omenn syndrome; SCID: severe combined immunodeficiency; XLA: X-linked agammaglobulinemia

Figure 5 Accuracy of reported cell count interpretation (dry format overall score (%)) in the PIDOT-EQAS dry format and trends. A The graph represents dry format overall scores per participating lab for the different centrally distributed well-defined PID cases classified per IUIS category. **B** The dry format overall scores in relation to the PID cases – left graph, analysis strategies – middle graph (i.e., manual gating (MG) vs automated gating and identification (AG&I)), and EuroFlow (EF) members vs EF followers – right graph. *$p<0.05$, **$p<0.001$, Mann Whitney U Test. Abbreviations: AG&I: Automated Gating and Identification; ALPS: auto-immune lymphoproliferative syndrome; AT: ataxia-telangiectasia; CID: combined immunodeficiency; CSR: class-switched recombination defect; CVID: common variable immunodeficiency; HIGM: hyper-IgM syndrome; MG: manual gating; TME: transplacental maternal engraftment; OS: Omenn syndrome; SCID: severe combined immunodeficiency; XLA: X-linked agammaglobulinemia