March 2026

Sponsored and reviewed by ICCS Quality and Standards Committee

Title: Between-Instrument Variability Assessment in Clinical Flow Cytometry Laboratories: How to Perform, and Acceptance Criteria

Authors:

Nithianandan Selliah, Ph.D., Cerba Research
Chrysanthi Tsamadou, MD., MLL Münchner Leukämielabor GmbH
Marci O'Driscoll, SCYM(ASCP), Tampa General Hospital
Chong Xu, M.D., Shanghai Center for Clinical Laboratory
Ahmad Al-Attar, Ph.D., University of Pittsburg Medical Center

INTRODUCTION

In clinical flow cytometry, maintaining reliable and reproducible assay performance across multiple instruments is critical for ensuring consistent data quality - within a single laboratory for longitudinal clinical trials. Between-instrument variability assessment evaluates the degree to which identical samples yield comparable results when acquired on different cytometers of the same or different models. This step is essential for maintaining assay robustness, harmonizing data acquisition, and supporting regulatory compliance under CAP/CLIA or GCLP or ISO15189 guidelines (1, 2).

The goal of ongoing between-instrument variability assessment is to ensure that a given assay produces comparable and reliable results when performed on two or more instruments used for sample acquisition. This evaluation helps identify potential sources of variability—such as differences in laser alignment, detector sensitivity, fluidics performance, or software compensation—and determines whether any observed differences fall within the assay's predefined acceptable limits. By demonstrating equivalency across instruments, laboratories can maintain consistent assay results, accurately monitor biomarker expression, and ensure equivalence in patient testing regardless of which instrument is used.

SCOPE

The aim of this module is to provide information, practical guidance, and examples related to the periodic practice of between-instrument variability assessment in clinical flow cytometry laboratories that use more than one flow cytometer to analyze the same assay. This module does not address the instrument-to-instrument precision studies required for the validation of new assays. Periodic between-instrument variability assessment must be performed in accordance with local regulatory requirements (for example, CAP requires this assessment at least twice per year). Beyond mandated requirements, each laboratory should determine whether additional between-instrument variability assessment is necessary following instrument repairs or other events that may impact assay performance.

Note: Assay settings established on the primary instrument should be transferred to additional instruments according to the specifications of each instrument model. Detailed procedure for transferring assay settings between instruments (whether identical or different models) is beyond the scope of this module.

COMPARISON TABLES AND ACCEPTANCE CRITERIA

Experimental requirements for between-instrument variability assessment depend on the assay. In this module, we describe examples relevant to both diagnostic applications of flow cytometry (e.g., leukemia/lymphoma evaluation) and enumeration assays (e.g., lymphocyte subset phenotyping). Traditionally, these categories have been labeled as *qualitative* for diagnostic tests, and *quantitative* or *semi-/quasi-quantitative* for enumeration assays. However, leukemia and lymphoma evaluations are almost never reported without accompanying numerical data, and lymphocyte subset assays are not strictly quantitative in the absolute sense. Therefore, we opted to use the more accurate descriptive terms **"diagnostic"** and **"enumeration"** throughout this document. Other types of assays, such as PNH evaluation by flow cytometry, or neutrophil oxidative activity assessment can be treated like diagnostic or enumeration assays, respectively.

DIAGNOSTIC TESTS

For flow cytometric tests that are used to establish the presence or absence of disease, a diagnostic accuracy 2x2 table is used. The results (as positive or negative for disease) from the two cytometers are entered as shown in Table 1. Guidelines for choosing sample size are available from several resources (3), but it must be approved by the medical director. For the diagnostic testing all the reportables used for diagnosis should be evaluated (and in agreement with local regulatory requirements).

TABLE 1: Agreement Analysis of Qualitative Disease Detection Between Two Cytometers (2×2 Contingency Table).

		Cytometer 1		**Total**
		Negative	Positive	
Cytometer 2	Negative	3 (TN)	0 (FN)	**3**
	Positive	0 (FP)	3 (TP)	**3**
Total		**3**	**3**	**6**

TP (True Positive) is the number of samples with the disease which are correctly identified.
TN (True Negative) is the number of samples without the disease which are correctly identified.
FP (False Positive) is the number of samples without the disease which are incorrectly identified (test positive)
FN (False Negative) is the number of samples with the disease which are incorrectly identified (test negative)

If the laboratory includes numerical values in the leukemia and lymphoma evaluation report, and such values may be used in making clinical decisions, it becomes incumbent on the lab to compare the data between the instruments used. The choice of the populations to compare, and the metric used to evaluate equivalence often depends on the assay, the population compared, the clinical relevance, and ultimately the decision of the laboratory medical director. An example of such a comparison is shown in Table 2.

