Westgard QC Order Form

Westgard QC, 7614 Gray Fox Trail, Madison WI 53717

CALL 1-608-833-4718 if you wish to pay by purchase order or other means.

Item	Price (US$)	Quantity	Subtotal
NEW! Basic Method Validation, 4th Edition	**$80.00**		
Most Popular! Basic QC Practices manual, 4th Edition	**$80.00**		
Six Sigma Quality Design & Control, 2nd Edition	$90.00		
Six Sigma Risk Analysis and Tools	$90.00		
Poor Lab's Guide to the Regulations, 2019 Edition	$80.00		
Basic Quality Management Systems	$80.00		
Westgard Online Courses			
NEW! Six Sigma Metrics	$195.00		
Quality Management and Design of Analytical Systems (5 credits)	$75.00		
"Westgard Rules" and Levey-Jennings Charts mini-course (3 credits)	$75.00		
Most Popular! Basic QC Practices – complete online course (14 credits)	$135.00		
Basic Method Validation – complete online course (15 credits)	$175.00		
		Grand Subtotal	
		Shipping & Handling: Within US = $9 Canada & Mexico = Add 10%; Europe, Asia, & all other countries = Add 20%	
		Sales Tax (Add 5.5% in WI, 6.0% in CT)	
		GRAND TOTAL	

Visit http://www.westgard.com/store.htm to place order online
Use coupon code SKIPTHEFORM to save $15 off any purchase

Your Name					
Institution & Department					
Street Address					
City		State		Zip Code	
Country					
Business Phone/Business Fax					
E-mail Address					
Credit Card Type (circle one)	VISA Mastercard American Express				
Credit Card Number				Exp. Date	
Signature					

Basic Method Validation

4th Edition

James O. Westgard, PhD

with contributions from

Nils Persons, PhD
Sten Westgard, MS

Westgard QC, Inc.
7614 Gray Fox Trail
Madison WI 53717

Phone 608-833-4718 Fax 608-833-0640
http://www.westgard.com

WESTGARDQC

Library of Congress Control Number: 2020941821
ISBN 1-886958-33-5
ISBN13: 978-1-886958-33-3
Published by Westgard QC, Inc.
7614 Gray Fox Trail
Madison, WI 53717

Phone 608-833-4718

Preface to the 4th edition

In 2020, the first full year of the SARS-COV-2 pandemic, hundreds of methods flooded the market to address the urgent demand. At one point in June, for every method that qualified for emergency use authorization (EUA), the FDA was disqualifying two methods (adding them to the "do not distribute" list). Thus, in a market where a majority of methods were of dubious quality, the need and the skills to evaluate the analytical quality of a test became paramount (again).

Basic Method Validation is part of a trilogy of "back to basics" books that focus on analytical quality management. The other two books are *Basic QC Practices* and *Basic Quality Management Systems*. When I teach these materials, I start with method validation because it introduces the basic concepts of analytical performance and the experimental and statistical techniques needed to describe performance in quantitative terms. These concepts carry through into the practice of QC and the selection of optimal QC procedures via quality design and planning.

The original source of this approach to method validation goes back nearly 50 years to a series of papers that were first printed in the *American Journal of Medical Technology* and later published as a monograph titled *Method Evaluation*. My co-authors were Diane J de Vos, Marian R. Hunt, Elsa F. Quam, R. Neill Carey, and Carl C. Garber, all of whom worked at the University of Wisconsin. We introduced this approach at workshops that were taught at the national ASMT and AACC meetings. David Koch, a past president of AACC, continues to teach this workshop at the annual AACC conference. This workshop holds the record for the longest running workshop in AACC history.

Statistics don't change much from year to year. It's rare to find new analytical approaches that are practical for laboratories to implement. What does change much more often are regulations, guidelines, and skill patterns. In the 20 years since the first publication of this book, new approaches have developed.

