The Advantages of Third-Party Testing

by Don Mays, Product Safety Insights LLC

Executive Summary

Engaging independent third-party laboratories to qualify a product's integrity is the gold standard for manufacturers. Third-party lab testing can eliminate confirmation biases that can lead to wrong conclusions, and can help provide a company with needed assurances before launching a new product into the marketplace. In addition, data from qualified third-party labs can help defend a company against product liability issues and government agency investigations. But failure to develop a comprehensive testing program with competent laboratory can result in severe consequences such as recalls, reputational damage, and lost sales for the company and its customers.

Due diligence is critical when selecting a third-party lab. Ideally, the lab should be accredited, and have the experience and qualifications to conduct the desired testing. It's not a one-size-fits-all situation. While some labs may be competent at certifying products against certain standards or at conducting quality assurance inspections, it may be necessary to work with different labs for evaluating performance or usability.

Manufacturers should be aware of natural variability in test results and should design enough margins into the product to compensate for that natural variability. Reality checks are also important to make sure that test data makes sense and would be predictive of real world experience under reasonably foreseeable use conditions.

Due diligence doesn't stop with a passing test result. **Continuous monitoring of product integrity and using good data management practices can help companies ensure their continued success.**

Introduction

One of the most exciting milestones for innovative companies is the day their new product is launched in the marketplace. It often represents the culmination of years of research and development, preparing multiple patent applications, writing manufacturing specifications, developing marketing and advertising campaigns, and managing distribution logistics, among the multitude of steps in the process of bringing a new product to market. Behind each step there is usually a team of people who have become emotionally invested in their part of the process and the launch overall. The executives in the C-suite are perhaps most emotionally invested, and financially invested, since they have to set the goals, make promises, and answer to investors and cutomers if promises aren't met.

Far too often, something goes wrong. The product may be a flop with buyers, real-world performance may not meet customer expectation, or supply chain issues could prevent timely delivery. Worst case—a serious safety problem is uncovered once the product is in consumers' hands, resulting in injuries, humiliating—and expensive—recalls, and future product liability lawsuits.

Recalls are far more than a daily occurrence. The chart below shows the number of recalls announced in 2021 by federal agencies, including the Food and Drug Administration, the U.S. Department of Agriculture, the U.S. Consumer Product Safety Commission, and the National Highway Traffic Safety Administration. In total, more than 2,500 recalls were announced last year, covering more than one billion units.

Number of Recalls in US in 2021[1]

Medical Devices	837
Pharmaceuticals	274
FDA - Food & Beverage	414
USDA - Meat & Poultry	47
CPSC - Consumer Products	218
NHTSA - Vehicles, etc.	770
TOTAL	**2560**

[1] Data source: Sedgwick Brand Protection – Recall Index

But federal agency-announced recalls typically involve only products that pose a health or safety problem, or don't comply with a regulation or standard. There may be other "secret recalls" handled by manufacturers for quality or performance issues. Those recalls are rarely made public and are generally orchestrated between the manufacturer and the retailer of the product.

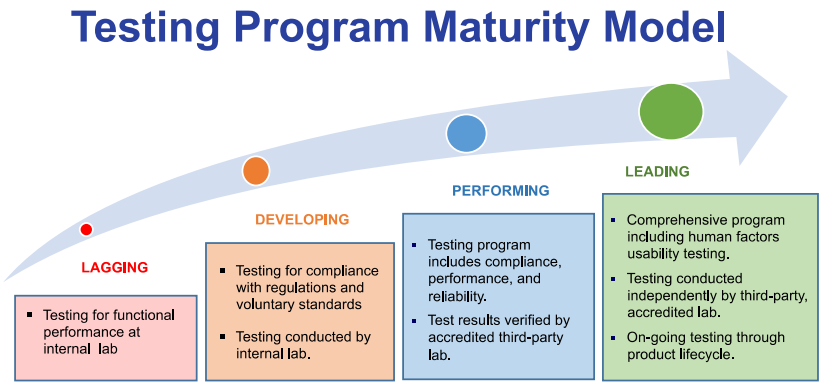
So why do we see so may recalls? In many cases, it's due to the lack of adequate premarket testing of the product. While testing is a critical stage in the research and development process, there are many reasons manufacturers get it wrong. This paper explores some of the ways testing can be can be mishandled.

Comprehensive Product Testing

Product testing takes on many forms. Some companies look at testing as just a method to confirm a product will perform its intended function. Others may view testing as a way to certify that the product conforms to all applicable regulatory requirements or safety standards. Quality assurance professionals will look at testing as a way to ensure production lots continue to meet their manufacturing specifications. Human factors scientists will sometimes test using a panel of users to determine any shortcomings in the product's design. A comprehensive testing program includes all those elements and more.

