4 Factors That Affect Research Reproducibility

Improve Reproducibility in Your Lab by Focusing on These Key Sources of Variability

WHITE PAPER

ASTARTE™ Biologics

Getting on with Discovery

The fiery debate over reproducibility in science has burned strong over the past several years, and the flames don't show any signs of dying down just yet. No matter how scientists view reproducibility in their respective fields — from a full-blown crisis to a minor issue that doesn't impact the credibility of published findings — the significance of reproducible experimental design is undeniable.

All fields of science, immunology included, were built on a foundation of replicating experiments and studying breakthroughs for expanding the knowledge base, inspiring more experimentation, and ideally, leading to additional breakthroughs. Without reproducibility, the flywheel of experimentation and discovery can't gain momentum.

According to a *Nature* survey of 1,500 scientists, more than 70% have tried and failed to reproduce the results of another scientist's experiment, and perhaps more shocking and concerning, over 50% have failed to reproduce their own experiments.[1]

An Answer to the Problem?

Widespread reproducibility mandates have been debated and undoubtedly would be costly, time-consuming and difficult to implement and enforce across the board. While we await an industry-wide standard, if one ever comes, there are many things we can take responsibility for in our labs and research practices to improve reproducibility.

We've outlined four of the primary sources of variation in experimental results and provided tips and research examples of how we're improving reproducibility in our immunology and inflammation research lab.

4 Factors That Affect Research Reproducibility

1. **Reagents**
2. **Equipment**
3. **Personnel**
4. **Methods**

Factor #1: Reagents

Antibodies and Cells

Verifying the quality of reagents used in research can prevent wasted time, wasted resources and unverifiable results. Unvalidated reagents can snowball down the research chain, leading to retracted publications and failures in pre-clinical and clinical trials. Front-end validation, while time-consuming, is an important step that many researchers choose to skip without thinking through the potential consequences.

Antibody quality control starts with procurement. Look for a vendor that offers detailed product documentation, has a substantial list of research publications using their products and is comfortable providing customer references.

The vendor should be able to provide test results to show that the antibody binds to its target and not to related proteins. This negative control is as important as a positive control since cross-reactivity is possible even with monoclonal antibodies.

Check every product's Certificate of Analysis before making a purchase to confirm the purity and viability of the cells and ensure the product and donor specifications match your experimental requirements. A good Certificate of Analysis will be lot-specific to reflect testing of that particular lot, not just a general product description.

Look for details like:

- Number of live cells per vial
- Percentage of viable cells
- Expression of cell surface antigens (types of cells in the vial)
- Cell purity
- Donor age
- Donor gender
- Donor race
- Donor height
- Donor weight
- Donor blood type
- Donor HLA type

Regardless of the product information given by a vendor, always do your due diligence by testing your antibodies and cell products before starting your experiments to account for discrepancies in the analysis and changes during shipping, handling or storage. Test your products using the technology you plan to use in your experiment. If you will be using antibodies for immunohistochemistry (IHC), test in IHC.

Culture Media

Culture media is an easily overlooked variable in experimental design as researchers typically use their favorite media or the one they deem most appropriate for the experiment at hand. However, differences in culture media environment can have a major impact on cell behavior and thus cause reproducibility issues.

The data in figures one and two show the effect of culture medium on monocytes. Human monocytes were cultured in either DME/F12 or IMDM for five days. Both media were supplemented with 10% human serum and 10 ng/mL recombinant human M-CSF. They were then stimulated by the addition of 100 ng/mL LPS and culture medium was collected after 48 hours. The monocytes cultured in DME/F12 made more IFNγ, IL-10, IL-6 and TNF-α, but the production of IL-13 was equivalent in the two media.

