Validation of Cell-based Fluorescence Assays: Practice Guidelines from the ICSH and ICCS – Part III – Analytical Issues

Shabnam Tanqri,[1*] Horacio Vall[2] David Kaplan,[3] Bob Hoffman,[4]
Norman Purvis,[5] Anna Porwit,[6] Ben Hunsberger,[7]
T. Vincent Shankey[8]; on behalf of ICSH/ICCS Working Group[1]

[1]Biogen Idec, San Diego, CA, USA

[2]USLabs/LabCorp, Irvine, CA, USA

[3]Case Western University, Cleveland, OH, USA

[4]Livermore, CA, USA

[5]Nodality, South San Francisco, CA, USA

[6]Toronto General Hospital, Toronto, ON, Canada

[7]Verity Software House, Topsham, ME, USA

[8]Shankey Biotechnology Consulting, Miami, FL, USA

Clinical diagnostic assays, may be classified as quantitative, quasi-quantitative or qualitative. The assay's description should state what the assay needs to accomplish (intended use or purpose) and what it is not intended to achieve. The type(s) of samples (whole blood, peripheral blood mononuclear cells (PBMC), bone marrow, bone marrow mononuclear cells (BMMC), tissue, fine needle aspirate, fluid, etc.), instrument platform for use and anticoagulant restrictions should be fully validated for stability requirements and specified. When applicable, assay sensitivity and specificity should be fully validated and reported; these performance criteria will dictate the number and complexity of specimen samples required for validation. Assay processing and staining conditions (lyse/wash/fix/perm, stain pre or post, time and temperature, sample stability, etc.) should be described in detail and fully validated. © 2013 International Clinical Cytometry Society

Key terms: Flow Cytometry; practice guideline; standardization; regulatory science; laboratory hematology

INTRODUCTION

Flow cytometers used for tests developed in an individual laboratory must be monitored for consistency of critical performance factors (1). If a LDT is performed on more than one flow cytometer in the laboratory, each instrument must be qualified to perform the test and shown to provide the same result. Although flow cytometers usually have robust performance, the latter can differ between instruments, even of the same model, especially on the extremes of measurement scales. Performance characteristics such as precision and fluorescence sensitivity that can change rapidly due to fluidic problems and, that in turn, can affect alignment of the sample in the optical path, should be checked each day the instrument is used. This is typically achieved using stable bead mix-tures during the daily start-up routine for each instrument.

Assay calibration establishes a measurement scale that can be used to report results quantitatively. Examples of calibrated measurements are fluorescence intensity measurements, cell concentration, DNA content or antibodies bound per cell. If the calibration

material, usually stable particles, used to establish the scale does not provide units for reporting results (i.e. Molecules of Equivalent Soluble Fluorochrome (MESF) for fluorescence intensity), the process is then termed standardization rather than calibration, but is often the only practical approach available.

If the laboratory has only one flow cytometer or only one designated flow cytometer will be used for the test, the absolute performance level of the instrument is not essential to characterize. However, objectively monitoring the instrument performance is essential to insure that consistent results will be obtained. If the test will be performed on two or more instruments, then knowing the relative performance of each instrument can be helpful, especially if the test makes measurements near the limit of performance or detection.

DEFINITIONS
Calibration

Process of adjusting an instrument so that the analytical result is accurately expressed in some physical unit of measure.

Calibrator

Material that has been manufactured or assayed to have known, measured values of one or more characteristics. The assayed values are provided with the material. Fluorescent manufactured particles can be assayed for diameter or for the amount of fluorescence they produce. A practical measure of particle fluorescence is the number of fluorochrome molecules in solution that produce the same amount of fluorescence as one bead (MESF).

Control Particle or Material

Material (e.g., sample of manufactured particles or autologous/allogeneic cell populations) that gives reproducible and predictable results when analyzed. Particles used to set up a flow cytometer can be used as a control even if they do not have an assayed value assigned to a physical characteristic. Controls can also be used to monitor the stability of an instrument and determine whether it is within calibration. A calibrator can be used as a control material, but a control material does not have to have an assigned value for a specific characteristic.

MESF (Molecules of Equivalent Soluble Fluorochrome)

Measure of particle fluorescence in which the assigned value to the signal from a fluorescent particle is equal to that from a known number of molecules in solution. This is a practical measure because a known concentration of particles can be compared directly with a solution of fluorochrome in a spectrofluorometer.

Precision or Reproducibility

Degree to which repeated measurements of the same thing agree with each other. In flow cytometry, precision of a measurement is estimated by the CV obtained when measuring replicates of a sample of particles (biological or nonbiological) with very uniform characteristics. Each flow cytometric measurement is typically the mean or median of 1,000–50,000 individual measurements; this is in contrast to typical fluorometric or spectrophotometric assay measurements, which are only a single physical measurement of the entire assay mixture.

Resolution

Degree to which a flow cytometer measurement parameter can distinguish two populations in a mixture of particles that differ in mean signal intensity. Fluorescence sensitivity can be considered a special case of fluorescence resolution for which the signals are very dim and at the lower limit of detection. Note that the resolution will appear different when data are acquired and/or displayed on a logarithmic rather than linear intensity scale. Depending on the maximum number of channels into which the signal intensity is acquired (e.g., 256 or 1,024 channels), a logarithmic display of the data may not have sufficient resolution to display populations that can actually be resolved by the instrument using a linear intensity scale.

Standard

1. noun. a. Acknowledged measure of comparison for quantitative or qualitative value. b. Something recognized as correct by common consent or by those most competent to decide. 2. adjective. a. Serving as a standard of measurement or value. b. Commonly used and accepted as an authority.

Standardize: verb. a. Cause to conform to a given standard. b. Cause to be without variation.

Instrument Performance Characterization and Standardization

Light scatter

Light scattering from cells depends on many factors including size, shape, internal microstructure, refractive index of the cells and surrounding fluid and the angles over which scatter is measured by the detectors. There is some truth in the general statement that forward scatter is a measure of particle size and side or right angle scatter is a measure of internal complexity, but these generalizations can also be very misleading in practice. The engineering design for light scatter measurements differs among flow cytometer manufacturers (and even between models) and results can vary considerably. The size of bead used as a standard for reproducibly setting the light scatter detector gains may not have a close relation to the size of cells being measured. The most useful standard for initially setting light scatter gains will be the actual type of cell used for the analysis, e.g., red cells, leukocytes or platelets.

The resolution of cell populations by light scatter is affected by the alignment of the sample stream, the excitation light and the optical system used to collect the scattered light. Resolution, particularly for small particles from background, is affected by purity of the excitation wavelength, cleanliness of the flow cell, consistency of the flow rate, and the debris-free sheath fluid composition.
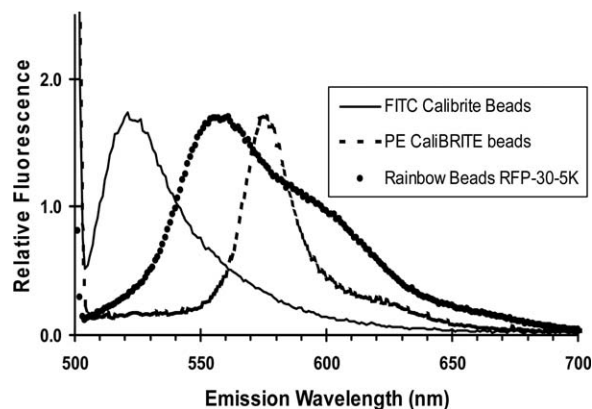
Fig. 1. Emission spectra of surface stained beads (FITC and PE Calibrite beads from BDB) and multi-fluorophore hard dyed beads (Rainbow beads from Spherotech).

A subtler source of scatter background for small particles is a difference in refractive indexes of the sample fluid and sheath. A similar issue may be encountered when the sample fluid has high protein concentration and the sheath fluid is protein-free.

Fluorescence

Fluorescent beads are either stained on the surface with a specific fluorochrome used in flow cytometry or internally with one or more fluorochromes. Internally stained beads contain fluorophores that are not water-soluble and are referred to as "hard dyed." Figure 1 shows emission spectra of beads surface stained with two dyes commonly used for immunofluorescence, fluorescein (FITC) and phycoerythrin (PE), as well as the emission spectrum from a hard dyed bead stained with multiple fluorophores. Since the emission spectra and excitation spectra of surface stained and hard dyed beads almost never match, hard dyed beads must be used with caution when used to standardize flow cytometer settings for assays using other fluorochromes as active reagent labels.

The advantage of hard dyed beads is their stability, which is much greater than that of surface stained beads. The stability of surface stained beads is improved considerably if they are freeze-dried and then kept at refrigerated temperature (2–8°C). Since the fluorophore on surface stained beads is exposed to the suspension buffer, the fluorescence emission may be affected by the pH, salt concentration and other factors in the buffer. It is always a good idea to suspend surface stained beads in the same buffer as the cells they are intended to standardize. This is especially important for FITC due to its pH dependent fluorescence.

Beads are also available with capture monoclonal antibodies because they have an anti-IgG or Fc receptor on the surface. These beads can be stained with fluorescent antibodies for setting compensation, and in some cases are used as calibrators for antibody binding.

Fixed or stabilized cells or nuclei are also available for staining with user provided reagents. Fixed nuclei are used as standards for DNA measurements or tools for lin-

earity of fluorescence verification (2,3). Stabilized cells can potentially be used as controls and calibrators (4–19). One approach being developed using CD4 T cells for immunofluorescence (see below), which assumes the antigen expression is truly stable across individuals and stability of the product is well validated (18,19).

Fluorescence Standards and Calibrators

There are currently no assigned fluorescent standard particles available from either the US National Institute of Standards and Technology or the National Institute of Biological Standards and Control (NIBSC, UK). Particles with assigned values are available from several commercial bead companies. However, it should be stressed that these are not uniformly assigned standards, nor certified by any recognized metrology organization. At present, fluorescence intensity is usually assigned in MESF units for beads that are surface stained with specific fluorophores such as FITC. Hard dyed beads that have assigned intensity values should not use the MESF unit if, as typically the case, the spectrum of the beads does not match the emission spectrum of the fluorochrome being calibrated. If intensity units are specified over a defined wavelength range, for a specific instrument model, then it is appropriate to assign calibration values to hard dyed beads. For example, Becton Dickinson (San Jose, CA) assigns intensity values in assigned arbitrary units (Arbitrary BD units or ABD) to Cytometry Setup and Tracking (CS&T) beads used to standardize setup of recent models of their flow cytometers. Bangs Laboratories (Fishers, IN) provides beads with FITC and PE with assigned spectrally matched MESF values. Spherotech (Lake Forest, IL) does not provide spectral region information, but distinguishes the MEF intensity units assigned to Rainbow beads from MESF units that are appropriate only for spectrally matched, surface stained beads. Bead suppliers, who do assign MESF units to spectrally matched beads, do not have a reference bead from an authoritative source, such as NIST. Variation in the MESF assignments among the bead suppliers can therefore be expected. It is always possible, however to cross-calibrate the beads from one supplier to the MESF value assigned by another supplier, which would be valid only between those specific two batches.