The establishment of acceptance criteria should account for differences in population sizes and assign appropriate allowable deviations. Acceptance criteria may be based on percentage

differences; however, for populations smaller than 0.1%, where percentage differences can be misleading, an absolute difference is recommended. Whatever method is used for these calculations, they should be fit-for-purpose and comply with the acceptance criteria approved by the laboratory director.

TABLE 2: Example of a comparison of cell populations between three instruments (relative % gated).

	Cytometer 1	Cytometer 2	Cytometer 3	Average	SD	Δ (1-2)	Δ (2-3)	Δ (1-3)
B cells	4.4	4.2	4.6	4.40	0.16	0.2	-0.4	-0.2
Kappa	66	65	66	65.67	0.47	1	-1	0
Lambda	30	33	31	31.33	1.25	-3	2	-1
k:l ratio	2.2	2.0	2.1	2.10	0.10	0.2	-0.2	0.1
CD5+ B cells	<0.1	<0.1	<0.1	<0.1	N/A	Acceptable*	Acceptable*	Acceptable*
CD10+ B cells	0.2	0.2	0.2	0.20	0.00	0.0	0.0	0.0
Plasma cells	1.1	1.2	1.4	1.23	0.12	-0.1	-0.2	-0.3
CD3+ T cells	72	77	76	75.00	1.16	-5.0	1.0	-4.0
CD4+CD8-	60	58	57	58.33	1.25	2.0	1.0	3.0
CD4-CD8+	34	38	32	34.67	1.49	-4.0	6.0	2.0
CD4+CD8+	0.9	0.7	0.8	0.80	0.08	0.2	-0.1	0.1
CD4- CD8-	2.7	2.4	2.9	2.67	0.21	0.3	-0.5	-0.2
CD4:CD8 ratio	1.76	1.53	1.78	1.69	0.12	0.2	-0.3	0.0
CD57+ T cells	3.5	3.8	3.2	3.50	0.24	-0.3	0.6	0.3
CD56+ T cells	5.8	5.9	5.5	5.73	0.17	-0.1	0.4	0.3
TCRαβ+	95	94	92	93.67	1.25	1.0	2.0	3.0
TCRγδ+	4.5	4.8	5.2	4.83	0.29	-0.3	-0.4	-0.7
NK cells	8.8	8.5	9	8.77	0.21	0.3	-0.5	-0.2
Monocytes	11.5	8.9	9.9	10.10	1.07	2.6	-1.0	1.6
CD34+ blasts	2.3	2.3	2.3	2.30	0.00	0.0	0.0	0.0
CD117+ cells	1.8	1.6	1.6	1.67	0.09	0.2	0.0	0.2

SD: standard deviation; $\Delta\,(1{-}2)$: difference between instruments 1 and 2; $\Delta\,(2{-}3)$: difference between instruments 2 and 3; $\Delta\,(1{-}3)$: difference between instruments 1 and 3.
Acceptance criteria were set at a SD ≤ 2 and an absolute difference (Δ) ≤ 10%.

Evaluation of inter-instrument comparability for rare event analysis assays, such as MRD, should follow the same structure (3). A B-ALL specimen with 0.01% MRD positive result could be set up in duplicate and run on at least 2 instruments. For myeloma MRD, create a mock sample by spiking a positive case into a negative (background) to create a positive sample (at/near the assay's sensitivity level), and acquiring it on multiple instruments.

ENUMERATION ASSAYS

For laboratories that perform enumeration assays, such as lymphocyte subset enumeration (TBNK), and CD34+ hematopoietic stem cell (HSC) enumeration, between-instrument variability asessment must evaluate both the population percentages and the absolute cell counts. It is recommended to place emphasis on comparisons around specific clinical cutoffs (for example 200-400 cell/μL for CD4 and 10 cells/μL for CD34). These are set based on the

context of use, manufacturer's performance claims, data from the verification/validation of the assay, external quality assessment, and medical director discretion. The acceptance criteria do not need to be the same for all populations; however, they should be set distinctively for the population percentages, and the respective absolute cell counts. As shown in the example (Table 3), %CV has the range from 0.45 to 9.3, depending on the population. Acceptance criteria set for each population should be reported. Between-instrument variability can be assessed by measuring the correlation between a comparative instrument and a test instrument and by calculating bias at relevant medical decision point(s). Statistical approaches, such as Deming regression with predefined allowable bias (e.g., ±10% or ±X cells/µL), may be used to support this assessment.