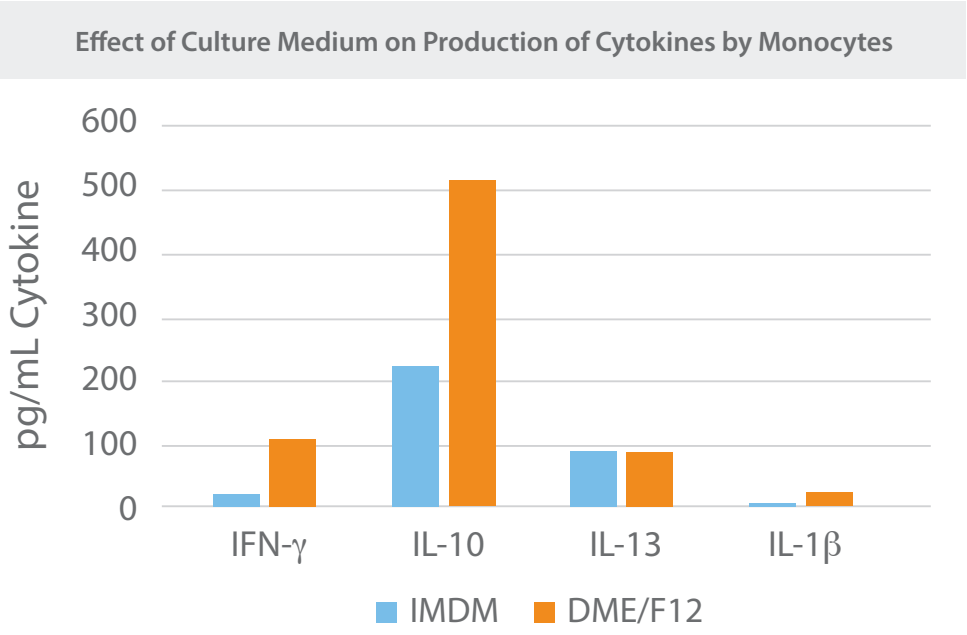


Figure 1
IFNγ, IL-10, IL-13 and IL-1β production measured after culture of monocytes in DME/F12 and IMDM.

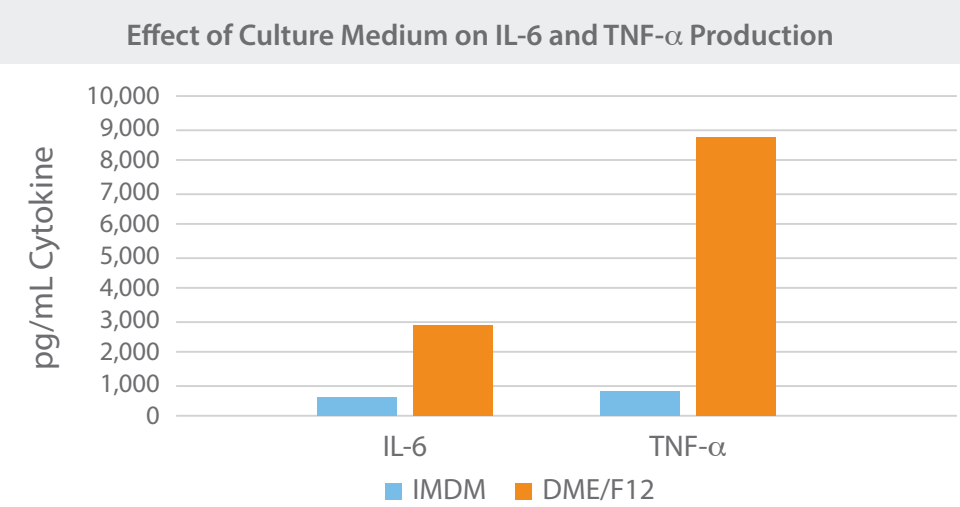


Figure 2
IL-6 and TNF-α production measured after culture of monocytes in DME/F12 and IMDM.

One way to reduce this variability is to use the same culture media as described in the original experiment when attempting to replicate a research finding. However, not all culture media are created equal.