An alternative generalized fluorescence intensity unit has been proposed by investigators at NIST to supplement the MESF unit (20). The Equivalent Reference Fluorophore (ERF) unit does not use the same fluorophore for reference material as the fluorophore used to stain beads. This allows one reference fluorophore to calibrate a large number of different fluorophores on beads. For example, the Nile Red dye could in theory be used as a fluorescence reference for PE. When ERF units are used, it is essential to identify the excitation wavelength and emission band over which the units are assigned. For example, if Nile Red is used to calibrate PE stained beads using excitation at 488 nm and an emission band of 560- 590 nm, the assigned units would also state that the ERF value is only correct in these conditions. The
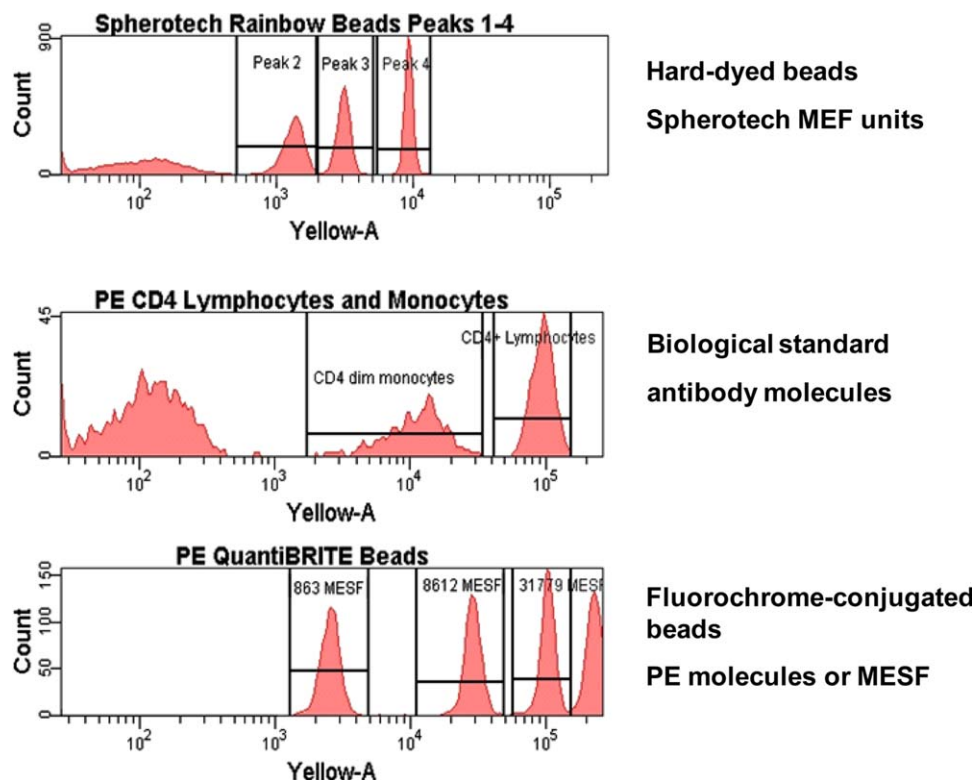
Fɪɢ. 2. Cross-calibration of different standards, each analyzed with the same gain setting on a flow cytometer. Comparison of mean or median fluorescent intensity for different samples analyzed on a flow cytometer with the same gain setting allows for comparison of the various bead types, along with a biologic standard under development for CD4 expression. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

ERF unit is an extension of the MEF unit and can be applied to hard dyed or surface stained beads. Unfortunately the variance between instrument models of the same type is quite high with CVs >15%, indicating the ERF system needs more improvement before it can be used in a practical sense (20).

Cross Calibration of Different Standards

It is also appropriate to cross-calibrate values of surface stained bead calibrators to hard dyed beads on a single flow cytometer. As long as the filters and other spectrally sensitive components in the instrument do not change, the intensity units cross-calibrated between the two bead types would be valid. There may also be situations where a biological standard, should a stabilized product be developed, is the most appropriate (or only) standard to which other particles should be referenced. Figure 2 illustrates the process of cross-calibration in which the mean or median fluorescence intensities of different bead types are compared after running the samples with the same fluorescence gain on a flow cytometer.

Alignment and Resolution for Bright Fluorescence

To check alignment of the sample stream to the excitation and emission optics, bright, uniformly stained beads are used. The intrinsic CV of the beads should be less than 3% so that the contribution of the flow cytometer to total measured CV can be reliably determined.

Low sample flow rates are used for best resolution and DNA analysis, where fluorescence CVs of 3% or less should be obtained. At high sample flow rates (e.g., greater than 5,000 events per second) when the biological variability is high (e.g., immunofluorescence) a CV of ~5% is acceptable with the beads used for alignment.

Figure 3 shows an example of alignment characterization using uniformly stained beads. Both scatter and all fluorescence channels can be evaluated and monitored. The particle manufacturer should provide lot specifications for the CVs of various parameters, scatter and fluorescence, that can be expected when the particles are analyzed on a flow cytometer.

Linearity

Validation of linearity of fluorescence measurements is important for quantitative tests and can be even more critical with fluorescence compensation, especially across instrument platforms. Most modern flow cytometers used for clinical applications acquire high resolution and wide dynamic range linear digital data. Compensation among fluorescence detectors and display of data on a log scale are computed from the acquired digital linear values.

If compensation is set using fluorescence at one point on the scale and measurements on other parts of the scale not in a proportional relationship, the calculated compensated values will be incorrect. It should be pointed out that a small error in proportionality can
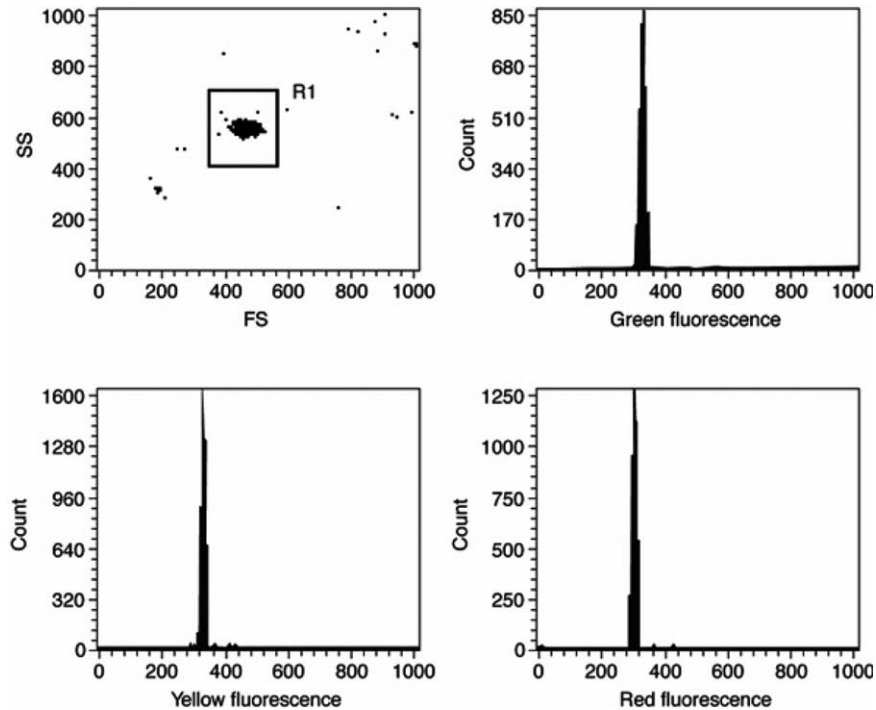
Fɪɢ. 3. Example of alignment characterization with uniformly stained hard-dyed beads. Both scatter and fluorescence channels can be evaluated and monitored over time.

cause a large absolute error in the calculated compensation value. This is why cross-instrument correlation of quantitative data, when compensation is utilized as part of data collection and/or analysis, can be very problematic for inter-instrument imprecision and bias.

To critically test linearity, it is best to monitor the intensity ratio of two different bead populations at various points on the fluorescence readout scale (3). If the response is proportional, the ratio should be the same at all values on the scale, by varying the PMT voltage and monitoring the median (preferred over mean) fluorescence intensity (MFI) ratio for the two different intensity beads (4). The method reliably measures any deviation from proportionality due to the electronics and

data acquisition system that follow the PMT. Figure 4 shows a plot of two beads over a measurement range of more than 3 decades for a flow cytometer.

Linearity can also be estimated from the intensities assigned to each population in a multi-bead set by plotting measured MFI vs. assigned value in a log scale. If the response is proportional, the slope of the line in the log-log plot should be exactly 1. For control or tracking purposes, measuring and monitoring the MFI of the various populations in the multi-bead set is a good way to determine if linearity is changing over time. Figure 5 shows a good example of how such beads provide information about the linearity. Broadening of the dimmer bead populations (as shown in Fig. 5) is not due to
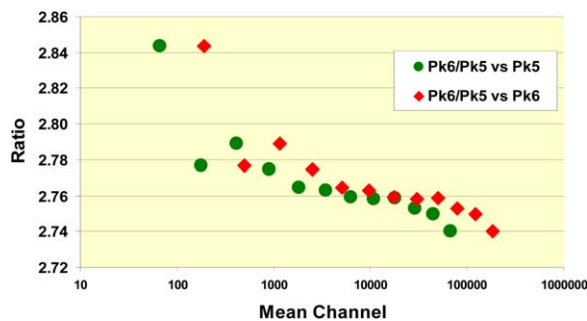

Fɪɢ. 4. Linearity validation determined by a ratio of brighter to dimmer MFI of populations 5 and 6 of the Spherotech Rainbow beads RCP-30-5A. The ratio of the two beads when plotted vs. channel number would show deviation from linearity across the measurement range. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]
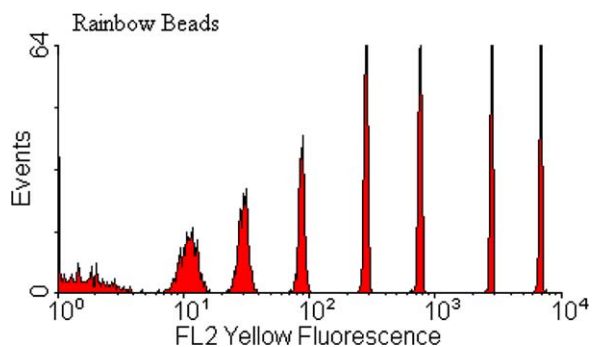