TABLE 3: Comparison of lymphocyte subset populations between instruments (relative % of lymphocytes and absolute count (cells/µL)). The same processed sample is acquired on the different cytometers. Acceptance criteria were set at %CV ≤10.

Sample 1	**CD3 %**	**CD3 count**	**CD4 %**	**CD4 count**	**CD8 %**	**CD8 count**	**CD19 %**	**CD19 count**	**NK %**	**NK count**
Cytometer1	55.17	692	10.94	137	36.62	460	19.59	246	22.76	286
Cytometer2	56.3	694	11.17	138	36.71	453	18.04	222	23.34	288
Cytometer3	55.47	622	11.69	131	35.44	398	19.46	218	22.15	248
mean	**55.65**	**669.33**	**11.27**	**135.33**	**36.26**	**437.00**	**19.03**	**228.67**	**22.75**	**274.00**
SD	*0.59*	*41.00*	*0.38*	*3.79*	*0.71*	*33.96*	*0.86*	*15.14*	*0.60*	*22.54*
%CV	***1.05***	***6.13***	***3.41***	***2.80***	***1.95***	***7.77***	***4.52***	***6.62***	***2.62***	***8.23***
Pass/Fail	***Pass***	***Pass***	***Pass***	***Pass***	***Pass***	***Pass***	***Pass***	***Pass***	***Pass***	***Pass***
Sample2	**CD3 %**	**CD3 count**	**CD4 %**	**CD4 count**	**CD8 %**	**CD8 count**	**CD19 %**	**CD19 count**	**NK %**	**NK count**
Cytometer1	71.93	1006	49.5	692	20.27	284	13.79	193	12.34	173
Cytometer2	71.93	1038	48.41	699	20.87	301	14.35	207	11.66	168
Cytometer3	72.49	1002	47.6	658	21.99	304	15.47	214	10.75	149
mean	**72.12**	**1015.33**	**48.50**	**683.00**	**21.04**	**296.33**	**14.54**	**204.67**	**11.58**	**163.33**
SD	*0.32*	*19.73*	*0.95*	*21.93*	*0.87*	*10.79*	*0.86*	*10.69*	*0.80*	*12.66*
%CV	***0.45***	***1.94***	***1.97***	***3.21***	***4.15***	***3.64***	***5.88***	***5.22***	***6.89***	***7.75***
Pass/Fail	***Pass***	***Pass***	***Pass***	***Pass***	***Pass***	***Pass***	***Pass***	***Pass***	***Pass***	***Pass***
Sample3	**CD3 %**	**CD3 count**	**CD4 %**	**CD4 count**	**CD8 %**	**CD8 count**	**CD19 %**	**CD19 count**	**NK %**	**NK count**
Cytometer1	92.36	2065	52.13	1166	38.5	861	1.59	36	5.61	125
Cytometer2	91.92	1990	53.39	1156	36.87	798	1.43	31	6.18	134
Cytometer3	91.05	1947	52.6	1125	36.38	778	1.57	34	6.76	144
mean	**91.78**	**2000.67**	**52.71**	**1149.00**	**37.25**	**812.33**	**1.53**	**33.67**	**6.18**	**134.33**
SD	*0.67*	*59.72*	*0.64*	*21.38*	*1.11*	*43.32*	*0.09*	*2.52*	*0.58*	*9.50*
%CV	***0.73***	***2.98***	***1.21***	***1.86***	***2.98***	***5.33***	***5.70***	***7.48***	***9.30***	***7.08***
Pass/Fail	***Pass***	***Pass***	***Pass***	***Pass***	***Pass***	***Pass***	***Pass***	***Pass***	***Pass***	***Pass***

Tip: Some laboratories acquire their external quality assurance (proficiency testing) samples on all cytometers - this provides data and to create a regular system for monitoring between-instrument concordance.

AUTOMATED SAMPLE PREPARATION SYSTEMS

Like cytometers, automated sample preparation systems should be included in the laboratory's between-instrument variability assessment procedures. The comparison method and acceptability criteria are similar to what is described for cytometers, above. If samples are routinely processed manually in addition to using automated sample preparation instruments, procedures need to be implemented to compare if manual set-up results are comparable to those of the automated sample preparation system(s). Table 4 shows an example of such a comparison. Similarly, if MFI is used for reporting or interpretation (e.g., CD64 expression on neutrophils and HLA-DR expression on monocytes), it should be included in the comparison, and the acceptance criteria should comply with those established by the laboratory director.