When using serum-containing media, there are natural variations in each animal's genetics, blood, environment and diet that create extremely high lot-to-lot variability. Acquiring enough serum from the same lot is challenging enough for a single researcher, and nearly impossible for a scientist in another lab to acquire.

So what can you do? While fetal calf serum (FCS) is a popular medium supplement, it does fall victim to extreme variability. There are several serum-free media on the market, which can be great fits for your immunology experiments and have far less variability. If FCS or other serum-containing media are required, focus on providing better documentation of your serum selection and lot characteristics.

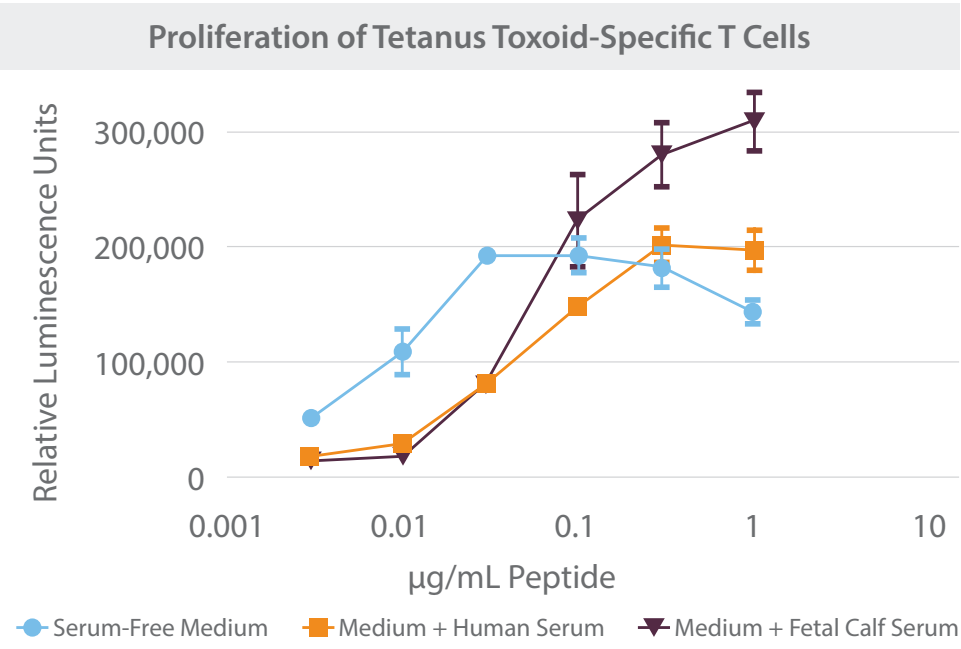


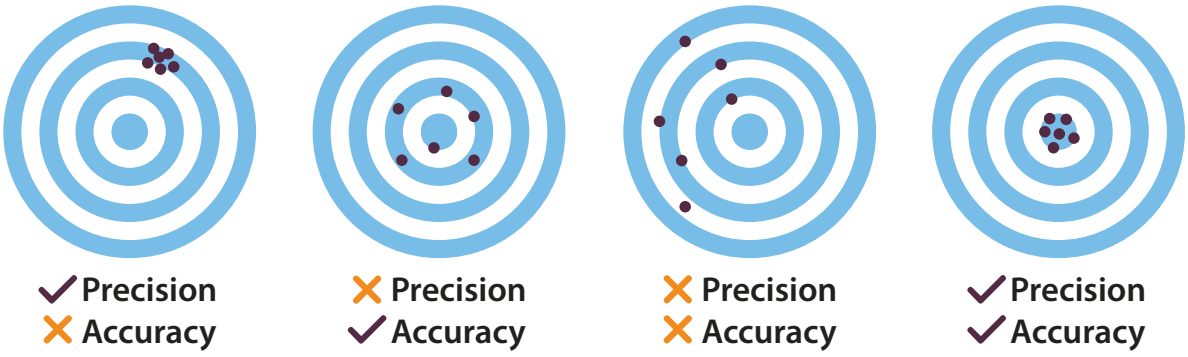
Figure 3
A medium designed to be used without serum was supplemented with human serum or with fetal calf serum and used to support a T cell proliferation assay. The medium supplemented with FCS supported greater proliferation at the highest peptide concentration. The medium that was not supplemented allowed greater proliferation with low peptide concentrations.