Fɪɢ. 5. Fluorescence histogram of Spherotech 8-peak Rainbow hard dyed beads. The bead positions can be monitored on a day-to-day basis to monitor instrument performance. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]
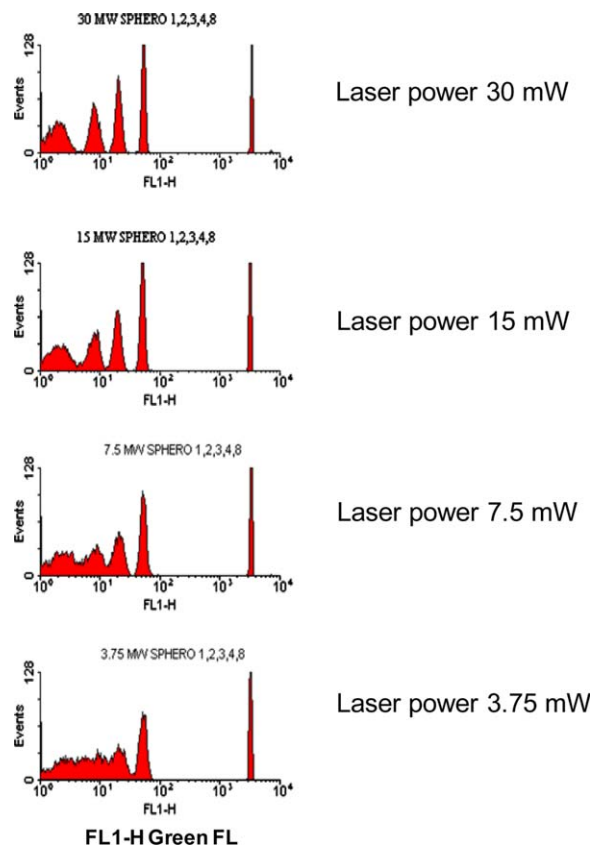
FIG. 6. Illustration of dim fluorescence bead populations broadening at lower laser excitation intensity using Spherotech Rainbow beads. The dimmest population is an unstained, autofluorescent bead. The other bead populations are stained uniformly with different amounts of fluorophores. PMT gain was changed to put the brightest bead at the same MFI at each laser power. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

increased variability in particle staining, but rather to less light being collected by the PMT, as described in the next section.

Sensitivity or Resolution of Dimly Fluorescent Populations

In practical terms, relative sensitivity for fluorescence with flow cytometers is measured by the ability to resolve dimly stained populations. At least three instrument factors affect the ability to resolve dimly stained populations: 1) fluorescence detection efficiency; 2) background light; 3) electronic noise within the instrument. Fluorescence detection efficiency and background light determine the number of signal and background photoelectrons that are amplified by the PMT, and the noise results in broadening of the populations in a fluorescence histogram. In physical terms, the noise is due to counting statistics. If 100 photoelectrons on average are generated when light strikes the photosensitive surface of a PMT, then the signal is 100 and the variation in the measured values or standard deviation is proportional to the square root of 100 or 10. The CV in this case would be $10/100 = 0.1$ or 10%. If background light

is also present, the photoelectron noise is also increased. For example, if the signal pulse from fluorescence generates on average 100 photoelectrons and the background above which that signal is measured generates on average 50 photoelectrons, the standard deviation of the measurement is contributed by both the signal and background photoelectrons and is the square root of $100 + 50$ or 12.25. The CV of the measurement in this case would be 12.25/100, or 0.1225 or 12.25%, if the only source of variation were photoelectron counting statistics. Electronic noise can also be present at a constant level, independent of the signal from the PMT, which can also contribute to the broadening of the measured signals in a histogram.

Figure 6 illustrates the broadening of dim populations that occurs when the fluorescence excitation signal is decreased. In this example, modifying the laser power changed the fluorescence signal from the sample. But at each laser power, the PMT gain was increased to place the brightest population at the same MFI. Even at the highest laser power, the dimmer populations get broader as the average signal intensity decreases.

Unfortunately, most flow cytometer manufacturers still specify fluorescence sensitivity in terms of the extrapolated MESF or MEF of an unidentified intercept or the calculated MESF or MEF of unstained, but autofluorescent beads. This approach is inappropriate for recent instruments that measure pulse area as the accurate and primary measure of cell fluorescence.

A better measure of sensitivity takes account of the broadness of an unstained bead and compares the broadness of the population in a histogram to the MFI of a stained bead. This concept is illustrated in Figure 7. A specific implementation of this concept has been used to define "Sensitivity" using a defined set of beads. In this case "Sensitivity" is defined as:

$$\text{Sensitivity} = \frac{\text{Positive MFI} - \text{Unstained MFI}}{2 \times \text{Unstained MFI SD}}$$

A quantitative characterization of detection sensitivity, background light and electronic noise provides a complete assessment of the factors that affect fluorescence sensitivity, but is probably not necessary when one or a few instruments are used for a lab-developed test.
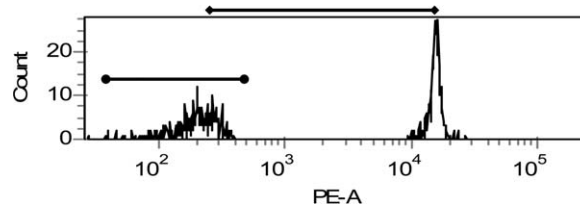

FIG. 7. Illustration of the concept of sensitivity based on the standard deviation of an unstained particle relative to a stained reference. The left sided peak is a true negative range and the ratio of the positive (right) side peak to the negative peak provides a measurement of the signal to noise.

Hoffman and Woods (5) provide a protocol for determining detection sensitivity and background light.

Quantitative Units

Several approaches have been developed to calibrate a flow cytometer for quantitative measures of antibodies bound per cell (6–17,21). The first uses beads that have been calibrated with a known number of fluorochrome molecules, a known number of mouse IgG antibodies, or a known binding capacity for anti-mouse IgG antibodies (16). The second approach uses cells that have a pre-determined average number of antibody binding sites per cell (17–19). A third utilizes software to compensate for lot-to-lot differences between fluorochrome labeled beads and reagents, using an index to quantitate molecular expression on cells (21).

Beads as Analyte Calibrators

In the first bead calibrator approach, beads labeled with a calibrated number of PE molecules per bead are used to calibrate the flow cytometer fluorescence scale in PE molecules (7) (Quantibrite beads from BD or Quantum PE beads from Bangs Laboratories). PE fluorescence from samples stained with a PE-conjugated antibody and analyzed with the same fluorescence detector setting as the beads can then be measured in terms of PE molecules per cell. Specially purified PE conjugates that have essentially all unconjugated and antibody conjugated to more than one PE molecule removed are essential in order to use this approach. Unconjugated antibody has a much higher affinity than when conjugated to PE, and hence a small fraction of unconjugated antibody in the staining reagent can therefore considerably reduce the fluorescence of the stained cells.

In the second bead calibrator approach (QIFI kit from DAKO (Glostrup, Denmark) or specifically targeted kits from BioCytex (Marseilles, France), beads are calibrated with a known number of mouse IgG antibodies, and a FITC-conjugated anti-mouse antibody prepared as part of the quantitation kit is used for indirect staining of both the calibrator beads and cells stained with mouse IgG antibodies (8,9). This approach is limited to indirect staining and therefore difficult to incorporate in multiparametric analyses that require staining with multiple different directly-conjugated mouse antibodies.

The third bead calibrator approach uses beads with calibrated binding capacity of mouse antibodies to anti-mouse antibodies on the bead surface (Quantum Simply Cellular/ QSC kit from Bangs Laboratories). The beads are incubated with the same antibody conjugate used to stain cells and are then washed to remove unbound antibody. The stained beads are used to calibrate the fluorescence scale, after which cells are analyzed. Since the same antibody is used on the beads and cells, the number of antibodies per cell can be determined from the number of anti-mouse antibodies bound based on the calibrated binding capacity of the beads (10–12). Concerns with the QSC approach include variation in measured ABC due to different antibody clones against the same antigen molecule and different ABC values have been observed with the same antibody conjugated with different fluorochromes (8,13). Variables of importance are to know whether your antibody is binding to one or two antigens and that Fc binding is not occurring. In addition, an apparent endless avidity of QSC beads that seem unable to saturate antibody binding may introduce further variability in quantitative fluorescence measurements (14).

Linear Scale to Biological Scale Transformation Using Biological Calibrators

The ultimate objective for quantitative fluorescence measurements with flow cytometry is to provide a calibration scheme such that the detected fluorescence signals in various fluorescence channels of a multicolor flow cytometer can be presented in terms of a number of analytes or antibodies bound per cell (ABC) (15). In the first cell calibrator approach, a biological standard such as a lymphocyte with a known number of antibody binding sites (e.g. CD4 binding sites) (4,5) can be used to translate the linear fluorescence intensity scale to an ABC scale. It is highly recommended that a single clone of the antibody amenable to labeling with different types of fluorophores associated with various fluorescence channels be used for the scale conversion. Assuming that different antibodies against different antigens have the same average fluorescence per bound antibody, a direct measure of antibodies bound per cell is obtained (16).

Considering the accessibility issues for fresh normal donor blood samples, potential cell reference standards, both cryopreserved, and lyophilized human peripheral blood mononuclear cells (PBMCs), could be practical alternatives that would allow for manufacturer antigen value assignments. The cryopreserved PBMCs are stored at −80 °C for a long period of time, i.e. a few years and are commercially available. Upon use, an optimal thawing protocol should be followed to ensure the cell viability. Due to the lack of a certified cryopreserved PBMC reference standard, individual users must evaluate the consistency of CD4 expression level on the cryopreserved PBMC with regard to the consensus value published for freshly prepared PBMC, ∼48,000 (7,17–19). Lyophilized PBMCs, on the other hand, are stored at 0-4 °C and are stable for at least a year. Because the lyophilization and fixation processes applied to PBMCs causes a decrease of the cell size and hence limits the accessibility of the binding sites, lower CD4 expression levels are reported for lyophilized PBMCs, i.e. CYTO-TROL control cells from Beckman Coulter (Miami, FL) (19). Another lyophilized PBMC pre-stained with anti-CD4 FITC, the first international reference reagent endorsed by the Expert Committee on Biological Standardization (ECBS) of WHO as a CD4+ Cell Counting Standard (WHO/BS/ 10.2153), will be soon available from the National Institute for Biological Standards and Control (NIBSC) in UK.

This reference cell standard provides not only the number of CD4+ cell count per unit volume, but also a mean CD4 expression level in terms of the equivalent fluorescein fluorescence value with an uncertainty estimate. However, use of the pre-stained calibration cells to determine antibodies bound will require use of a FITC CD4 conjugate with the same properties as the FITC conjugate used to stain the CD4 counting standard. These values were obtained through an international pilot study on "Quantification of Cells with Specific Phenotypic Characteristics" co-organized by NIBSC, UK, the National Institute of Standards and Technology (NIST) and Physikalisch-Technische Bundesanstalt (PTB), Germany under the Working Group on Bioanalysis of the Consultative Committee for Amount of Substance (CCQM/BAWG) in the pilot study CCQM-P102 (in progress).