Below is a testing program maturity model that shows how companies can move from lagging practices to leading practices. Leading practices include testing for compliance,

safety, performance, quality, usability, and sustainability. Much of that testing is best handled by independent, third-party, accredited laboratories. In addition, ongoing testing through the product's lifecycle can capture deficiencies that might not be captured on newly produced products.



A robust testing program begins with ideation of the product and continues through its lifecycle. At the beginning of product development, designers often set up a stage gate process that requires certain tests to be passed before proceeding to the next stage of development. It's not until the first prototype of the product is produced that more thorough testing can be executed. But the further along the design path, the harder it is to make changes … and the more invested the development team and company becomes. Testing at that point is often designed simply to confirm expected results. A critical test failure at the prototype stage sometimes results in disbelief, questioning, and blame. It's a human tendency to place more weight on preconceptions than on information that contradicts those prior beliefs. This is called confirmation bias.

Confirmation bias replaces

critical thinking with wishful thinking

It is strongly recommended that manufacturers get an independent opinion where confirmation biases or emotional attachment may exist. Frequently, manufacturers set up their own in-house testing labs to qualify their products. In my experience, most of those labs do not follow Good Laboratory Practices[2] and are not accredited by outside agencies to perform the necessary testing.

What is third-party testing?

One might wonder about how first-, second-, and third-part testing differ. Test data developed by the producer of the product would be considered first-party data. It comes with risks of confirmation bias and possible lack of competence. Second-party test data can be developed by the first-tier customer of the product. For example, if a factory under contract produces a product for the brand manufacturer, data developed through testing by the brand manufacturer are considered second-party data. Third-party testing involves

[2] "Guidelines for Industry, Good Laboratory Practices – Questions and Answers"; U.S. Department of Health and Human Services/Food and Drug Administration/Office of Regulatory Affairs, June 1981

testing by an independent laboratory and is the gold standard to support conformity assessment, including compliance with regulations and standards. The ISO/IEC 17000 standard provides the following definitions:[3]

First party *– a conformity assessment activity that is performed by the person or organization that provides the object*

Second party *– a conformity assessment activity that is performed by a person or organization that has a user's interest in the object*

Third party *– a conformity assessment activity that is performed by a person or body that is independent of the person or organization that provides the object, and of the user interests in that object*

While internal or in-house testing may be sufficient for qualifying a product through the stages of research and development, it can fall short of providing a comprehensive unbiased evaluation of the product's integrity. Product integrity includes at least five interrelated elements:

- Compliance with regulations and standards;
- Safety for the end user throughout the product's life;
- Performance of the product to function as intended and expected by the consumer;

[3] ISO/IEC 17000 – Clauses 2.2, 2.3, and 2.4

- Quality in regard to reliability, durability, and freedom from defects; and
- Sustainability in regard to longevity and environmental impact.

In my experience, many companies focus testing on compliance with regulations. But compliance is just the starting place, since it simply provides the license for bringing a product to the market. It provides no guarantee of safety, performance, quality, or sustainability.

I recently studied CPSC (U>S> Consumer Product Safety Commission) recalls from 2016 through 2020, and found that only about 10 percent of the published recalls were for products that violated a regulation, standard, or rule. The great majority of recalls were conducted because a safety hazard was identified after the product went to market, sometimes resulting in injury to the consumer.

This research clearly shows the need to expand a test program well beyond that of compliance testing. It is likely that the manufacturers of the recalled products did not have a comprehensive testing program. Whatever testing was undertaken may not have been conducted by a qualified, independent laboratory. The need for independent, third-party, comprehensive, and competent testing is essential to ensuring the product's integrity.

Human factors analysis

One undervalued area of a product-testing program is human factors analysis. Products need to be safe under all reasonably foreseeable use conditions, and understanding how a product might be used requires special skills. Laboratories that specialize in conformity assessment may not have the expertise for human factors analysis. However, some third-party laboratories have specialists who understand human behavior and can be effective at identifying issues that may not be discovered otherwise.

The lack of comprehensive human factors analysis has resulted in a crisis for many companies. Fitness equipment maker Peloton, for example, suffered reputation damage and wound up recalling their treadmill after it was discovered that young children could be pulled underneath a moving treadmill belt. One child died and several were injured. Fisher-Price recalled their Rock 'n Play inclined sleeper after dozens of infant deaths were associated with its use. In both cases, a serious hazard was identified after the

product went to market, perhaps because the manufacturer didn't fully understand how the product might be used. Plus, they may have placed overreliance on warning labels and instructions, which are often not read or heeded by consumers. Comprehensive human factors analyses are effective in uncovering safety or usability issues that are not detected under common testing regimes.