TABLE 4: Comparison of cell populations between four automated sample preparation systems and manual preparation (relative % population gated shown).
Acceptance criteria is set at %CV ≤10.

	Manual	Auto1	Auto2	Auto3	Auto4	Avg	SD	%CV	Pass/Fail
B cells	12.0	12	11.4	10.8	10.2	11.3	0.70	6.2%	**Pass**
Kappa+	6.6	6.0	6.0	5.4	5.4	5.9	0.45	7.6%	**Pass**
Lambda+	4.8	4.8	4.8	4.8	4.2	4.7	0.24	5.1%	**Pass**
K:L ratio	1.5	1.3	1.3	1.2	1.3	1.3	0.10	7.4%	**Pass**
CD5+ B cells	<0.1	<0.1	<0.1	<0.1	<0.1	*Acceptable			**Pass**
CD10+ B cells	1.2	1.2	1.2	1.1	1.3	1.2	0.06	5.3%	**Pass**
CD38++ PC	<0.1	<0.1	<0.1	<0.1	<0.1	*Acceptable			**Pass**
CD3+ T cells	75.6	69	82.8	74.4	71.4	74.6	4.69	6.3%	**Pass**
CD4+CD8-	48.0	43.2	52.8	47.4	43.8	47.0	3.45	7.3%	**Pass**
CD4-CD8+	21	20.4	24.0	20.4	21.0	21.4	1.35	6.3%	**Pass**
CD4+CD8+	1.8	1.8	1.8	1.8	1.8	1.8	0.00	0.0%	**Pass**
CD4-CD8-	4.2	4.2	4.8	4.2	4.8	4.4	0.29	6.6%	**Pass**
CD4:CD8 Ratio	2.3	2.1	2.2	2.3	2.1	2.2	0.09	4.1%	**Pass**
CD57+	4.2	4.2	4.2	4.2	3.6	4.1	0.24	5.9%	**Pass**
CD56+	1.8	1.8	1.8	1.8	1.8	1.8	0.00	0.0%	**Pass**
TCRαβ+	72.0	65.4	79.2	70.2	67.2	70.8	4.78	6.8%	**Pass**
TCRγδ+	3.6	3.6	3.6	3.6	4.2	3.7	0.24	6.5%	**Pass**
NK cells	13.8	13.8	13.2	13.2	13.8	13.6	0.29	2.2%	**Pass**
CD14+ monos	10	10.6	10.2	10.0	9.6	10.1	0.32	3.2%	**Pass**
CD34+ blasts	3.6	3.6	3.6	3.6	3.6	3.6	0.00	0.0%	**Pass**

CLINICAL TRIALS

Assays for clinical trials are in general validated in two instruments. During the assay validation, between-instrument variability assessment is determined with 3-4 samples processed by the

same operator in two instruments. % difference between instruments is calculated and acceptance criteria could be set up to 20% or based on the laboratory director's discretion (there is no published guideline for acceptance criteria (4)). All the reportables within acceptance criteria for Repeatability are assessed for between-instrument variability during the assay validation (4,5). Rare populations (<5% of parent or <100 events) are not considered for between instruments as high variability is expected. Any specific population below the lower limit of quantitation (LLoQ) is also not considered in the assessment. If any rare population is critical for a particular clinical trial, then laboratory director should evaluate and identify other means of testing between-instrument variability (using patient samples, cell lines or control materials).

EXPERIMENT SETUP

One instrument could be defined as "Comparative" instrument, while other instruments are considered "Test" instruments. Test instruments are then compared to the comparative instrument for between-instrument variability assessment.

Number of samples

Sample size selection should be determined by the laboratory and approved by the Medical Director. Between 1 and 3 samples should be tested to assess between-instrument variability. The sample should include positive and negative populations, as appropriate for the assay.

Sample Types

Samples for comparability studies should reflect sample types tested. Residual patient samples are preferred and can be pooled if needed. Using clinical samples can avoid matrix effects from fixed or simulated samples. However, in the event these are unavailable, alternative samples such as reference materials can also be considered.