Factor #2: Equipment

Calibration

Lab equipment and instruments lose their calibration over time, which can lead to unreliable and imprecise results. An often overlooked component of responsible experimentation is making sure all equipment is calibrated and capable of measuring its intended results with accuracy, precision and safety.

Precision vs Accuracy



Following a proper calibration schedule will make your results more reliable and reproducible. But how do you know if your equipment is calibrated? And how often should you check your equipment for accuracy?

Everything from your largest, most expensive equipment to your smallest instruments should be calibrated consistently to ensure proper measurement. Reports should include the calibration at the outset of measurements as well as after adjustment. This habit will help inform your calibration frequency and let you know if your instruments are off. Follow these suggestions for some popular lab equipment.

Equipment	Calibration Frequency	Suggestions
Biological Safety Cabinets Incubators Centrifuges	Manufacturer-recommended calibration intervals or before and after major experiments	Hire a metrologist to do a professional, NIST-traceable calibration
Pipettes	Biannual or before major experiments	Test the full range of volume the pipette can dispense
pH Meters	Before every use	Refer to the original manual for exact calibration standards

Variation

Two different pieces of the same equipment are going to differ — whether they are the same model and year, or two models from different decades. All equipment wears differently, even if used for the same purposes over the same time period in the same lab conditions.

Figures four and five show the difference in measuring luminescence using two microplate readers, a Packard Fusion™ and a BioTek Synergy™ HT. The proliferation values as measured by the Packard are lower but more consistent within triplicates as shown by the smaller error bars based on standard deviation. The BioTek plate reader returned values that were higher but also had a greater standard deviation and coefficient of variation.

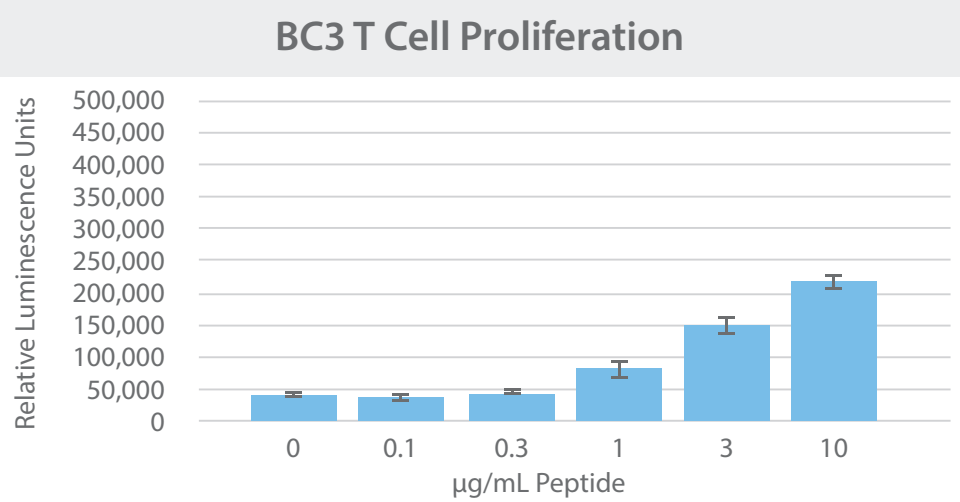


Figure 4
BC3 T cell proliferation using a Packard Fusion™ microplate reader to measure luminescence.

