Technical Note

Flow Cytometry Controls, Instrument Setup, and the Determination of Positivity

Holden T. Maecker* and **Joseph Trotter**

BD Biosciences, San Jose, California

A frequent goal of flow cytometric analysis is to classify cells as positive or negative for a given marker, or to determine the precise ratio of positive to negative cells. This requires good and reproducible instrument setup, and careful use of controls for analyzing and interpreting the data. The type of controls to include in various kinds of flow cytometry experiments is a matter of some debate and discussion. In this tutorial, we classify controls in various categories, describe the options within each category, and discuss the merits of each option. © 2006 International Society for Analytical Cytology

Key terms: control; gating; specificity; compensation; instrumentation setup

Control samples are essential in flow cytometry, since they provide the context within which one can interpret test samples. Any experiment may, and probably should, contain at least three types of controls: setup (or instrument) controls, specificity (or gating) controls, and biological comparison controls. In some cases, the same control sample may serve more than one of these purposes. *Setup or instrument controls* are those that are used to properly set up (or at least check the setup of) the instrument, including photomultiplier tube (PMT) voltage gains and compensation. *Specificity or gating controls* are those used to help distinguish specific from nonspecific binding. These are often used to set the location of gates or graphical regions used to classify the cells. In other words, they are used to determine positivity or negativity for particular markers. *Biological comparison controls* are those that provide biologically relevant comparison conditions, for example, unstimulated samples or healthy donor samples. In some cases, these can function as gating controls; for example, an unstimulated sample can function to define a positive/negative threshold for a cytokine response. In the sections below, we will discuss these three types of controls in some detail.

SETUP CONTROLS

Certain aspects of cytometer setup, such as laser alignment, laser time delay, sensitivity, etc., need to be checked periodically, but may not require controls in every experiment. Controls for checking and monitoring these parameters are beyond the scope of this discussion, as we will assume that the cytometer in question is in good working order, and that a reasonable instrument quality control routine is performed.

Each experiment will, however, require certain setup controls specific to the markers and fluorochromes used. Instrument setup and optimization strategies can vary between different types of cytometers (analog vs. digital) and for different types of experiments (one or two color vs. multicolor). However, any strategy for multicolor setup (in which spectral overlap occurs) must include two elements: (1) setting the instrument gain (i.e. PMT voltages) and (2) determining the degree of spectral overlap and required compensation for that experiment. Because calculation of compensation values is voltage dependent, these steps must occur in sequential order, unless the cytometer software has a means of recalculating compensation when PMT voltages are altered.

SETTING PMT VOLTAGES

PMT voltages are often set using an unstained sample of the cells in question. In this strategy, the voltages are adjusted such that the unstained cells appear in the first decade (or first quartile) of a 4 decade logarithmic scale for each fluorochrome to be measured. While this strategy can provide good results for some situations, it is not universally optimal. In particular, it can be suboptimal for

WILEY InterScience®
DISCOVER SOMETHING GREAT

fluorochromes with longer wavelength emissions, such as APC, Cy7, Alexa 700, etc. This is because most cells emit little autofluorescence at these wavelengths, so the fluorescence intensity of unstained cells is near zero, and the variance is dominated by photon counting statistics and electronic noise. Any attempt to adjust voltages in these channels based on visual placement of an unstained sample is inherently difficult and subjective, and may not correlate with optimized detection of true signals in that channel.

Thus, it becomes prudent to determine the minimal voltage required to ensure that each detector has enough gain applied to sufficiently boost dim signals above a level where underlying electronic noise significantly contributes to the measurement. These detector settings can then be used as baseline starting values when setting up an experiment. Resolution sensitivity at the lower end, or the ability to resolve discrete dim populations, can be correlated with the CV of dim particles (1). In each fluorescence detector, the higher the population mean, the lower the contribution of electronic noise to the population CV. By choosing particles with similar fluorescence signals as dimly stained cells, and sequentially varying detector gain to move the population mean over a reasonable range, one can plot the population CV vs. the voltage applied (Fig. 1). The resulting curves have a predictable shape, with CV initially decreasing as gain is increased, until sufficient voltage is applied such that the CV is stabilized with minimal electronic noise contribution. The inflection point of each curve can be taken as a reasonable minimum voltage to use for optimal resolution sensitivity when the brightness of the unstained or dim cells is similar to that of the standard particles. This PMT voltage may be regarded as a baseline voltage to ensure that the instrument will do the best job of distinguishing dim from unstained events.