Frequency of testing

The frequency of between-instrument variability assessment depends on local regulatory requirements. For laboratories under CAP certification, instruments are checked against each other at least twice in a calendar year for comparability of results.

DATA ANALYSIS

Reportables to compare

For diagnostic testing any of the reportables used to aid in the diagnosis can be evaluated (and in agreement with local regulatory requirements).

For an assay validated for clinical trial, all the primary reportables which are within acceptance criteria for repeatability are assessed for between-instrument variability.

Statistical calculations

Percentage difference is calculated as follows for each reportable:

$$\%\ \text{difference} = \frac{(v\text{alue from Comparative instrument} - \text{value from Test instrument})}{\text{value from Comparative instrument}} \times 100$$

Reportable refers to parameters reported for the assay, for example, B cells (% of lymphocytes).

For laboratories assessing more than two instruments for comparability, the Mean, Standard Deviation (SD), and Coefficient of Variation (%CV) are calculated across all instruments for all reportable parameters. Data are accepted if individual instrument values fall within an acceptable range (e.g., ±20% or ±2SD) of the Mean. Additionally, the %CV is used to assess overall performance, with an acceptance threshold typically of ≤10%.

SUMMARY

Between-instrument variability assessment in clinical flow cytometry ensures that an assay produces consistent, reliable, and clinically interchangeable data across all cytometers used within a laboratory or clinical trial. These assessments should be performed using representative patient samples or validated reference materials when adding an assay to a new instrument, and semiannually per CAP requirements. Each laboratory should clearly define its frequency of testing, reportables to compare, and acceptance criteria within its standard operating procedure to ensure harmonized performance across all instruments.
The key considerations differ by assay type:

- Diagnostic assays require qualitative agreement (positive vs. negative) and, when reported, can include population percentages.
- Enumeration assays require evaluation of both relative frequencies and absolute counts, with particular attention to clinically meaningful thresholds.
- Clinical trial assays rely on predefined acceptance criteria, typically ≤20% difference, and may include operator-specific comparisons.

Ultimately, between-instrument variability comparison provides documented evidence that results are stable, reproducible, and suitable for clinical interpretation.

REFERENCES (additional references listed below serve as informational resource).

1. https://documents.cap.org/documents/CAP15189_Accreditation_ProcessChecklist_web.pdf
2. https://www.cap.org/laboratory-improvement/accreditation/accreditation-checklists
3. Monaghan SA, Eck S, Bunting S, Dong XX, Durso RJ, Gonneau C, Hays A, Illingworth A, League SC, Linskens E, McCausland M, McCloskey TW, Rolf N, Shi M, Wallace PK, Litwin V, Kern W, Deeb G, Nash V, Olteanu H. Flow cytometry assay modifications: Recommendations for method validation based on CLSI H62 guidelines. Cytometry B Clin Cytom. 2025 May;108(3):252-266. doi: 10.1002/cyto.b.22202. Epub 2024 Aug 21. PMID: 39165120.
4. Selliah, N., Nash, V., Eck, S., Green, C., Oldaker, T., Stewart, J., Vitaliti, A., & Litwin, V. Flow cytometry method validation protocols. Current Protocols, 2023:3, e868. doi: 10.1002/cpz1.868.
5. CLSI (Ed.). Validation of assays performed by flow cytometry, 1st ed. CLSI document H62. Wayne, PA: Clinical Laboratory Standards Institute, (2021).
6. Clinical and Laboratory Standards Institute. Verification of Comparability of Patient Results within One Healthcare System: Approved Guideline (Interim Revision). CLSI document EP31-A-IR. Clinical and Laboratory Standards Institute, Wayne, PA; 2012.

7. Clinical and Laboratory Standards Institute (CLSI). Protocol for the Evaluation, Validation, and Implementation of Coagulometers: Approved Guideline. CLSI document H57-A (ISBN 1-56238-656-5) Clinical and Laboratory Standards Institute, 940 West Valley Road, Suite 1400, Wayne, Pennsylvania 19087-1898 USA, 2008.

Reviewed by:

1. Bruce Greig, MLS(ASCP) SCYM CM, Associate in Pathology, Microbiology, and Immunology (retired), Vanderbilt Medical Laboratories
2. Jolene Cardinali, MLS(ASCP), Hartford Hospital
3. Natalia Kokorina, MD, Spring Analytics
4. Ulrika Johansson, PhD FRCpath, University Hospitals Bristol and Weston NHS Foundation Trust

For any questions on this module or any other suggestions, please email info@cytometry.org