Over the years, there has been a changing emphasis on method validation and method verification. Method validation is concerned with the question whether performance meets defined quality goals for the intended medical use. Method verification focuses on whether the performance observed in a laboratory is consistent with the performance claimed by the manufacturer.

Method verification is a less rigorous form of performance assessment. The experiments require fewer data points, but sometimes more complicated data analysis. The assumption of these studies is that the performance claimed by the manufacturer is acceptable for the intended medical use, therefore the laboratory only needs to verify the manufacturer's claims. The verification process is supposed to confirm that the method performs "as advertised."

Method validation is still required for Laboratory Developed Methods, known as LDTs. An LDT may be a new method developed for use in a single laboratory or it could be an FDA-approved method that has been modified by a laboratory. The act of modifying or creating a new test method is classified as "highly complex" and it requires all the studies detailed here in method validation, PLUS as many other studies as necessary to prove the analytical and clinical utility of the method. Nevertheless, laboratories generally received wide latitude in regulation of these tests. These new methods have been seen as crucial for innovation in laboratory medicine, and to impose more stringent regulations, such as a mandate for FDA approval for every LDT, was viewed as deeply inhibitive. Thus, LDT regulation was a bit of a blind spot in the US.

Into this blind spot came Theranos.

The whole story of Theranos is too complex to discuss here. Suffice it to say that Theranos was a fraud, and they weaponized the LDT exclusions in their fraud. By claiming LDT status, they exempted themselves from FDA approval requirements of their instrument. Ultimately, their fraud was discovered, and the CEO was banned from running a laboratory or serving as an officer in a public corporation. But in the wake of that scandal, better regulation of LDTs was given a higher priority.

That is, until 2016. Regulation of medical devices is always a political issue, and during times when the regulation of any industry is seen as an impediment, the problem of the LDT goes unresolved. And while the philosophy of any particular administration may be pro- or anti-regulation, the issue has been further complicated by a regulatory turf war between the FDA and the CDC (as represented by CMS and CLIA). Both agencies argue for the regulatory authority over LDTs. Proposals and draft legislation has been introduced. But no progress has been made, and there is no sign of a resolution on the horizon.

In the meantime, we recommend that laboratories and organizations that develop LDTs treat them as highly complex methods, and conduct all the validation studies listed in this book.

This 4th edition provides important updates based on these regulatory requirements and emerging standards of practice, particularly the latest guidelines from CLSI (Clinical and Laboratory Standards Institute).

A significant and timely addition to this manual is the discussion of qualitative testing validation and verification. For those labs seeking to evaluate PCR tests, serology tests, and antigen tests related to SARS-COV-2, we provide succinct and practical advice.

For more than fifty years, I have worked on quality control and method validation. While statistics, equations and calculations may not change, the context and the environment are constantly evolving. I hope this fourth edition helps you understand these method validation and verification numbers in the proper context of your laboratory.

James O. Westgard
Madison Wisconsin

Acknowledgments

Dr. Nils Persons, of Siemens Diagnostics, provided a critical eye in this latest revision and update of the book. As someone who worked for a diagnostic manufacturer, Nils has had ample experience interacting with customers and their verification and validation challenges. As a key figure on the CLSI committees, he also knows the latest guidelines and their latest recommendations.

Several colleagues have helped in the development of these materials, including Elsa Quam and Patricia Barry who are long-time associates of mine in the Clinical Laboratories at the University of Wisconsin Hospital and Clinics.

The antique maps that appear in this book are part of a small personal collection. I hope you find them helpful for illustrating key ideas in the book, as well as interesting and beautiful historical documents.

About the Authors and Contributors

James O. Westgard, PhD, **FACB** is an Emeritus Professor in the Department of Pathology and Laboratory Medicine at the University of Wisconsin Medical School. In addition to pioneering the use of validation protocols, he is best known for popularizing the multirule QC procedure, often called the "Westgard Rules."