Qualifying laboratories

It is not always easy to find laboratories that are qualified to test unique products. While some laboratories specialize in routine testing for common products, they may stretch beyond their limits of competency to accommodate a client's specific needs. It is important to thoroughly assess a lab's competency before engaging it to do testing. Depending on the extent of the test program, it may be necessary to work with several laboratories with different capabilities.

Some regulators require that testing for compliance with regulations is conducted by approved or accredited laboratories. There are several accreditation bodies in the U.S. The most widely recognized are the American Association for Laboratory Accreditation (A2LA), ANSI National Accreditation Board (ANAB), and the National Voluntary Accreditation Program (NVLAP).[4] Those organizations are members (signatories) of ILAC, the International Laboratory Accreditation Cooperation. Accreditation bodies will assess independent laboratories for compliance with the requirements specified in the

[4] https://www.nist.gov/standardsgov/accreditation-bodies-mras

ISO/IEC 17025 standard. While this is a complex scheme, leading practices dictate using an ISO/IEC 17025 accredited laboratory for critical conformity assessment purposes.

When qualifying independent laboratories, it's important to make sure that their accreditations extend to the type of testing needed by the manufacturer and that their accreditations are up to date. It is also important to assess that the laboratory personnel selected to conduct the testing have been adequately trained.

The location of the laboratory may also be a factor, particularly when testing large, difficult-to-ship items. It is sometimes more expedient to use laboratories that are geographically close to the point of manufacture.

Turnaround time and cost may also be critical factors for selecting a third-party lab. Testing costs often vary by country. It's usually less expensive to have testing done in Asia than in the U.S. While that may save money, it presents a challenge to U.S.-based companies to qualify a lab or to witness testing when questions arise.

Developing the contract

Engagements with third-party laboratories should be handled through a legally binding contract. The contract should detail the scope-of-work, the expected deliverables and timeline, and, of course, the cost. When turnaround time is critical, it may be possible to

insert a penalty clause for missed deadlines, allowing the client to reduce the payment to the lab if the schedule is not followed.

Indemnification clauses are often a point of contention when negotiating a contract. If a laboratory makes an error in their testing by missing a safety hazard per se, and that later results in a recall, the manufacturer may try to hold the laboratory financially responsible. Independent laboratories are reluctant to agree to such liabilities, but they sometime acquiesce to keep in the client's good graces.

Test sample selection

Selecting representative test samples is a critical part of the testing process. For premarket qualification testing, it is critically important to ensure that the samples selected are representative of the full run production. That often means selecting random samples off the production line or shipping pallets.

Typical quality assurance inspections involve sampling at four different points: pre-production, first-run production, in-process production, and at pre-shipment. In many cases, post-shipment samples are also inspected. Samples are selected randomly and evaluated for critical, major, or minor defects. A critical defect should stop production. Particularly when products are produced overseas, quality assurance inspections are very often conducted by third parties. The more qualified labs that do this work apply

ANSI/ASQ quality assurance standards[5] for defining acceptance criteria. Conducting this work independently prior to shipping prevents disruption when defective products are found on the receiving end of the shipment.

Independent laboratories are often in fear of testing "golden samples" – ones specially made to ensure they pass all the tests. That is why organizations such as Consumer Reports buys most of the products they test on the open market instead of getting them directly from the manufacturer. However, that doesn't mean that the samples that they do test are representative of typical production; it's always possible that a "lemon" ended up in the lab. However, a manufacturer's good quality assurance program should ensure that each and every sample will provide the desired test results.

Repeatability and reproducibility are also concerns when selecting and testing products. Repeatability, in this case, refers to getting consistent test results within the same laboratory. Reproducibility refers to getting the same test results at different laboratories. It's important to recognize naturally occurring variability from sample-to-sample, test-to-test, and lab-to-lab. Enough margin of acceptance should be planned for the product's design to compensate for normal variations that might provide different test results.

Verifying test results

[5] ANSI/ASQ Z1.4-2003 (R2018): Sampling Procedures and Tables for Inspection by Attributes, and ANSI/ASQ Z1.9-2003 (R2018): Sampling Procedures and Tables for Inspection by Variables for Percent Nonconforming

Manufacturers are well advised to conduct a reality check when reviewing test results. Do the results make sense? Would the same results be expected each and every time? Would the product be expected to perform the same way in the real world as it did in the laboratory? Are there mistakes in the lab report? Relying on a single test result could be a dangerous practice. And to reiterate a previous point, it is important not to let confirmation bias interfere with critical judgment. It always important to have an independent verification process to ensure the test results are right.