On analog instruments, the electronic baseline correction circuits make it impossible to optimize PMT voltages based on measured dim particle CVs, since the signals are clipped at the low end (i.e., no values below zero are reported). Alternatively, one can make similar sequential measurements of signal:noise (median of the positive population/median of the negative population) for a mixed population containing both negative and positive events, and the result will be a series of curves with a shape that is inverted compared to those in Figure 1. Again, the inflection point of each curve represents the minimum voltage for optimal resolution sensitivity at the low end of the expected range of fluorescence.

The above strategies for determining minimum baseline PMT voltages need not be carried out more than once for a particular instrument with a given optical configuration, unless there is a significant change in either the instrument or the background fluorescence of the cells being analyzed. It need only be repeated when filters, PMTs, lasers, or fluorochromes are changed. However, these baseline voltages are starting points and will not automatically be optimal for every experiment. Some adjustment should be made for significantly different cell types, differ-
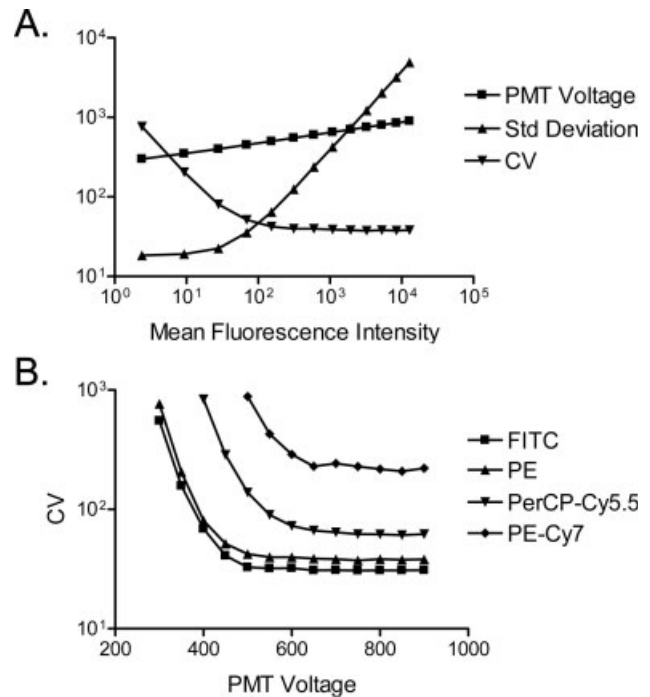

Fig. 1. (**A**) Characterization of a particular detector when analyzing dim fluorescent particles [SPHERO Rainbow Calibration Particles, Peak 2 (Spherotech, Libertyville, IL)]. There is an expected linear relationship between log PMT voltage and fluorescence intensity. The standard deviation and CV also relate to fluorescence intensity in a predictable manner. Because of the highly linear relationship between log PMT voltage and fluorescence intensity, and because standard deviation is imbedded in the CV, one can simplify the characterization of each detector by simply plotting CV vs. PMT voltage. This is shown in (**B**) for several detector channels excited by a 488 nm laser, using the Peak 2 beads. The inflection point for each curve indicates the minimum baseline voltage for optimal resolution sensitivity in each channel. Note that the suitability of this method is dependent upon reasonable matching of fluorescence intensity of the particles used and autofluorescence or dim staining of the cells of interest in each channel.

ent antibody-fluorochrome panels, and for significantly different experimental protocols (e.g., surface vs. intracellular staining). This requires the use of a fully stained sample to characterize the populations, hopefully containing the complete range of signals (negative and positive) to be present in that experiment.