Nils B. Person, PhD, FACB is a Senior Clinical Consultant for Consultant, Siemens Healthcare Diagnostics, is a board-certified clinical chemist with over 40 years' experience in laboratory medicine. He spent 15 years directing hospital laboratories prior to joining Siemens and has spent the last18 years supporting Siemens technical staff and customers. Dr. Person has also been part of a number of Clinical and Laboratory Standards Institute (CLSI) Standards development teams and was involved in the development of EP23 Laboratory Quality Control Based onRisk Management and EP26 User Evaluation of Between Reagent Lot Variation, EP21 Total Analytical Error and C24 Statistical Quality Control for Quantitative-Measurement Procedures.

Patricia Garrett,PhD is Principal at Pat Garrett Consulting. She worked in hospital labs and diagnostics companies for more than 30 years, focusing on QC, method evaluation, and infectious disease tests. She led the working group that produced CLSI's EP12-A2, User Protocol for Evaluation of Qualitative Test Performance.

Sten Westgard, MS, is the Director of Client Services and Technology for Westgard QC, Inc.

There's more online at Westgard Web

In order to squeeze in all the updates, revisions, and entirely new chapters into this book, yet still keep it a reasonable length, we had to make some cuts. Notably, we had to cut out the glossary and reference lists.

But don't worry, you can still view these resources online. Visit **http://www.westgard.com/bmv/extras.html** for access to:

- Frequently-Asked-Questions (FAQs)
- Glossary of terms
- Complete reference list for this book
- Spreadsheets, worksheets and other downloads
- Links to Method Validation calculators, including some exclusively available to the owners of this book.

Table of Contents

1: Is quality still an issue for laboratory tests?

In this introductory chapter, Dr. Westgard challenges current thinking that analytical quality is already better than needed for today's medical care. Using historical maps that were regarded as the most authoritative and accurate records of the time, he illustrates that popular "truths" can be well documented and widely believed, yet entirely wrong. He sets out the need to define requirements for quality in order to manage quality in a quantitative manner.

Objectives:

- Begin thinking about quality in a critical way.
- Recognize that current beliefs about quality may not be grounded in fact.
- Identify the critical issue for managing quality in a quantitative way.

Lesson materials:

- **Myths of quality**, by James O. Westgard, PhD

Things to do:

- Study the lesson.
- Find out what quality is needed for a cholesterol test.

The Mythical Island of California!
NOUVEAU MEXIQUE ET CALIFORNIA, *by Alain Mallet, Paris 1686. A miniature French map showing California as a flat-topped island – a myth that persisted from 1620 for over 100 years.*

Myths of Quality

A MYTH is a Mistaken Yarn, Theory, or Hypothesis!

James O. Westgard, PhD

Historical Myths of Cartography

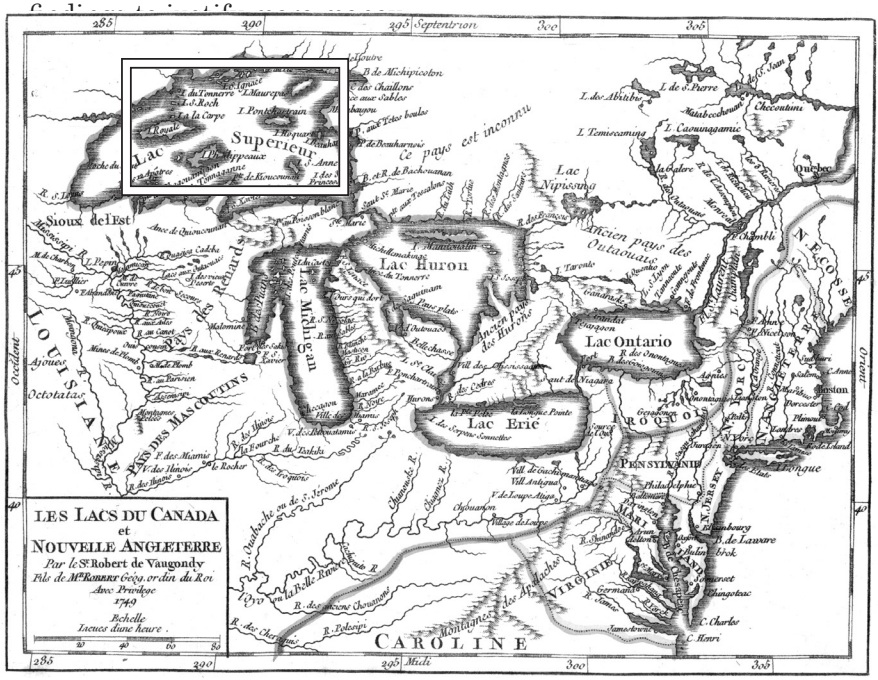
Mythical island of California. Did you know that California was an island? It's well documented on the most reputable maps of the 1600s that California was completely surrounded by water. For example, see the accompanying map that shows the *Isle de Californie*. There it is, documented in black and white, proof that California was an island.