An alternative method for analyte quantitation is described in a recently patented method integrating the use of calibration beads, fluorescently labeled monoclonal antibodies and software for data analysis (U.S. patent 8,116,984), (21,22). This approach is used for a commercial assay for CD64 quantitation on leukocytes and used in the determination of infection/sepsis (21,23,24). Bead value assignment is integrated into the software in a lot-specific manner, which allows for strict control over lot-to-lot variations in both bead and antibody fluorescence properties (F/P ratio, labeling efficiency, etc.). The bead can then be used for the creation of arbitrary index values or translated into ABC or molecules per cell units, should a universal standard or reference material again be made available. Possible additional advantages to this method are: 1) the use of internal cell populations for purposes of assay process control (positive and negative control cell populations) and; 2) the integration of the bead into the specimen analyzed by the flow cytometer; thus both the calibrator and controls are in the same listmode file of individual specimen results. Stipulating the use of uncompensated data collection, thereby minimizing any compensation bias, further reduces differences between instrument models or baseline offsets. Further, the ability to use instrument specific protocols for data analysis with the lot-specific software component of the method can remove any other inter-instrument bias. The method also stipulates to spectrally match fluorochromes between the calibration bead and the monoclonal antibody directed to the cell-based analyte of interest; this appears important given the apparent high imprecision (CV >15%) among users with the same instrument model using hard dyed, nonspectrally matched beads for quantification (20).

Cell Concentration

Cell concentration, also called the absolute count, can be measured in two different ways. Either the instrument can analyze a measured volume of sample or a known number or concentration of a reference particle can be added to the sample. In the latter case, concen-

tration will be given in terms of the ratio of sample cells to reference bead count. To test the accuracy of the cell concentration measurement, most bead suppliers provide bead suspensions at known concentrations that can be used to either calibrate or test the accuracy of the concentration measurement. CLSI H-42 provides details on the various methods for cell concentration measurement (25).

REAGENT PERFORMANCE
Optimization of Antibodies, Key Reagents, and Assay Systems

The successful design, development, validation and implementation of LDT are dependent on properly defining the assay instrumentation, reagents, procedures, and measurands. For flow cytometric assays, selecting which antibody combinations best delineate, distinguish and measure key differences within the target populations of interest and the number of simultaneously measured antibodies is a critical step. The numbers of lasers, spatially separated interrogation sites and available fluorochromes have significantly increased the number of signals that can be measured simultaneously. Once the antibody combinations are defined, an optimal fluorochrome must be selected for use with each antibody. One should objectively review the expected antigen expression on each of the target populations to be delineated and classify antigen density based on lowest to highest. Use of antibodies against intracellular and nuclear antigens in combination with surface markers also factors into fluorochrome optimization decisions.

Typically, one would choose a fluorochrome with the best quantum efficiency/yield as the antibody conjugate to identify the lowest antigen density so as to obtain the best possible signal to noise ratio. Fluorochromes with lower quantum efficiencies/yields should be chosen as the antibody conjugates used to identify the highest antigen densities. Population autofluorescence and spectral overlap from all fluorochromes must also be considered. These decisions may be influenced on which fluorochrome is dedicated to quantitating the measurand or analyte. Additionally, if absolute quantification of fluorescence intensity or reporting metrics related to changes in antigen expression is desired, the fluorochrome choices should be limited on those readouts to fluorochromes where appropriate Type IIb and IIIb fluorescence standards exist, specifically FITC and PE (16,21,26).

Reagent optimization is the process of selecting the best combination of reagents for use in the application and optimizing the performance of reagents within the design objectives and constraints of the assay. These parameters should be well defined within the assay design control documents. Reagent optimization must account for all assay design objectives and constraints including, but not limited to sample type, cell isolation method, lysing reagents, buffers, fixatives and permeabilization requirements. Selecting the appropriate reagents

to work together as a reagent system is often an iterative process evaluating different buffers, lysing reagents, isolation methods, fixatives, permeabilization reagents and antibodies to achieve optimal sensitivity, specificity and reproducibility. Design objectives often impose constraints on whether staining will be performed pre or post cell isolation (lysing), stabilization, fixation and/or permeabilization. It is important to realize that not all antibodies work equally well within all reagent systems and that not all antigen epitopes are equally available, expressed or recognized following cellular isolation, preparation, and particularly fixation and storage. Additionally, it is important to note that some fluorochromes are sensitive to reagent systems and stringency conditions (pH, temperature, detergents, fixatives, alcohols) used to permeabilize cells post-surface staining, thereby affecting any subsequent intracellular staining step.

Antibody and Fluorochrome Conjugate Optimization

Once the panel of antibodies is identified, the laboratory must determine which antibodies should be measured simultaneously. Often the same anchor-gating antibodies are used in every tube of a multi-tube panel, thereby allowing consistent gating strategies. In the diagnosis of leukemia/lymphoma, CD45 versus linear or log right angle light scatter (SSC), used in combination with maturation markers and population delineation markers, has proven very valuable for the diagnosis of various hematopoietic disorders and detection of minimal residual disease (27,28). Decisions related to how many antibodies and which antibodies should be measured simultaneously are often limited by instrument constraints (i.e. number of lasers and detectors), assay design control specifications and available fluorochrome conjugates. The choice of an anchor-gating strategy is dependent on the specific population of interest within the assay and given panel.

Laboratories should develop quantifiable antibody performance specifications, listing qualified and nonqualified antibody clones, vendor sources, fluorochrome conjugates and all appropriate acceptance criteria for antibody performance as it relates to the assay design objectives, specifications and constraints (25,28). The antibody performance specification should define the positive target populations and internal negative control populations contained within the assay samples, if possible; otherwise external controls must be defined. The antibody specification should define the appropriate positive and negative quality control samples (29), as well as the performance of any required gating reagents to be used, during antibody QC processes. Gating fluorescence parameters on the cells of interest should be carefully chosen to minimize the spectral overlap into the primary measurand antibody detector (30).

The antibody specification should contain acceptance criteria for both specific binding metrics and nonspecific binding metrics. Common specific binding metrics include population percent positivity, signal to noise ratio, saturation requirements, fluorescence intensity

ranges and other calculated metrics such as fold change, percent specific fluorescence signal, ratio metric changes between cells or spectral changes and rank order based metrics. Nonspecific binding metrics are often referenced to autofluorescence or isotype control intensity and limits to percent population expression within known negative populations. Negative and positive controls should bracket the expected fluorescence intensity of the target analyte, having additionally validated the linearity range of the assay. Should all cellular populations express the analyte of interest, then beads may need to be integrated into the assay design to serve as a surrogate limit of detection or negative control.

Fluorochrome selection is determined based on the antigen density expressed by the target cell population of interest (31). Antigen/antibody evaluations requiring highly quantitative assessment should be performed using fluorochromes measured in detectors with the lowest amount of spectral overlap contributions from all other antibody fluorochromes used in the panel design especially when positive gating antibody selections are used to identify the target population of interest (26,30–37).

The primary factors to be considered during antibody optimization are antigen/antibody saturation, optimal signal-to-noise ratios, minimized background fluorescence, antibody specificity and steric hindrance of antibody binding due to antigen density and close proximity of multiple antigen epitopes. Simple serial dilution antibody titrations against both positive and negative cellular targets are invaluable for antibody concentration optimization. It is preferable to perform antibody titrations against multiple sample sources that cover the expected range of population percentages and antigen expression as will be observed in the sample population (38). Use of cell lines is only acceptable for the evaluation of antibodies not expressed in healthy samples and to evaluate antibody performance at extreme limits of expected ranges. Cell lines may also be used as a secondary surrogate in spiking experiments to simulate varied population percentages seen in pathologic conditions. Titrations should always be performed using the same sample preparation, reagent systems, staining conditions, final staining volumes and cell concentration to be used during the assay. Titrations indicate antibody saturation staining concentrations while allowing identification of optimum signal-to-background staining concentrations at the same time (Fig. 8). Though saturation is desirable, some antigen/antibody binding pairs do not exhibit complete saturation. In such cases, signal-to-background is often used to identify the optimal staining concentration. Additionally, antibodies against highly expressed antigens may require including cold (unlabeled) antibody during the reaction to obtain proper labeling while limiting the fluorescence intensity and avoid fluorescence quenching. For antibodies such as activation markers or phospho-specific antibodies, it may be necessary to perform antibody titrations against multiple unmodulated and modulated samples under controlled conditions. Testing activation marker and phospho-antibody specificity has been performed
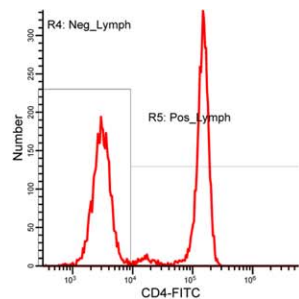
Antibody & Fluoro: CD4-FITC
Vendor: BDIS
Vendor Cat#: 340133
Vendor Lot#: 1683

Evaluation Date: 01/24/08
CAI Accession #: 1a08480
Expiration Date: 7/31/2009
Neat Volume (µL): 20