When examining such a fully stained sample using baseline instrument gains, what conditions should prompt further adjustment of the detector voltages? Most importantly, if the positive signals are so high as to be off-scale in a particular detector, the PMT voltage should be reduced to bring all events on-scale. This is not only for aesthetic reasons, but is essentially required since the true extent of their fluorescence cannot be measured and both their brightness and compensation in other dye dimensions would be incorrect. Second, if known negative events occur very high on the fluorescence scale, one may wish to reduce the PMT voltage in that detector and bring the negative population mean down until its lower left approaches that of the right side of an unstained population at baseline gain. However, this is not always advisable, because it can result in an increase in measurement

error that can ultimately impact compensated populations, not to mention a false impression of the location of true background staining. In situations of high background (when positive events are still on-scale), it is generally advised to first investigate and reduce sources of the high background. This might include reducing the titer of the antibody (and/or choosing a higher avidity antibody), increasing the number of washes, and using an unlabeled irrelevant antibody or Fc-binding reagent to block nonspecific staining.

SETTING COMPENSATION

Fluorescence spillover is inherent in most multicolor experiments, is easily measured, and the data is subsequently "compensated" to remove the effects of that spillover. The net result is that populations with equivalent fluorescence in a given channel will have similar means (or medians), regardless of their fluorescence in other channels. Even though many analog instruments have compensation circuits, it is important to consider compensation as experiment-associated rather than an instrument setting per se. Once the proper gains are determined for each immunofluorescence detector, determining the experiment compensation is straightforward since the spillover characteristics of each fluorochrome may be directly measured. While compensation can result in visual spread of negative populations in one or more compensated fluorochrome dimensions (2,3), the display is still easier to interpret than an uncompensated display. Estimating compensation properly requires the use of some type of control, most commonly a series of bead or cell samples, each stained with a single fluorescent marker used in the experiment. Correct compensation is then calculated and applied either through the instrument itself or via the software. In either case, it is imperative that the compensation control samples are run after the PMT voltages for the experiment have been set, since the degree of compensation required is gain-dependent. However, some newer cytometers have a limited ability to automatically adjust compensation settings when PMT voltages are changed.

Use of single-stained cells as compensation controls is common, and works well in many situations. One drawback of this approach is that it requires a source of cells: either extra cells from the test samples are used or another source of cells must be procured. A second drawback is that, when certain markers are dim and/or found only on rare subsets of cells, it becomes difficult to accurately set compensation using antibodies to those markers. In such cases, a different antibody conjugated to the same fluorochrome is often used as a compensation control. In fact, some investigators routinely use an antibody like CD8 (which is very bright and present on a reasonable subset of cells) for compensation in each channel, regardless of what antibody will be used experimentally in that channel. In many cases, this is a perfectly acceptable practice, since use of a signal that is equal or brighter in intensity to the signal in the test samples will allow for accurate calculation of compensation for a given dye.

A difficulty arises when certain tandem dyes are used, especially APC-Cy7 and PE-Cy7. These fluorochromes tend to have more lot-to-lot variability in their emission spectra, and their spillover is also affected by handling and storage of the antibody and stained samples. When using these dyes, substitution of a different antibody for compensation than is present in the test samples is a risky practice and may result in inaccurate compensation.

Two types of beads can be used to set compensation in lieu of stained cells: beads that are stained with a particular fluorophore, or capture beads that can be stained with a labeled antibody of the user's choice. The former are convenient for many applications where emission characteristics of the dyes are stable, and have been used by manufacturers along with dedicated software to provide automated compensation routines that are often used in clinical laboratories. However, when tandem dyes such as APC-Cy7 or PE-Cy7 are included in an experiment, these prestained beads suffer from the same problem as mentioned above: they may not be able to accurately set compensation for every antibody conjugate that can be used in these channels. Thus, they need to be used with caution and with careful control of the APC-Cy7 and PE-Cy7 conjugates that will be used in the experiment.