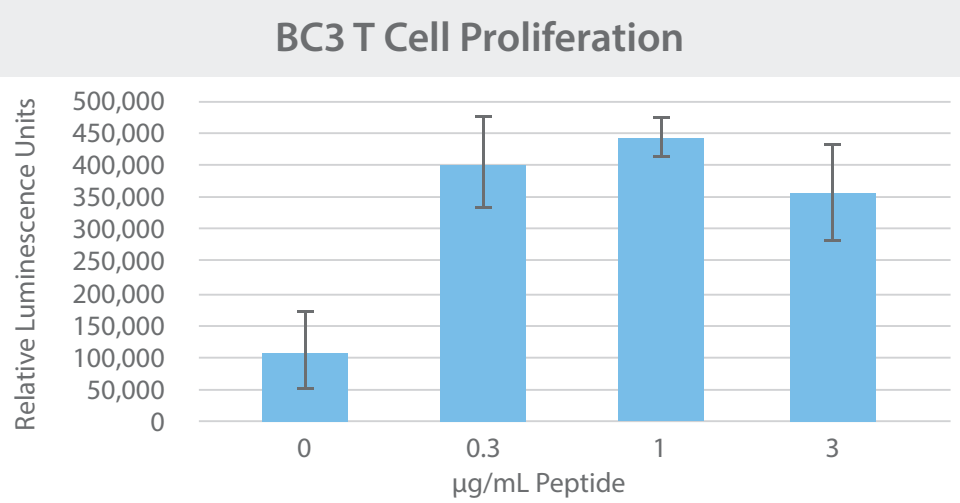


Figure 5
BC3 T cell proliferation using a BioTek Synergy™ microplate reader to measure luminescence.

The first step in controlling machine-to-machine variation is to follow a consistent calibration schedule. Beyond calibration, there are a few things you can do to reduce the impacts of equipment variability.

1. **Contain your experiments to one set of equipment.** If your lab has two incubators, for example, keep all of your samples for one experiment in the same incubator. This will eliminate any variability from jumping back and forth between machines throughout the study. We even see variation in evaporation from the front to the back of the incubator, so keep positioning in mind.
2. **Provide thorough documentation.** Always note which equipment you used in your methods, from the brand, model and serial numbers to the date of last calibration and any other pertinent information. Since equipment is a likely source of variation between labs, it should be noted.
3. **Understand the idiosyncrasies of each machine.** Realize that each machine can give slightly different results. For example, a used microplate reader for luminescence may give lower values than a new machine. Understand these differences and know how to account for them within and between experiments.

Also be selective when making new equipment purchases to be sure you're getting the best available product for your budget from a reputable vendor that offers a reasonable warranty.

Buying used equipment can be a major cost savings if you know what to look for. When evaluating previously owned lab equipment, do your research and ask the right questions.

- ✔ How long has the equipment been in use and in what capacity?
- ✔ Does the machine have a history of damage or refurbishment?
- ✔ What are the terms of the warranty? Even used equipment should have a warranty.

Factor #3: Personnel

Training and Mentorship

Ensure every researcher under your management is adequately trained in experimental design. Provide training and establish design standards to which everyone must adhere.

Every scientist has an individual style when executing even the simplest procedures, like pipetting liquids, mixing solutions or counting cells. Draw attention to those differences by comparing results and discussing the reasons for any discrepancies. This hands-on approach, coupled with basic lab-wide standardization guidelines, will help your scientists appreciate the proper procedures.

Beyond training, put in place an expectation and foundation of strong mentorship from senior researchers. This is usually built into the hierarchy in larger labs. In smaller research labs, mentorship needs to come from the top. Owners, managers or experienced scientists should make a concerted, defined effort to understand the professional goals of each researcher, identify their strengths and opportunities for improvement and set aside time to work with each individual to advance their knowledge, skills and confidence.

Researchers who feel valued and are given the tools they need to succeed will be more fruitful and loyal in the long term.