"Trust but verify"	*"Distrust and verify"*
--Ronald Regan	-- Hillary Clinton

The recent trial of Theranos founder Elizabeth Holmes uncovered not only fraud but also the investors' lack of due diligence. The much-ballyhooed Theranos blood-testing machine was claimed to run dozens of diagnostics with just a single drop of blood. This was too good to be true of course, and investors lost millions of dollars. Appropriate independent, third-party testing may well have uncovered the fact that the machine did not work as claimed.

Test data management

A good lawyer will advise companies to make sure that their test data are protected from legal discovery. While that might not always be possible, it's important to be aware that

test data can be used against a company in the event of a product liability lawsuit or government agency investigation. In many cases, sensitive data should be labeled "Privileged and Confidential" and sent to the legal team, which provides a certain level of protection. However, in the converse, testing that was developed by independent, third-party labs can be used to defend a company by demonstrating that the company has exercised due diligence.

Managing mountains of product data is a very labor-intensive task. Engineering drawings, change notices, laboratory test data, warranty claims, and customers service data often reside in disparate, non-relational databases. Leading practices includes setting up centralized data repositories with access restricted to those with a need to know. More advanced processes include "big data" analytics that make the practice of mining product data much more efficient.

Continuous monitoring

Product safety and quality assurance practices should not stop at the point of product launch. It is critically important to continue to monitor how the product performs in the real world. Capturing and analyzing consumer complaints early in the life of the product can allow a manufacture to tweak the design and drive better customer satisfaction. Similarly, monitoring and analyzing warranty claims may also help to identify problems with a product early in its life.

Mining social media has become a common practice to measure how consumers respond to a new product. Consumers are quick to complain on social if they are dissatisfied. Websites that allow consumer reviews can paint a different picture; ratings are not always posted by verified purchasers and may have been seeded by the manufacturer. Plus, marketing teams know that asking a consumer to provide a review of the product shortly after it is purchased is likely to result in a more positive review than what the consumer might write after using the product longer and discovering all its shortcomings.

But hidden in all that data are nuggets of information that can be used by a manufacturer to make product improvements and immediately address safety and quality concerns before they become a crisis. Government agencies, including the CPSC, expect companies to monitor social media and all other channels of information post product launch to make sure that safety issues are captured and reported to the agency on a timely basis, as required by law.

Using third-party labs for forensics

Even the most diligent companies will sometimes make mistakes. Thing can and will go wrong. When they do, it's not always easy to identify the root cause of the problem. That may be due to the lack of technical expertise to diagnose a problem, or the lack of will to admit that there actually is a problem with the product. That's where third-party forensic labs can help. Since they have no emotional or financial attachment to the product, they may be able to identify root causes without the biases that could drive wrong conclusions.

Selecting a third-party forensic laboratory is more difficult than finding a third-party lab for conformity assessment. Like medical doctors, forensic labs have specialties. It's important to fully understand a forensic lab's capabilities and experience before engaging it to conduct a root cause analysis.

Before I was hired by Samsung Electronics America, the company suffered through one of the most noteworthy and embarrassing recalls. Their newly launched Galaxy Note 7 smart phone had defects that resulted in battery fires. Samsung worked with three different forensic labs to quickly identify the problem, which eventually resulted in a major recall and subsequent improvements in their testing program.

Conclusion

There are many advantages to using third-party labs through various stages of a product lifecycle. Third-party data can help validate internal test data, provide needed assurances before product is launched, protect a company in the event of a product liability lawsuit or government investigation, and provide independent root cause analysis when something goes wrong. Using third-party labs also demonstrates a high level of due diligence to ensure product integrity. But selecting the right lab or labs is a critical first step. It's important to qualify labs before engaging them to conduct any testing and to verify that have the proper experience and accreditations to perform the desired work.

Don Mays is Founder of Product Safety Insights LLC, a consulting enterprise that advises companies on product safety assurance. He has a rich background of experiences that included serving as Chief Safety and Quality Officer at Samsung North America, and Managing Director, Product Safety and Quality for Deloitte. For nearly 20 years, Don worked for Consumer Reports, most recently as its Senior Director, Product Safety and Technical Policy. He was also Vice President of Retail and Consumer Product Services at Intertek, as well as the Technical Director for the Good Housekeeping Institute. Don is currently President of the Society of Product Safety Professionals, Chairman of ASTM's Committee on Consumer Products, and a board member for Kids In Danger.