Capture beads differ in that they provide essentially the best of both worlds: they can be stained with antibodies much like cells, but because the antibodies are captured without regard to their specificity, they provide a bright, uniform signal for each antibody regardless of how bright that antibody would stain cells. It is rare that an antibody appears brighter on cells than on capture beads. The beads can also be handled along with the test samples so that they are exposed to the same experimental conditions (fixation, permeabilization, temperature, light exposure, etc.) as the test samples, which can again be important for sensitive tandem conjugates. The only major caveat to using capture beads is that they require a different scatter gate than cells, and they will only capture antibodies of a certain class [e.g. those with mouse kappa light chains for an anti-mouse kappa capture bead (BD Biosciences, San Diego, CA)]. Knowledge of the source species and class of antibody used in the experiment is thus important to successful use of capture beads for compensation.

SPECIFICITY (GATING) CONTROLS

Once PMT voltages are set and compensation samples are run, the experimental samples can be collected and analyzed. At this point, controls may be required to help set gates, or positive/negative boundaries in the data. These gates are used to answer fundamental questions in cytometry: are cells positive or negative for a certain marker, or what proportion of cells is positive? As such, gating controls can take on great importance to the interpretation of the data.

One might first ask when gating controls are actually needed. Certainly, some gates can be drawn unambiguously, without reference to any control sample. For example, any marker that has clearly bimodal expression, with

no overlap of positive and negative populations, does not require a control for accurate determination of positive and negative populations. This would usually include markers such as CD4 and CD8 on T cells, for example. The only caution here is that, in some cases, a small subset of dimly stained cells may be important (e.g. in functional responses) and consistent inclusion of these dim cells in the positive gate would be advised (3,4).

Gating controls become important when there is no clear division between positive and negative populations. Certain activation markers, such as CD25, CD38, CD69, and some cytokines might be included in this category. The gating control for these markers might be a biological comparison control (unstimulated or irrelevantly stimulated cells, see next section), or it might be a more generic control such as an isotype control or fluorescence-minus-one (FMO) control (3).

ISOTYPE CONTROLS

Isotype controls have a long history in flow cytometry, and are meant to account for nonspecific staining of an antibody of a particular isotype conjugated to a particular fluorochrome. Matching the isotype of the test antibody is important; Figure 2 demonstrates that antibodies of different isotypes can have different levels of background staining. However, even when the control antibody is isotype-matched to the test antibody, there are two main limitations to the usefulness of this type of control. The first limitation is that individual antibody conjugates have various levels of background staining, depending upon their specificity, concentration, degree of aggregation, and fluorophore:antibody ratio, among other variables. It is thus a hit-or-miss prospect to find an isotype control that truly matches the background staining of a particular test antibody. And, remembering that we are using the isotype control to help us define the true level of background staining, this becomes a circular proposition.

The second limitation of isotype controls is that, by themselves, they do not account for fluorescence spillover from other channels. This limitation can be overcome by including all relevant antibodies in the other channels along with an isotype control antibody in a single channel of interest. However, even this approach still suffers from the first limitation, namely that the isotype control may not match the test antibody for background staining.

FMO CONTROLS

When high-quality monoclonal antibody conjugates are used at appropriate concentrations, they tend to have relatively low background staining. As such, in experiments of >4 colors, the major source of background staining tends to be fluorescence spillover. Because of this, the use of FMO controls has become both popular and prudent. FMO controls are samples that include all of the antibody conjugates present in the test samples except one (3). The channel in which the antibody conjugate is missing is the one for which the FMO provides a gating control.