This map of *Nouveau Mexique et Californie* by Alain Mallet was published in 1686 in the *Description de l'Univers* (Paris). Mallet copied the flat-topped model of California that appeared in an earlier map by Sanson, who was one of the most distinguished French cartographers (It was very common for mapmakers to copy each other's work). When a new discovery appeared on one map, it was widely disseminated on the other maps of the time. The discovery that California was an island was first documented in 1622 and persisted on maps as late as 1750, even though evidence in 1705 clearly established that this was not true.

Mythical island of Friesland. Actually, there is quite a history of mythical islands, suggesting that these myths are not as rare as you might expect. In the late 1500s, one of the most famous mapmakers, Abraham Ortelius, prepared a map of the Northern Atlantic that showed an island of Friesland lying a bit west and south of Iceland, complete with a detailed description of the coastline, the harbors, the people who lived there, what they looked like, and what they did for a living. It's a beautiful map, decorated with sailing ships and sea creatures, and was the most authoritative map of the area at that time. The only problem was that Friesland didn't exist. When people sailed to the new world and passed Iceland, they ascribed more and more details and reality to Friesland because they expected it was the next body of land.

The mythical island of Friesland!
***SEPTENTRIONALES REGIONES**, by Philip Galle, Antwerp, 1595. A miniature of Ortelius' famous map of the north Atlantic region showing the mythical island of Friesland (see box) located to the southwest of Iceland. [Emphasis Added]*

Mythical islands in Lake Superior. Another example that is of interest to those of us in the midwest are the islands of Ponchartrain and Phillipeaux in Lake Superior. When the border between the US and Canada was settled by the Treaty of Paris in 1783, it was decided that these islands would be part of the US In the early 1800s when Wisconsin was being settled, the US government sent out surveyors to map this area more completely, but they couldn't find these islands! They appeared on all the maps of the time, but they didn't show up above the water. It seems that the explorers created these islands and named them for the government minister who was funding their investigations. They were probably trying to get more funding for further explorations and needed some preliminary



The mythical islands in Lake Superior!
LES LACS DU CANADA et NOUVELLE ANGLETERRE, *by Robert de Vaugondy, Paris, 1749. This map shows Lac Superieur containing the real Isle Royale and the mythical islands of Phillipeaux, Pontchartrain, Maurepas, St. Anne. [Emphasis added]*

Modern myths of quality

These cartographic myths are amusing in retrospect, but they were taken very seriously at the time and created problems later on. There are myths today that are also taken seriously and will cause us significant difficulty in the future. Some of them hit close to home – the quality of healthcare and the quality of laboratory testing.

Myth: QA assures quality in healthcare. It's a *mistaken yarn* that puts a good spin on current efforts to measure the quality of healthcare. As healthcare providers, we all talk about quality assurance (QA), but our quality assurance programs (which are often required by regulation and accreditation) primarily deal with **measuring** performance. Quality Assessment would be a better name for these efforts. While it is important to **assess** quality to know how well we're doing, measuring quality doesn't **assure** that the necessary quality will be achieved. Achieving quality actually requires quality design and planning, which starts with defining the quality that is needed, then builds that quality into the process.