Titration 1/X	Negative Population			Positive Population			Noise to Signal Ratio
	Gated <#>	%Gate <%>	linMeanX <MESF>	Gated <#>	%Gate <%>	linMeanX <MESF>	
1	8064	46.7	3025	8064	53.3	146976	48.59
1/2	7626	47.2	3071	7626	52.8	145216	47.29
1/4	7862	48.8	3025	7862	51.2	142396	47.07
1/8	7950	48.5	3093	7950	51.5	123213	39.84
1/16	7448	47.6	3094	7448	52.4	81694	26.41
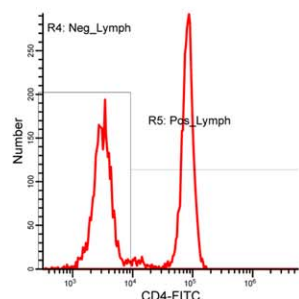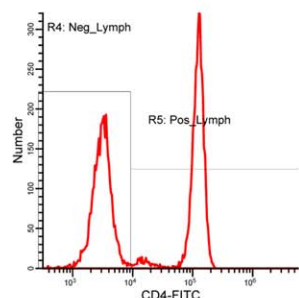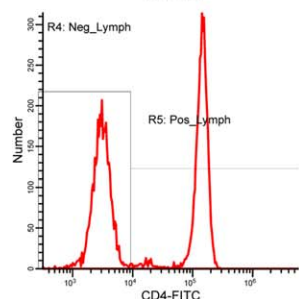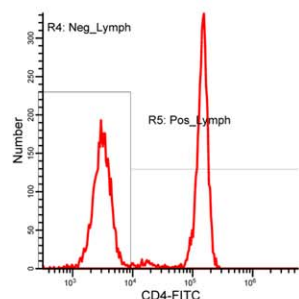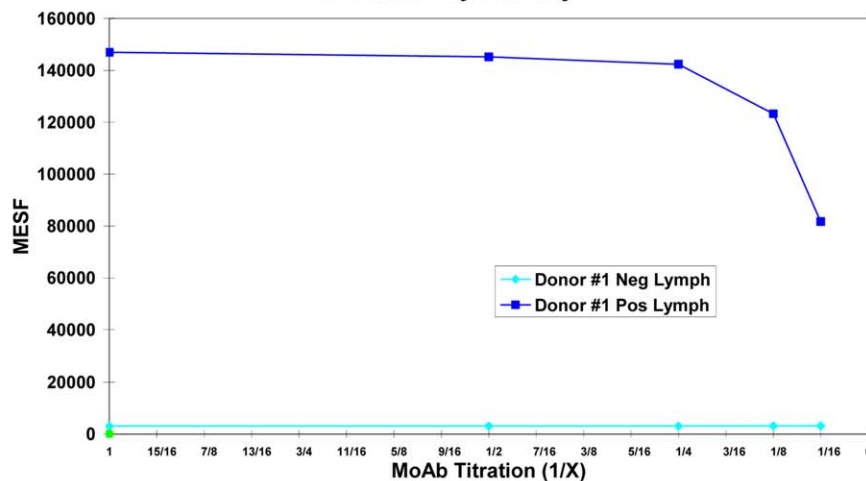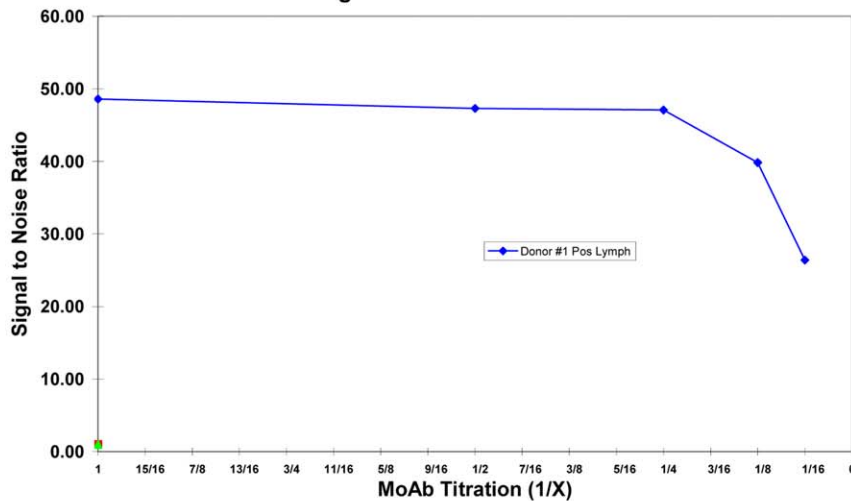
FIG. 8. Titration of CD4 antibody. Optimal dilution would be no lower than 1:4, as indicated by the decrease in signal to noise ratio and drop in fluorescence intensity. Single parameter histograms of the dilutions from lowest (top) to highest (bottom) are shown on left. The ratio of the positive to lymphocyte linMeanX provides a signal to noise axis in the lower right plot. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

using peptide-blocking experiments that employs matched phosphorylated and nonphosphorylated peptides in combination with irrelevant peptides as a means of demonstrating specific antibody binding. Additional approaches are the use of activation pathway-specific inhibitors or excess (≥100-fold) unlabeled antibody or

purified antigen in its native conformation to determine background or nonspecific binding levels of appropriate specific antibodies.

Antibody specificity and nonspecific binding characteristics should be compared to established antibody specification data sheets, as well as available literature citations. Side by side evaluation of multiple antibody clones and fluorochrome conjugates is often required to identify the best conjugate and clone for optimal detection and quantification. Once optimal saturation concentrations for all antibodies are determined, pilot combinations of all antibodies to be paired for simultaneous measurement should be tested against the same samples used for titrations to determine any potential interference that would artificially reduce the absolute fluorescence staining of any single antibody or proportions of cells stained (39). The use of fluorescence minus one (FMO) panel testing may also be instructive, if compensation is required for the assay (29).

Controls

It is of utmost importance to reliably distinguish between antigen-positive and antigen-negative cell populations in order to accurately measure the population of positive cells. The level of background derived from instrument (noise), autofluorescence, spectral overlap, and nonspecific antibody binding should be established using proper controls. Addition of beads to establish low fluorescence detection limits maybe necessary for assays targeting measurands that are ubiquitously expressed in cells.

Isotype controls are antibodies of the same class (isotype) of immunoglobulin as the specific antibody, but with specificity towards an antigen not present on the cells under study. Isotype controls should also match the fluorochrome type and number of fluorochrome molecules per immunoglobulin (F/P ratio) of the test antibody. These controls can determine whether there is undesirable antibody binding through Fc receptors and/or fluorochrome binding. However, it is difficult to include a properly matched control for each antibody in a multicolor assay. It is now an expert consensus that isotype controls should not be used to set positive gating regions in many situations and that internal cellular controls (i.e., negative cells) can be used to equally well determine the lowest level for which the antibody binding will be considered negative (21,22,25).

Isoclonic controls consist of a mixture of fluorochrome conjugated antibody and an excess amount (>100-fold excess) of the same, unlabeled antibody. This control can detect a fluorochrome-induced binding by demonstrating a lack of competition by the unlabeled antibody.

Internal negative controls are populations of cells within the specimen that do not express the studied antigen and thus should remain unlabeled in a given assay. In properly titrated assays this population should have the fluorescence intensity nearly as low as unstained cells. Since autofluorescence may differ in var-

ious cells types, background is optimally assessed if the negative control population is of the same cell type (e.g. B- or T- lymphocytes) or incorporates a means to compensate for autofluorescence. Comparing the fluorescence intensity of the internal negative control to an unstained control of the same cell type allows for an estimation of the level of nonspecific antibody binding. Cell types distinct from the assay target population may also serve as a negative process control, as it still may have advantages over batch type external controls.

Internal positive controls are populations of cells within the specimen that express the studied antigen and thus remain highly labeled in a given assay. In properly titrated assays the positive population should have a fluorescence intensity higher than or comparable to that expected for test sample stained cells. Comparing the fluorescence intensity of the internal positive control to an unstained or negative control of the same cell type allows for an estimation of the level of nonspecific antibody binding. Cell types distinct from the assay target population may also serve as a positive process control, as it still has advantages over stabilized external controls.

Surface and Intracellular Staining: Lysis and Permeabilization Procedures

Cell surface staining

Whole blood/bone marrow lysis methods are recommended over cell separation pre-analytical procedures by CSLI, as immunophenotyping after Ficoll isolation gives selective loss of different leukocyte populations and lower counts of lymphocyte subsets (37).

Four methodological variants of surface staining and red cell lysis are currently in use. (1) Stain-lyse-wash methods give the best signal discrimination but should be avoided when cell-loss due to washing is an issue. Samples to be stained for surface immunoglobulins should be thoroughly washed before incubation with monoclonal antibodies, in order to avoid the artifacts of cytophilic antibody. (2) Stain-lyse-no wash methods are recommended when unadulterated enumeration of cell populations is required, but can give higher background and may need to be avoided when "dim" antigens are investigated. (3) Lyse-stain-wash methods are used when cell concentration has to be adjusted before staining or red cells need to be removed. Some authors favor macrolysis, i.e. lysis in a large volume of reagent before wash and concentration of the sample. (4) No lyse-no wash whole blood methods have been developed when enumeration of leukocytes subsets is necessary. The large load in red blood cells may require use of a nuclear dye to positively select with assays on leukocytes or to exclude nucleated cells for red cell assays.

In a study of lymphocyte subset enumeration, the results indicated no significant difference between "lyse no wash" and a few observations using "no lyse no wash" methods from those obtained using "lyse and wash" methods (38). Some consider, no lyse-no wash methods better suited for absolute counting, for

example enumeration of CD34+ cells or CD4+ T-cells (39,40).

Modern commercialized lysis solutions can be safely applied and give comparable results, particularly if automated (41). Many commercial reagents include a fixative that should be validated and used if the acquisition is not immediate (<1 h) or necessary to follow universal precaution guidelines.

Intracellular Staining

Flow cytometric evaluation of specific intracellular epitopes, including proteins, epigenetic protein modifications (e.g., protein phosphorylation or methylation), DNA or RNA generally require that the target cell population be fixed and permeabilized in order to allow antibodies or target-binding dyes to cross the cytoplasmic and nuclear membranes. It is possible to permeabilize some cells without prior fixation and still measure intracellular antigens that are anchored within the cell (42). In general, it is necessary to fix the target cells to ensure that target epitopes do not escape from the cell and are optimally expressed for detection. In some cases, it is also useful if the target epitope is maintained in its native cellular or nuclear location. To accomplish this, cells generally require fixation, which can be accomplished using either a cross-linking fixative (formaldehyde or paraformaldehyde) or a low molecular weight alcohol (methanol or ethanol) (43–47), which fix cells based on their ability to coagulate proteins, and other cellular components.

Cell Fixation and Permeabilization

Some protein epitopes are denatured by alcohol treatment alone. The extent of cell fixation using formaldehyde is dependent on the fixation time, temperature, formaldehyde concentration and presence of other proteins in the suspending media (e.g. buffer vs. serum). Some intracellular epitopes are not detected unless high formaldehyde concentration or longer reaction times are used (46,47), while other intracellular epitopes show a decrease in expression in flow cytometry, with longer fixation times or high formaldehyde concentration. Formaldehyde fixation is attractive for intracellular epitopes, as it fixes the target epitope rapidly while maintaining its tertiary structure at the time of fixation.

Cell permeabilization is usually required to allow antibody or other probes access to the cytoplasmic or nuclear compartments of cells. In general, permeabilization is accomplished using detergents or alcohols. Thus, ethanol can both fix and permeabilize cells in a single step. But there are limitations of alcohol fixation and permeabilization, such as causing denaturation and loss of complex target epitopes. Additionally, alcohol fixation of samples containing significant numbers of red blood cells can result in a reduction in the recovery of leukocytes caused by an aggregation and trapping of cells. For these reasons, the majority of studies of intracellular epitopes or DNA content using clinical samples have used fixation with formaldehyde (or paraformaldehyde) followed by permeabilization with lysing reagents, such as saponin, NP-40, or Triton X-100 (43,44,47). For samples lacking significant red cells (e.g. isolated cells, peripheral blood mononuclear cells, cerebrospinal fluid, broncho-alveolar lavage, disrupted tissues), optimal epitope preservation may be obtained by fixation with formaldehyde or paraformaldehyde, followed by permeabilization using methanol or ethanol (45). Some phospho-epitopes generated by activation of signal transduction protein kinases, require treatment with relatively high alcohol concentration (after formaldehyde fixation) in order to be detected (48). In short, different specimens have different issues that must be addressed during validation of the assay.