It is important to recognize what an FMO control does and does not provide. It does provide a means to measure
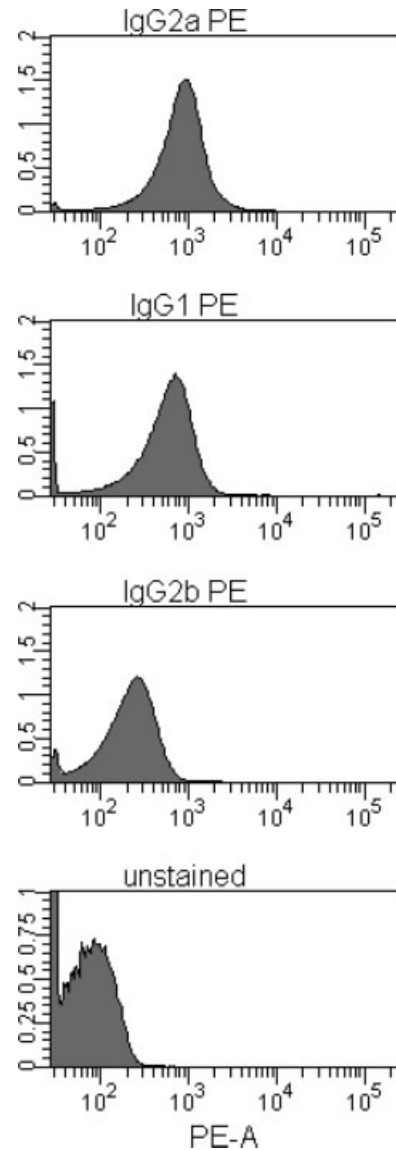

FIG. 2. The staining of isotype controls can vary significantly from each other and from unstained cells. Three different isotype controls, each used at the recommended concentration, are compared with unstained cells in intracellular staining. All plots are gated on small lymphocytes.

the effects of spillover from populations in other dye dimensions on a particular channel of interest. This can be very important when trying to develop new multicolor panels, where maximizing resolution sensitivity in certain channels is required. Because it does not contain an antibody in the channel of interest, an FMO control does not provide a measure of background staining that may be present when an antibody is actually included in that channel. As long as background staining is insignificant in comparison to background caused by spillover, this is not an issue. But if antibody-dependent background becomes significant, the FMO alone will no longer be appropriate for determining a negative/positive boundary.
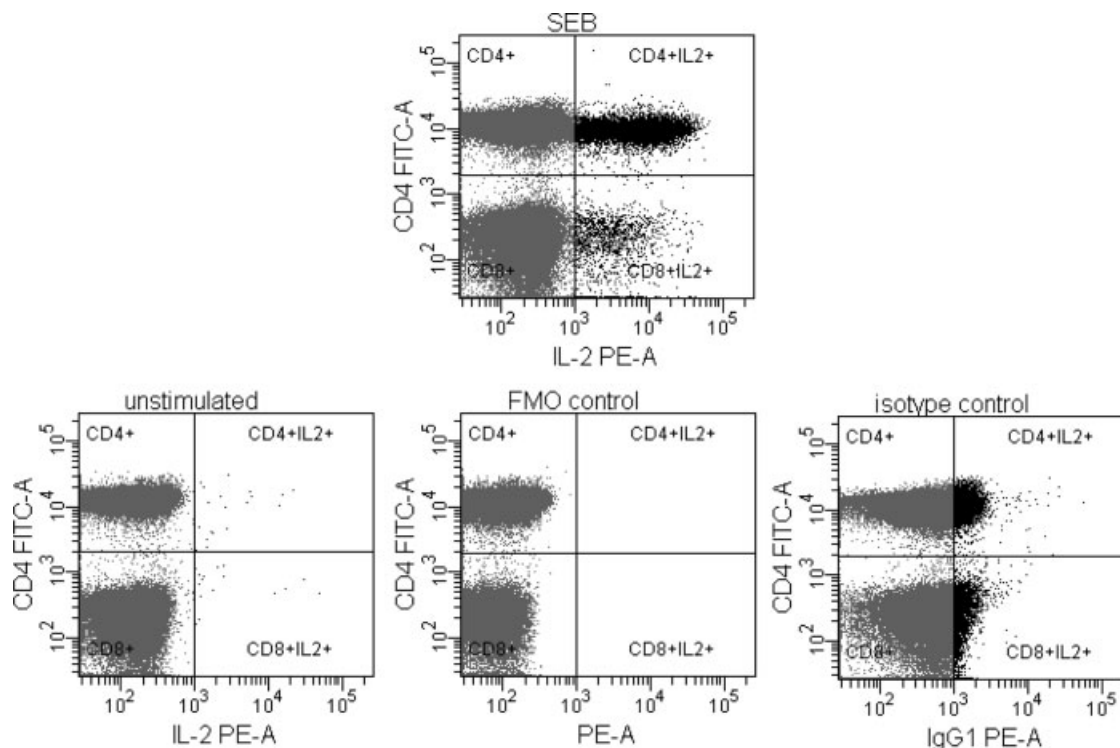
Fig. 3. The best control for a marker stimulated *in vitro* is often an unstimulated (or negative control stimulated) sample, stained with the same antibodies as the test sample. For the response seen in the top panel, the unstimulated control (lower left) best teaches where to place the quadrant markers in the PE dimension. An FMO control (lower middle) or an isotype control (lower right), if used to set the quadrant markers, would set them too low or too high, respectively. Note that no control is necessary to set the quadrant marker in the FITC dimension, since there are two well-separated populations in each panel.