Statistical Analysis

Researchers too often fall into the trap of sensationalizing conclusions with the goal of publishing their works in high-visibility, prestigious journals. This can lead to cherry-picking results that, while promising or interesting, may not be reproducible.

If you have a sound hypothesis, your experiments are properly designed and executed and you stick to the original intent of your research, your results should speak for themselves and garner the attention they deserve.

In *Nature*'s survey, nearly 90% of respondents endorsed "better understanding of statistics," "better monitoring/supervision" and "more robust design" as the top approaches to boost reproducibility in science.

"Being at the cutting edge of science means that sometimes results will not be robust. We want to be discovering new things but not generating too many false leads."

Marcus Munafo, a biological psychologist at the University of Bristol, UK

Factor #4: Methods

Variation in Cell-Based Assays

Cell-based assays, when developed and executed correctly, are an important tool in the drug discovery process that can confirm how a drug candidate interacts with and responds to an organism. It's important to follow the assay instructions closely and use the materials and methods exactly as described, otherwise risk errors or unreliable outputs.

Cell-based assays have a much higher potential for variation than a chemical assay due to the nature of the immune system. Understanding the degree of variability in a cell-based assay and the sources of it will help researchers account for inherent variations in the assay to avoid improper conclusions.

Here are three common factors that should be carefully controlled to reduce variability in any cell-based assay.

- **Selecting Cells**
  Cell type will affect the conditions and outcomes of the assay. There are major differences between primary cells and cell lines that make each suitable for different types of assays. Follow the kit instructions or protocol carefully and procure the ideal cell type to ensure proper measurement.

- **Pipetting Accuracy**
  Pipetting errors are very common and can lead to false data. It's important to calibrate your pipette instruments and follow proper techniques to ensure even distribution in all wells.

- **Controlling Edge Effects**
  Edge effects are errors resulting from uneven distribution of cells in a well plate, where cells in inner wells are exposed to different conditions than those in outer wells. The causes are many, including evaporation during incubation, temperature control and plate stacking. Researchers have concluded that simple changes, such as pre-incubation of newly seeded plates at room temperature, can reduce edge effects.[2]

Inconsistent Methods

Methods and measurement should be carefully considered and tailored to each experiment. Any changes in methods or measurement techniques can yield inconsistent results, whether in your lab or in others.

One common example of inconsistent methods impacting results is measuring cell proliferation. When using CellTiter-Glo®, a reagent based on ATP content, increased cell metabolism and higher ATP per cell is interpreted as proliferation. However, increased cell metabolism can also increase ATP per cell, which may complicate interpretation.

If measuring proliferation with the uptake of 3H-thymidine or bromodeoxyuridine, which label cells replicating their DNA, only the cells in S phase will be labeled. These two methods can produce vastly different results, especially with immune cell populations in which a small subset may proliferate with the majority simply maintaining their metabolism.