While validating a fixation and permeabilization technique for a new intracellular epitope and/or different target cell population, it is useful to check whether the obtained intracellular staining is associated with the expected subcellular localization using fluorescence microscopy or image analysis. For the detection of antigens located in the nucleus, the use of large probes (e.g. pentameric IgM antibodies, large fluorophores, such as PE or APC, or large Quantum dots (Q-dots) should be carefully validated, as restrictions in the size of nuclear pores may limit the diffusion of large molecules into or out of the nucleus. Finally, it is important to be aware that fixation and/or permeabilization techniques can change the expression of cell surface molecules. For example, alcohol treatment can reduce or eliminate the expression of CD14 (49).

Negative Controls

If nonspecific binding is suspected by an antibody-conjugate, an isotype control may be of use to determine the level of background staining. However, it is generally better to use an internal cell population that lacks the target antigen as a negative control (49). One approach that has proven useful for phospho-epitope measurements is the use of targeted inhibitors that are specific for inhibition of expression of the target of interest (50).

Measuring Cell Surface Plus Intracellular Targets

Three approaches have been reported for the simultaneous measurement of cell surface and intracellular epitopes. The most common approach is to first fix and permeabilize the sample, and then simultaneously stain both surface and intracellular epitopes (47). This approach does not differentiate cell surface from cytoplasmic expression. Two approaches for antigen detection and localization are possible. The first is to initially stain cell surface epitopes, wash away the excess surface antibodies, then fix and permeabilize, and finally stain cytoplasmic epitopes (51). The alternative approach when methanol is used in high concentrations (50-80%) to permeabilize cells following formaldehyde fixation, is to first fix, then wash, then stain with antibodies to surface markers, wash, and then permeabilize with

methanol and subsequently stain with antibodies to cytoplasmic markers. Not all conjugates are amenable to this approach. Assay development should include a matrix of experiments to optimize analyte detection.

DNA Content Analysis

DNA content analysis has been performed on clinical samples for over thirty years, with the peak of its clinical application occurring in the 1990s. At that time, a DNA Consensus Conference published a set of guidelines that provides a good source of information for clinical implementation of tests using DNA content measurements, including cell cycle analysis or sub-G1 apoptotic cells (2). The expression of specific cell cycle-related proteins (such as histone H3 protein phosphorylated on Ser10 and cyclin A2) in the context of DNA content may provide additional information (52). The fixation and permeabilization technique used is frequently a compromise that allows DNA content measurements with sufficient quality while preserving target protein epitopes (45).

Compensation

Compensation is a complex exercise in flow cytometry, but very much an integral part of performing some successful flow multicolor immunofluorescence measurements. The advent of digital cytometers and the availability of different fluorochromes readily used in combinations for different purposes make this exercise complex, but now fairly automatic. Appropriate use of compensation requires a basic understanding of fluorochrome excitation and emission spectra, and each individual instrument's compensation set up (53–55). Most digital instruments come with their own compensation setup procedures, which when strictly followed and suited for the laboratory's purpose, can be useful in making the process simpler and reproducible for different operators. Quantitative assays of molecular expression may benefit from an assay design that avoids or minimizes compensation wherever possible.

Digital instrument compensation

It is not recommended to use manual compensation for complex experiments without the use of stringent standards or calibrators. Digital compensation can be achieved either pre- or post-data acquisition. Pre-data acquisition requires the setup of compensation tubes (based on the desired compensation methodology i.e. beads or live cells) and use of the instrument's compensation setup software.

After major maintenance is performed on an instrument, a new compensation matrix must be acquired and saved for subsequent assays using that specific panel, and settings. Even with post data acquisition compensation, it is recommended to run a compensation control to ensure proper compensation of the experiment. This can be achieved using a set of files that has already been used in previous compensation experiments and is known to the operator to represent a properly compensated analysis.

Compensation using cells

Compensating using cells can be advantageous and a challenge at the same time. Using live cells can be the closest target to actual test sample. Compensation using live cells often yields less manipulation post acquisition, since it is closest to the actual in vivo testing environment. Coupled with the actual antibodies used for patient testing, the upfront work would save a lot of time later since the post-acquisition verification could potentially be less work for the operator.

Single color with representative pure dye conjugate

It is a very good practice to pick the antibody that needs the most adjustment post compensation for each fluorochrome even compensating using pure dye conjugates. It is imperative that the compensation settings yield from this antibody still work for the other antibodies with the same fluorochrome.

Single color with specific tandem dye conjugates

The availability of tandem dye makes multi-color flow cytometry immunophenotyping protocols easily doable, however, optimal compensation requires understanding the tandem dye staining characteristics. This may translate to the operator performing frequent or tube by tube compensation adjustments as the antibody ages. Degradation of some tandem dyes in multi-color cocktails may be faster compared to when pipetted singly, particularly if exposed to more or repetitive light or oxidizing conditions.

Compensation Using Beads

Compensating using fluorescence beads has several advantages. First and foremost, there is no need to find an appropriate sample specimen. Also, the bead format is consistent, allowing compensation to be done without concern for sample stability or viability. The use of beads also homogenizes the protocols used and removes operator subjectivity when choosing a sample. Another added benefit is that it is much easier to track tandem dye degradation as there is no additional interference from the sample chosen. However, beads present their own problems that must be dealt with during compensation set-up, including cellular samples needing further refinements from instrument setting initially suggested by beads (56). The basic method for compensation using cells can also be applied to beads.

In specific compensation, extra tubes are stained with specific antibodies from a panel. As this is only generally needed with tandem dyes, most frequently PE-Cy7, it would be prudent to use the same marker in the tandem dye slots, such as CD45. The compensation matrix can be generated for specific tubes within the panel and applied individually to each tube as they are being run.

In bead-based compensation, the effect of tandem dye degradation is much easier to track since there are no issues with sample degradation or variation between runs.

Usually only a single unstained cell or reference bead tube is required, rather than contained in every tube, for proper compensation. Positive beads are stained with the relevant markers, one marker to a tube or as a mixture in a single tube.

There are also situations where cells may be preferred for certain specific compensation markers and in order for compensation to be calculated correctly, both negative and positive populations have to be in every tube, which is often achieved for instance with peripheral blood or bone marrow.

The biggest danger in using the bead format is autofluorescence. This may be caused by any number of reasons but is usually related to the dye conjugate and lack of spectral matching. This will cause cellular samples to be overcompensated when beads have been used as the sole format to establish compensations. As in all assays, cells in a variety of specimen should also be used to validate any bead-based compensation set-up.

Compensation Validation

At time of setup

Prior to formulating the validation plan for the assay, during verification, an adequate amount of time should be allotted to study and perform different possible compensation setup procedures for the assay. Commercial assays and LDTs should have clear instructions for use that cover compensation, if required for the assay. The ultimate procedure to be considered for validation should be reproducible, cost effective, and easy to follow by different level of operators.

Fluorescence minus one (FMO) compensation controls are samples labeled with all antibodies of the multicolor test sample except one in all possible combinations (57). This helps to determine both nonspecific antibody binding and background due to compensation for spectral overlap. It also allows determining positivity and setting regions in samples that contain multi-labeled populations. The "baseline offset" or bi-exponential data visualization methods that scale the axes on histograms and two-dimensional plots to enable visualization of signals from all cells, may facilitate proper use of FMO controls and setting boundaries between positive and negative cell subsets.

It is recommended to:

- determine autofluorescence using an unstained but fully processed cell sample, using the same settings as in the assay
- verify the expected binding characteristics of the antibody
- apply the proper titration assay to determine the antibody concentration resulting in the best resolution

- use isotype or isoclonic controls in assays known to give unexpectedly high background
- optimize combinations of fluorochrome-conjugated antibodies to minimize the need for spectral compensation
- use single labelled controls to determine the degree of spillover fluorescence into other detectors
- use FMO controls to set regions in multicolour samples
- use additional, specific parameters that allow cell populations of interest to be "pulled-out" of the overlapping population through sequential Boolean gating strategies in case of weak expression of antigens

While it is important to determine the optimal compensation set-up for each assay, it is also important to determine whether such data transformation is required for the assay intent. Compensation can compromise inter-instrument correlation due to differences in how spectral overlap is handled by a specific instrument model or software version and "baseline offset" or bi-exponential data visualization methods may also introduce a bias to the data.

Verification and Monitoring

Assessment of a 6+ multi-color flow panel with tandem dyes must be done upfront so that a proper verification and monitoring schedule for QC can be performed, including performance of compensation easily done by the laboratory personnel.

Most common practice for laboratories in performing and monitoring compensation are divided into two parts. One is a long procedure using the instrument's recommended compensation matrix set for the laboratory and the other is a short verification procedure performed more routinely in between the full calibration procedures. Most common reasons in performing a full compensation procedure are a change in the major hardware components (lasers, PMTs, alignment, flow cells, etc.) of the cytometer or a new lot of reagents or assay kit.

Verification using live cells is needed after each compensation procedure. The full panel needs to be tested after compensation is performed. Verification in between the long procedures can be achieved by checking a well characterized specimen on each instrument in the laboratory. In addition to assessing staining characteristics or expected staining patterns, it should be verified that each tube within the panel for this particular sample does not present over-compensation or under-compensation. This is normally a quality control procedure performed for each sample in most laboratories. If no problems are noted, the compensation daily monitoring is performed this way. If issues are observed, it has to be determined whether it is specific to one instrument, to the sample used, or the lot of antibody being used including operator error and instrument problem. The laboratory must have the verification procedure

clearly written and well documented for assurance of reproducibility and ease in monitoring.

There is still no substitute for the data analyst and interpreter to be familiar with the correct expected positive and negative staining characteristic for each antibody within the panel. The visual presentation of the data must be concise and easily interpretable and the compensation procedure facilitating the ease of interpretation, not complicating it.

Single vs. Multi-Instrument Compensation

When using a single instrument for the assay in question, the compensation will address the unique relationship between cells, reagents and instrument. When using two or more instruments, setting the compensation will be further challenged due to stability of the stained compensation tubes and number of runs that the lab needs to perform. Preparing the compensation tubes for particular experiments with more than 2 instruments will demand a large volume or multiple set-up tubes to provide adequate compensation samples for all the instruments. It is best to keep one instrument with a valid compensation matrix while the other instruments are being adjusted for new compensation settings. It is an additional advantage to use the same sample for verification post compensation between instruments, as it provides a stable point of comparison between instruments.

If using a single digital instrument, it may be acceptable to acquire high resolution FCS 3.0 data uncompensated and collect associated compensation controls to perform software compensation post acquisition, but this procedure also needs to be validated for every software version. To avoid problems in performing post acquisition compensation, standardized voltage and PMT settings must be used for the run for both samples and compensation controls.

Data Analysis

Data analysis is typically the last step of a diagnostic assay, yet errors introduced early in the assay can confound any strategy for interpreting the data. Therefore, when developing the data analysis portion of an assay, it is critical to ensure that the technique is robust, instrumentation properly configured, and the analysis strategy appropriate. For flow cytometric diagnostic assays identifying and quantifying cell subsets, the primary concern is to find appropriate strategies. Cell counts and percentages are typically reported for many assay types, but some tests results are reported in ABC, molecules per cell, MESF, or arbitrarily defined indexed units.