An open question is whether an FMO or isotype control needs to be performed on each specimen used in an experiment. In general, one would not expect significant variability in fluorescence spillover between donors or sample sources, although conceivably different staining intensities in various markers could affect the degree of spillover. Similarly, background staining by an irrelevant antibody ought to be relatively constant between samples, but prudence suggests that there are likely to be sample-specific factors that contribute to variability. Thus, the most conservative approach would be to provide the chosen control for each specimen; but if a large study is undertaken, and this would add significant cost and effort, a pilot study might be used to verify whether there is significant donor variability in controls for the markers being studied.

In summary, two types of gating controls (other than the biological comparison controls discussed below) that are in common use are isotype controls and FMO controls. The former address background due to nonspecific antibody binding, but will not be accurate for all antibodies of a given isotype. The latter address spillover-induced background, but not nonspecific antibody binding. FMO controls are generally more relevant in multicolor experiments where the major source of background variance after compensation is spillover-induced. Neither of these controls needs to be run for every marker in an experiment, but only for those markers where determination of a positive/negative boundary is otherwise ambiguous.

BIOLOGICAL COMPARISON CONTROLS

Some investigators may believe that an experiment is not complete if it does not contain a specificity control, like an isotype control. However, there are often biological comparison controls that are more appropriate for properly setting positive/negative boundaries. For example, in stimulation assays, the unstimulated (or irrelevantly stimulated) sample usually provides the best means to distinguish positive from negative events (Fig. 3). There may be exceptions where high background in the unstimulated sample makes it more difficult to set a clear positive/negative boundary, but in most cases this control is far more relevant than an isotype control or FMO for this purpose. This is because, like an FMO control, the unstimulated control accounts for spillover effects on the channel of interest by including all of the antibody conjugates present in the other test samples. And, like an isotype control, it accounts for nonspecific staining in the channel of interest. But since the same antibody is present as in the other test samples (and the only difference is the sample stimulation), there are no issues of matching the background of the control antibody to the test antibody. Thus, a biological comparison control in this situation is usually the most relevant control for determining positivity of the test sam-

ples. In fact, some groups consider this control so important that they run multiple replicates of it, so that the level of background to be subtracted from the test samples can be determined with greater accuracy.

In addition to unstimulated controls, positive biological comparison controls are also important. In functional assays, these may be samples stimulated with a mitogen, such as SEB, PMA+ionomycin, or a set of peptides with high reactivity among a broad spectrum of individuals (e.g., CEF peptide pool, (5)). The purpose of the positive control is to validate the potential finding of negative results in the other test samples. In other words, it ensures that the experiment was performed in such a way that responses could be detected, and that the cell sample in question was capable of being activated. This protects against misinterpretation of negative data in cases where a reagent was not added, the cells were nonviable, etc. The most appropriate positive control stimulus is thus one for which (1) all individuals to be tested (or at least as many as possible) are positive, and (2) the response is as close as possible in its physiological requirements to the response being measured in the other test samples. A set of positive control peptides may best meet requirement no. 2, but may or may not meet requirement no. 1. PMA+ionomycin meets no. 1 but definitely not no. 2. SEB or an equivalent superantigen is reasonably satisfactory, though not a perfect match for either requirement.