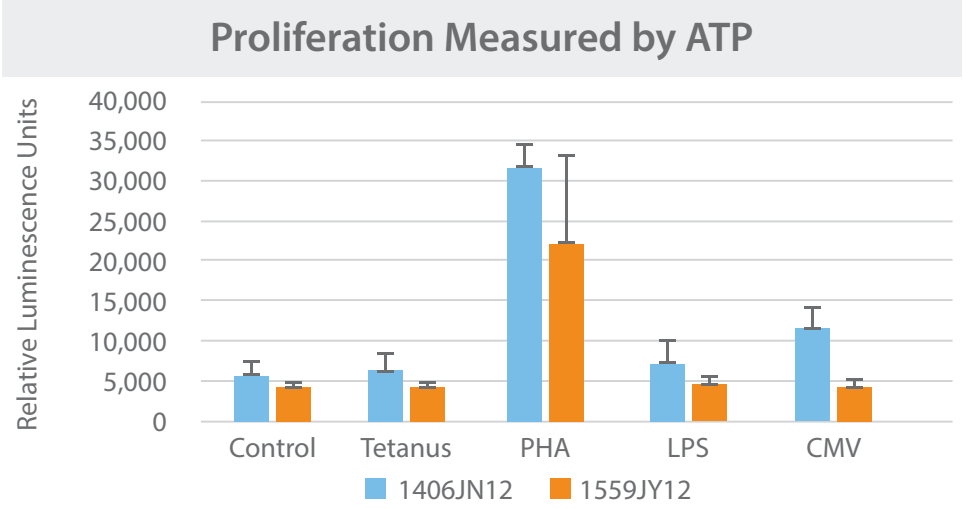


Figure 6
PBMC proliferation measured by ATP content using CellTiter-Glo®.

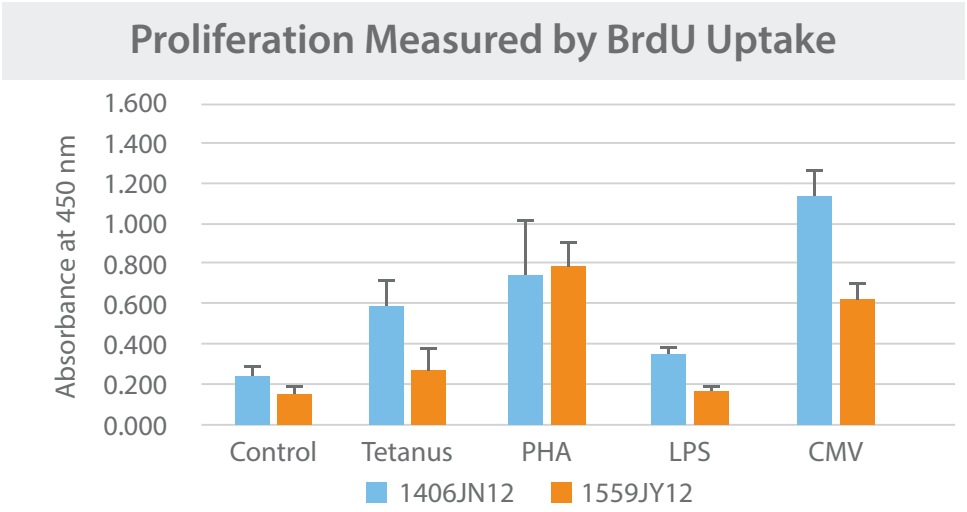


Figure 7
PBMC proliferation measured by uptake of bromodeoxyuridine.

We also see differences in the amount of cytokines measured using ELISA versus the Meso Scale method, as a second example. This can be traced, in some cases, to the nature of the cytokine used to prepare the standard curve and the accuracy of its concentration.

Another variable that should be carefully selected is the analysis software. While several tools may be designed with the same basic principles and analytic concepts, each could return different results when analyzing the same dataset. Flow cytometry hardware combined with software can increase or decrease sensitivity of detection for rare cell types. Consider using more than one type of software to get a better picture of your full dataset.

Beefing up your design standards, method selection and measurement techniques will not only increase your intra-lab reproducibility but should help outside scientists replicate your findings as well.

Conclusion

Depending on the size of your lab and type of research you perform, sweeping overhauls in the way you run your lab and design experiments may or may not be necessary. Only you can be the judge of your needs.

However, you should be conscious of your responsibility to publish reproducible findings or deliver valid data to your customers. Even implementing some of the suggestions mentioned above could give your research a significant boost in reproducibility.

References

[1] Baker M. 1,500 Scientists Lift the Lid on Reproducibility. Nature. 2016.

[2] Lundholt B, Pagliaro L, Scudder K. A Sample Technique for Reducing Edge Effect in Cell-Based Assays. J. Biomol. Screen. 2003; 8: 566-70.

ASTARTE™ Biologics

Getting on with Discovery

866-513-1894 | www.astartebio.com