Validation of analysis procedures compares the new assay results with a accepted standard or predicted values. In case of analyses that include some subjectivity such as setting gates, it is important to determine the degree of reproducibility in data analysis between 2 or more operators.

Data Analysis Strategies

The Clinical Laboratory Standards Institute published guidelines for the enumeration of immunologically-defined cell populations (25), which detail the steps required to identify and analyze lymphocyte and stem cell subsets, which can be applied to other cell subsets. The most fit for purpose strategies must be considered for each specific analyte or intent of the assay.

Every flow cytometer has at least one piece of software included, which at a minimum allows the operator to acquire data. These packages are also usually capable of performing some post-acquisition analysis, typically by drawing regions and gates to partition the data into cell subsets. Cell-based assay designers usually adopt operator-defined gating as part of analysis strategy, but this is subjective (58). Assay designers should consider other options available and determine which offer the most robust and objective solution for each assay design.

Automatic Gating and Clustering Software

Automatic gating software is designed to eliminate subjective decisions (58–63), and is an available option in commercial products. If a gating approach is appropriate for an assay, it is possible to design the test initially with a manual gating strategy, and then to evaluate automatic gating solutions prior to putting the test into routine use. Software validation guidelines are available from U.S. FDA and other regulatory bodies, typically divided into categories of off-the-shelf software and customized, assay-specific software.

Modeling software has long been the accepted strategy for analysis of flow cytometry DNA cell cycle data (64,65) offering the advantage of accounting for overlapping populations. Modeling software has more recently been used in multi-parameter data analysis. Most modeling approaches do not require subjective operator decisions.

Select Strategies that Minimize Subjectivity and Maximize Reproducibility

When designing or deploying an assay, it is important to consider the impact of subjective decisions on the robustness of the test. Any time that an operator must make a choice, draw a boundary, or remember to perform a step, there is an opportunity for the assay to produce different results from the same data file. In some cases, the differences may be inconsequential to the test results. In other cases, such as rare-event analysis, the differences can radically change the end result. The differences should be considered in how a gate might be drawn if the operator is instructed to "create a loose gate that encompasses all monocytes" versus "draw a region that only includes monocytes." The purpose and importance of each step the operator is asked to perform must be delineated.

Well-trained operators must understand fully each step in the test and be able to perform the steps accurately and reproducibly. Assay designers should use the

training process as a means to edit and refine the assay's indications and instructions for use.

Analysis templates are a useful tool for ensuring that the assay is always performed in a reproducible manner. Most software used in tests allows templates (documents, models, etc.) to be created, saved and reused for analysis of subsequent test files in a similar manner. Templates help ensure that all critical elements are included, provide a QA parameter for the assay and they can serve as an example of how the assay is performed.

Sources of Error in Selected Strategies and Impact on Results

Any analysis approach will have sources of error associated with it, in addition to preanalytic variables, preparation, and acquisition that must be identified. For some analysis strategies, errors are compounded; that is, each analysis error is propagated to each subsequent step. Gating is a good example of this kind of error. Each dependent gate will include the error(s) introduced by the gates before it. Exclusion errors where cells of interest are accidentally excluded are particularly troublesome for rare-event types of analysis. Inclusion errors, where cells not of interest are included and influence the accuracy derived from the assay. Additionally, poor compensation can introduce errors and impair the ability to identify the subsets of interest.

Assay instructions should require that antibody and fluorochrome labels be entered prior to acquisition so that they are properly stored in the data files that will be analyzed. These labels are vital information for most software applications that will be used for subsequent analysis in assays.

Use Appropriate Data Transformations for Visualization of Data

When presenting flow cytometry data, it is often useful to include graphical displays of the important features from the acquired data files. Univariate histograms, bivariate dot and contour displays, and other plots can convey a great deal of information about the sample being analyzed. A key element of providing useful graphics is to select the appropriate transformation for the data, i.e. linear or log-like. Linear transformations are most often used to display light scatter parameters, if similarly sized particles, or DNA cell cycle data, where linear relationships are expected between populations. Log-like transforms are generally used for parameters that have a large dynamic range of intensities (typically ten-fold or more). Bin widths are wider for the low intensity values and narrow as intensity increases. This allows populations of vastly different intensities to be visualized and distinguished from one another on the same graph.

A true logarithmic transform does not allow values less than or equal to zero to be displayed, which is an important limitation in modern flow cytometry. Since it is now quite common to find negative values (i.e. values below zero intensity) in data files, other log-like transformations are generally recommended. Biexponential and hyperlog transforms are secondary alternatives to true

log transforms. These transforms are designed to maintain the useful characteristics of log transforms while addressing the main limitation. Zero and negative values are legal for these transforms.

Statistical Methods

Intensity measurements

Intensity measurements are typically reported using mean or median. Median identifies the value in the subset where half of the events are of higher and half are of lower intensity. For most flow cytometry applications, the median is preferred as it is less affected by outliers. Mean is the average value of the subset and can be computed as an "arithmetic mean" or as "geometric mean." The latter is commonly incorporated into flow cytometry data analysis software.

Intensity measurements are reported in "linear channels," but it is possible to use a quantitative conversion to report ABC or MESF. These quantitative measures typically require a bead standard or calibrator, as discussed above.

Proportional measurements

The relative amounts of different cell subsets are reported as cell counts and percentages. When reporting percentages, the parent population should be identified, i.e., "percent of total cells," "percent of lymphocytes." Alternatively ratiometric measurements can be made by comparing the fluorescence intensity of two cell populations to determine the relative fluorescence intensity. Most analysis software gives the user control over the number of decimal places to display which must be set so that it reflects that actual sensitivity and precision of the measurements being made.

Validation of results

Once an assay has been designed, it must be validated to ensure that it performs as expected. Other sections of this document discuss methods for verifying instrument performance. It is best to isolate and test each of the sources of variability one at a time as the overall variability of the entire assay will be greater than any individual element causing variability. It may be easiest to think of the overall variability as the sum of the individual sources, though this is an oversimplification of the total variability. Typical sources of assay variability include:

- Variability with same sample, same operator, and same instrument can be tested by splitting a sample and performing replicate runs to acquire several data files from the same experiment. No more than six replicates should be run since each "measurement" actually is derived from measurements of tens of thousands of cells for each replicate analysis.
- Machine-to-machine variability can also be tested with a split sample. Operator-to-operator variability may occur at any step of the assay: sample handling and

preparation, acquisition and data analysis. The former can be tested by splitting the sample prior to preparation and the latter by performing data analysis on the same sets of data files.

- Software-to-software variability. It is sometimes useful to compare the LDT results produced by two or more software applications. It is also important to verify that the software application itself produces consistent results.

Conditions for Valid Analysis and Definition of Out-of-range Cases

Part of the validation process is to determine the range of results that the assay should be able to include and provide a statistically valid measurement. If results are compared to "normal" results, normal ranges must be established with an adequate sample of normal donors. It is important to identify interfering substances or conditions that could interfere with the assay.