Nonstimulated assays may require biological comparison controls for proper interpretation as well. For example, any comparison of phenotypes between a disease state and healthy individuals should include some of each in every experimental run, if possible. Because of the variabilities of sample preparation, direct comparisons between patients or conditions should generally be run together.

Finally, standardization of longitudinal studies can be facilitated by the use of control cell populations. These might be cryopreserved PBMC from a well-characterized donor, frozen in multiple aliquots, which can be thawed and processed with each experimental run. The data obtained with the control cells can then be required to fall within a certain range in order for the test data to be deemed interpretable. Additionally, the control cells could be used for instrument setup in a variety of ways. Preserved and lyophilized cell standards are also commercially available for these purposes (6).

CONCLUSIONS

One of the most fundamental questions asked in flow cytometry experiments is whether a subset of immunofluorescently stained cells is positive or negative, or what proportion of cells are positive or negative for a given marker. Answering this question accurately requires careful instrument setup and use of controls. First, detector gains should be reasonably set for optimum resolution sensitivity, and compensation should be correctly determined and applied, so as to avoid spurious or unnecessarily ambiguous results. Second, gating controls need to be present for those markers that do not have an obviously separated bimodal distribution. In some cases, a biological control (such as unstimulated cells) provides the best gating control for determining positivity. In other cases, where a biological comparison control is either not possible or not appropriate, an isotype or FMO control may be used. Neither of these is fully optimal, and the choice of which to use may depend upon whether background from dye spillover or from nonspecific antibody binding is of greater concern within the experiment in question. Initially, this might mean running multiple controls to get a sense of which one is most reliable for determining the positive/negative cut-off. In any case, judgement should be exercised so that an obviously inappropriate control is not used simply because there was no consideration of alternatives. Consistent application of thoughtfully determined gating criteria will go a long way toward standardizing the use of flow cytometry to answer clinically important questions.

LITERATURE CITED

1. Chase ES, Hoffman RA. Resolution of dimly fluorescent particles: A practical measure of fluorescence sensitivity. Cytometry 1998;33: 267–279.
2. Maecker HT, Frey T, Nomura LE, Trotter J. Selecting fluorochrome conjugates for maximum sensitivity. Cytometry A 2004;62:169–173.
3. Roederer M. Spectral compensation for flow cytometry: Visualization artifacts, limitations, and caveats. Cytometry 2001;45:194–205.
4. Maecker HT, Rinfret A, D'Souza P, Darden J, Roig E, Landry C, Hayes P, Birungi J, Anzala O, Garcia M, Harari A, Frank I, Baydo R, Baker M, Holbrook J, Ottinger J, Lamoreaux L, Epling CL, Sinclair E, Suni MA, Punt K, Calarota S, El-Bahi S, Alter G, Maila H, Kuta E, Cox J, Gray C, Altfeld M, Nougarede N, Boyer J, Tussey L, Tobery T, Bredt B, Roederer M, Koup R, Maino VC, Weinhold K, Pantaleo G, Gilmour J, Horton H, Sekaly RP. Standardization of cytokine flow cytometry assays. BMC Immunol 2005;6:13.
5. Currier JR, Kuta EG, Turk E, Earhart LB, Loomis-Price L, Janetzki S, Ferrari G, Birx DL, Cox JH. A panel of MHC class I restricted viral peptides for use as a quality control for vaccine trial ELISPOT assays. J Immunol Methods 2002;260:157–172.
6. Ghanekar SA, Maecker HT, Maino VC. Immune monitoring using cytokine flow cytometry. In: Rose NR, editor. Manual of Clinical Laboratory Immunology, 7th ed. Washington, DC: ASM Press; 2005; 353–360.