LITERATURE CITED

1. Hoffman RA. Standardization, calibration and control in flow cytometry. In: Robinson JP, Darzynkiewicz Z, Dean P, Dressler L, Tanke H, Rabinovich P, Stewart C, Wheeless L, editors. Current Protocols in Cytometry. New York: Wiley, 2005.
2. Shankey TV, Rabinovitch PS, Bagwell B, Bauer B, Duque RE, Hedley DW, Mayall BH, Wheeless L. Guidelines for the Implementation of Clinical DNA Cytometry. Cytometry 1993;14:472–477.
3. Bagwell CB, Baker D, Whetstone S, Munson M, Hitchcox S, Ault KA, Lovett EJ. A simple and rapid method for determining the linearity of a flow cytometer amplification system. Cytometry 1989;10:689–694.
4. Wang L, Gaigalas AK, Marti G, Abbasi F, Hoffman RA. Toward quantitative fluorescence measurements with multicolor flow cytometry. Cytometry Part A 2008;73A:279–288.
5. Hoffman RA, Wood JC. Characterization of flow cytometer instrument sensitivity. Current Protocols Cytometry. 2007, Chapter 1:Unit 1.20.
6. Bikoue A, Janossy G, Barnett D. Stabilised cellular immunofluorescence assay: CD45 expression as a calibration standard for human leukocytes. J Immunol Methods 2002;266:19–32.
7. Gratama JW, D'Hautcourt JL, Mandy F, Rothe G, Barnett D, Janossy G, Papa S, Schmitz G, Lenkei R. Flow cytometric quantitation of immunofluorescence intensity: Problems and perspectives. Cytometry 1998;33:166–178.
8. Hoffman RA. Standardization and quantitation in flow cytometry. Methods Cell Biol 2001;63:299–340.
9. Davis KA, Abrams B, Iyer SB, Hoffman RA, Bishop JE. Determination of CD4 antigen density on cells: Role of antibody valency, avidity, clones, and conjugation. Cytometry 1998;33:197–205.
10. Bikoue A, George F, Poncelet P, Mutin M, Janossy G, Sampol J. Quantitative analysis of leukocyte membrane antigen expression: normal adult values. Cytometry 1996;26:137–147.
11. Poncelet P, Carayon P. Cytofluorometric quantification of cell-surface antigens by indirect immunofluorescence using monoclonal antibodies. J Immunol Methods 1985;85:65–74.
12. Rossmann ED, Lenkei R, Lundin J, Mellstedt H, Osterborg A. Performance of calibration standards for antigen quantitation with flow cytometry in chronic lymphocytic leukemia. Cytometry Part B, Clin Cyto 2007;72B:450–457.
13. Lenkei R, Gratama JW, Rothe G, Schmitz G, D'hautcourt JL, Arekrans A, Mandy F, Marti G. Performance of calibration standards for antigen quantitation with flow cytometry. Cytometry 1998;33:188–196.
14. Serke S, van Lessen A, Huhn D. Quantitative fluorescence flow cytometry: A comparison of the three techniques for direct and indirect immunofluorescence. Cytometry 1998;33:179–187.
15. Lenkei R, Andersson B. Determination of the antibody binding capacity of lymphocyte membrane antigens by flow cytometry in 58 blood donors. J Immunol Methods 1995;183:267–277.
16. Wang L, Gaigalas AK, Abbasi F, Marti GE, Vogt RF, Schwartz A. Quantitating fluorescence intensity from fluorophores: Practical use of MESF values. J Res Nat Ins Stand Technol 2002;107:339–353.
17. Wang L, Abbasi F, Gaigalas AK, Vogt RF, Marti GE. Comparison of fluorescein and phycoerythrin conjugates for quantifying CD20 expression on normal and leukemic B-cells. Cytometry Part B Clin Cyto 2006;70B:410–415.
18. Poncelet P, Poinas G, Corbeau P, Devaux C, Tubiana N, Muloko N, Tamalet C, Chermann JC, Kourilsky F, Sampol J. Surface CD4 density remains constant on lymphocytes of HIV-infected patients in the progression of disease. Res Immunol 1991;142:291–298.
19. Wang L, Abbasi F, Ornatsky O, Cole KD, Misakian M, Gaigalas AK, He HJ, Marti GE, Tanner S, Stebbings R. Human CD4+ lymphocytes for antigen quantification: characterization using conventional flow cytometry and mass cytometry. Cytometry A 2012;81A:567–575.
20. Hoffman RA, Wang L, Bigos M, Nolan JP. NIST/ISAC standardization study: Variability in assignment of intensity values to fluorescence standard beads and in cross calibration of standard beads to hard dyed beads. Cytometry A 2012;81A:785–796.
21. Davis BH. Improveddiagnostic approaches toinfection/sepsis detection. Expert Rev Mol Diagn 2005;5:193–207.
22. Clinical and Laboratory Standards Institute (CLSI). Fluorescence Calibration and Quantitative Measurement of Fluorescence Intensity; Approved Guideline. NCCLS document I/LA24-A (ISBN 1-56238-543-7). CLSI, 940 West Valley Road, Suite 1400, Wayne, Pennsylvania 19087-1898 USA, 2004.
23. Gibot S, Béné MC, Noel R, Massin F, Guy J, Cravoisy A, Barraud D, De Carvalho Bittencourt M, Quenot JP, Bollaert PE, Faure G, Charles PE. Combination biomarkers to diagnose sepsis in the critically ill patient. Am J RespirCrit Care Med 2012;186:65–71.
24. Icardi M, Erickson Y, Kilborn S, Stewart B, Grief B, Scharnweber G. CD64 Index provides simple and predictive testing for detection and monitoring of sepsis and bacterial infection in hospital patients. J Clin Microbiol 2009;47:3914–3919.
25. Clinical and Laboratory Standards Institute (CLSI). Enumeration of Immunologically Defined Cell Populations by Flow Cytometry; Approved Guideline—Second Edition. CLSI document H42-A2 (ISBN 1-56238-640-9). Clinical and Laboratory Standards Institute, 940 West Valley Road, Suite 1400, Wayne, Pennsylvania 19087-1898 USA, 2007.
26. Schwartz A, Wang L, Early E, Gaigalas A, Zhang Y, Marti GE, Vogt RF. Quantitating fluorescence intensity from fluorophore: The definition of MESF assignment. J Res Nat Inst Stand Technol 2002;107: 83–91.
27. Borowitz MJ, Guenther KL, Shults KE, Stelzer GT. Immunophenotyping of acute leukemia by flow cytometric analysis: Use of CD45 and right-angle light scatter to gate on leukemic blasts in three-color analysis. Am J Clin Pathol 2003;100:534–540.
28. Wood B. 9-Color and 10-Color Flow Cytometry in the Clinical Laboratory. Arch Pathol Lab Med 2006;130:680–690.
29. Maecker HT, Trotter J. Flow cytometry controls, instrument setup, and the determination of positivity. Cytometry Part A 2006;69A: 1037–1042.
30. Hulspas R, O'Gorman MRG, Wood BL, Gratama J, Sutherland DR. Considerations for the control of background fluorescence in clinical flow cytometry. Cytometry Part B 2009;76B:355–364.
31. Maecker HT, Frey T, Laurel E. Nomura LE, Trotter J. Selecting fluorochrome conjugates for maximum sensitivity. Cytometry Part A 2004;62A:169–173.
32. Bigos M, Baumgarth N, Jager GC, Herman OH, Nozaki T, Stovel RT, Parks DR, Herzenberg LA. Nine color eleven parameter immunophenotyping using three laser flow cytometry. Cytometry 1999;36:36–45.
33. McLaughlin BE, Baumgarth N, Bigos M, Roederer M, De Rosa SC, Altman JD, Nixon DF, Ottinger J, Oxford C, Evans TG, Asmuth DM. Nine-color flow cytometry for accurate measurement of t cell subsets and cytokine responses. I. Panel design by an empiric approach. Cytometry Part A 2008;73A:400–410.
34. McLaughlin BE, Baumgarth N, Bigos M, Roederer M, De Rosa SC, Altman JD, Nixon DF, Ottinger J, Li J, Beckett L, Shacklett BL, Evans TG, Asmuth DM. Nine-color flow cytometry for accurate measurement of t cell subsets and cytokine responses. II. Panel performance across different instrument platforms. Cytometry Part A 2008;73A:411–420.
35. Purvis NP, Stelzer GT. Multi-platform, multi-site instrumentation and reagent standardization. Cytometry 1998;33:156–165.
36. Stelzer GT, Marti G, Hurley A, McCoy P Jr, Lovett EJ, Schwartz A. U.S.-Canadian Consensus recommendations on the immunophenotypic analysis of hematologic neoplasia by flow cytometry:

Standardization and validation of laboratory procedures. Cytometry 1997;30:214–230.

37. Renzi P, Ginns LC. Analysis of T-cell subsets in normal adults. Comparison of whole blood lysis technique to Ficoll-Hypaque separation by flow cytometry. J Immunol Methods 1987;98:53–56.

38. Levering WH, van Wieringen WN, Kraan J, van Beers WA, Sintnicolaas K, van Rhenen DJ, Gratama JW. Flow cytometric lymphocyte subset enumeration: 10 years of external quality assessment in the Benelux countries. Cytometry B Clin Cytom 2008;74B:79–90.

39. Alvarez-Larran A, Jover L, Marin P, Petriz J. A multicolor, no-lyse no-wash assay for the absolute counting of CD34 cells by flow cytometry. Cytometry 2002;50:249–253.

40. Greve B, Cassens U, Westerberg C, Raulin C. A new no-lyse, no-wash flow-cytometric method for the determination of CD4 T cells in blood samples. Transfus Med Hemother 2003;30:8–13.

41. Mekouar H, Mullier F, Dinon Q, Bailly N, Cornet Y, Chatelain B. No effect of lysis solutions on absolute CD19+ lymphocytes count and CD45 index in chronic lymphocytic leukemia. Cytometry Part B Clin Cytometry 2011;80:369–372.

42. Schroff RW, Bucana CD, Klein RA, Farrell MM, Morgan AC Jr. Detection of intracytoplasmic antigens by flow cytometry. J Immunol Meth 1984;70:167–177.

43. Clevenger CV, Bauer KD, Epstein AL. A method for simultaneous nuclear immunofluorescence and DNA content quantification using monoclonal antibodies and flow cytometry. Cytometry 1985;6:208–214.

44. Schimenti KJ, Jacobberger JW. Fixation of mammalian cells for flow cytometric evaluation of DNA content and nuclear immunofluorescence. Cytometry 1992;13:48–59.

45. Pollice AA, McCoy JP Jr, Shackney SE, Smith CA, Agarwal J, Burholt DR, Janocko LE, Hornicek FJ, Singh SG, Hartsock RJ. Sequential paraformaldehyde and methanol fixation for simultaneous flow cytometric analysis of DNA, cell surface proteins, and intracellular proteins. Cytometry 1992;13:432–444.

46. Mann GJ, Dyne M, Musgrove EA. Immunofluorescent quantification of ribonucleotide reductase M1 subunit and correlation with DNA content by flow cytometry. Cytometry 1987;8:509–517.

47. Chow S, Hedley D, Grom P, Magari R, Jacobberger JW, Shankey TV. Whole blood fixation and permeabilization protocol with red blood cell lysis for flow cytometry of phosphor-epitope expression in leukocyte subpopulations. Cytometry A 2005;67A:4–17.

48. Woost PG, Solchaga LA, Meyerson HJ, Shankey TV, Goolsby CL, Jacobberger JW. High Resolution kinetics of cytokine signaling in human CD34/CD117-positive cells in unfractionated bone marrow. Blood 2011;117:e131–e141.

49. Keeney M, Gratama JW, Chin-Yee IH, Sutherland DR. Isotype controls in the analysis of lymphocytes and CD34+ stem and progenitor cells by flow cytometry—Time to let go! Cytometry 1998;34:280–283.

50. Hedley DW, Chow S, Shankey TV. Cytometry of intracellular signalling: From laboratory bench to clinical application. In: Darzbybkiewicz Z, Robinson JP, editors. Methods in Cell Biology, 4th ed. San Diego, CA: Academic Press; 2011.

51. Shults KE, Miller DT, Davis BH, Flye L, Hobbs LA, Stelzer GT. A standardized ZAP-70 assay—Lessons learned in the trenches. Cytometry B Clin Cytom 2006;70B:276–283.

52. Jacobberger JW, Sramkoski RM, Stefan T. Multiparameter Cell Cycle Analysis. Meth Mol Biol 2011;699:229–249.

53. Roederer M. Spectral compensation for flow cytometry: Visualization artifacts, limitations, and caveats. Cytometry 2001;45:194–205.

54. Bayer J, Grundwald D, Lambert C, Mayol JF, Maynadié M for the Association Francaise de Cytometrie Thematic Workshop on Fluorescence Compensation Settings in Multicolor Flow Cytometry. Cytometry B Clin Cytom 2007;72B:8–13.

55. Baqwell CB. Hyperlog-a Flexible Log-like Transform for negative, zero and positive valued data. Cytometry Part A 2005;64:34–42.

56. DiGiuseppe JA, Cardinali J. Improved compensation of the fluorochrome Am Cyan using cellular controls. Cytometry B Clin Cytom 2011;80B:191–194.

57. De Rosa SC, Brenchly JM, Roederer M. Beyond six colors: A new era in flow cytometry. Nat Med 2003;1:112–117.

58. Naumann U, Luta G, Wand M. The curvHDR method for gating flow cytometry samples. BMC Bioinformatics 2010;11:44–49.

59. Lo K, Hahne F, Brinkman RR, Gottardo R. FlowClust: A Bioconductor package for automated gating of flow cytometry data. BMC Bioinformatics 2009;10:145.

60. Walther G, Zimmerman N, Moore W, Parks D, Meehan S, Belitskaya I, Pan J, Herzenberg L. Automatic clustering of flow cytometry data with density-based merging. Adv Bioinformatics 2009:686759. Epub 2009 Nov 19.

61. Herbert DJ, Miller DT, Bruce Bagwell C. Automated analysis of flow cytometric data for CD34+ stem cell enumeration using a probability state model. Cytometry B 2012;82B:313–318.

62. Miller DT, Hunsberger BC, Bruce Bagwell C. Automated analysis of GPI-deficient leukocyte flow cytometric data using gemStone™. Cytometry B ClinCytom 2012;82B:319–324.

63. Wong L, Hunsberger BC, Bruce Bagwell C, Davis BH. Automated quantitation of fetomaternal hemorrhage by flow cytometry for HbF-containing fetal red blood cells using probability state modeling. Int J Lab Hematol 2013 [Epub ahead of print]

64. MultiCycle, Phoenix Flow Systems, San Diego, CA, USA.

65. ModFit LT, Verity Software House, Topsham, ME, USA.