Myth: Statistical QC controls the quality of laboratory tests. It's a *mistaken theory* that the mere use of statistics assures that laboratory test results have the necessary quality. Virtually all laboratories apply statistical quality control (QC) as part of their efforts to assure the quality of laboratory tests. While we may not understand the theory or the statistics, we still seem to believe that something magical happens because of those statistics. We act as if analyzing controls and plotting results on control charts will assure the quality of our testing processes, even if we don't understand any of the numbers.

Myth: Quality can be managed even if the required quality isn't known. It's a *mistaken hypothesis* that quality can be managed even if **we don't know** the quality that is needed. Few laboratories have defined the analytical quality that is needed for the tests they perform. How is it possible to know we are achieving an unknown? Can you manage the finances of the laboratory without knowing the budget? Don't you need to know the quality required for a laboratory test if you are to manage the quality of the testing process?

Myth: Quality requirements need to consider only imprecision and inaccuracy. This problem with quality requirements gets to be even more complicated. Current thinking about quality goals and requirements is flawed because it considers only the stable method performance characteristics (imprecision and inaccuracy). If performance is always stable, why bother doing quality control at all? If QC is necessary, don't we have to consider the performance characteristics of QC procedures (probabilities for error detection and false rejection) in our goal-setting models?

Myth: Current laboratory methods have better imprecision and inaccuracy than needed. The net effect of all these myths is the belief that the performance of current laboratory methods is better than required for medical needs. This belief is based on a mistaken theory for setting quality goals, a mistaken hypothesis in equating all medically tolerable variation with analytical variation, disregarding the subject's own biological variation, and the mistaken assumption that QC procedures have ideal response curves and can detect any change in performance, regardless how small.

Myth: Analytical quality is a given today. As a consequence of these myths, there is a common feeling today that analytical quality is a given, i.e., analytical quality itself is being assumed today. In the midst of programs on Six Sigma, Lean, Risk Management, and Total Quality Management (TQM), it is often mistakenly assumed that the problems in technical quality management have already been solved. This represents the mistaken hypothesis that past efforts have resolved any technical difficulties, so now we can get on to new issues that are in vogue, such as monitoring customer satisfaction, measuring patient outcomes, etc.

Myth: No further improvements in analytical quality are needed. The collective result of all these myths is a false sense of security regarding the quality of laboratory testing processes. Many think analytical quality is so good that there is no need for further improvement. This is the most serious myth of all because it makes us complacent about what we are doing and hinders efforts to further improve the analytical quality of laboratory tests.

Myth: The government regulates laboratory tests to make sure quality is acceptable. While it is true that the Food and Drug Administration (FDA) does approve new tests and analytic systems, it is important to understand that this clearance is based only on "truth in labeling." Manufacturers are required to make claims for certain performance characteristics, such as reportable range, precision, accuracy, interference, detection limit, and reference range and to submit data to support those claims. The FDA's process focuses on whether or not the data supports the manufacturers' claims, not whether or not the quality of the testing process is acceptable for patient care. We may believe that manufacturers would not submit a new test for FDA clearance unless the quality was acceptable, but that assumption is not always true.

Myth: Laboratories today should focus and pre-analytic and post-analytic errors since analytic errors are no longer a problem. This idea surfaced in 1990 from the CDC in an effort to broaden the quality assessment efforts in clinical laboratories. Later in the 90s, CMS adopted that perspective in revising the CLIA Final Rules to include quality management of the "total testing process," i.e., pre-analytic, analytic, and post-analytic parts of the testing methodology. By the second decade of the 21st century, this belief became "common wisdom" in the laboratory community. Today that belief today is used (and accepted) as an argument for reducing the amount of statistical QC performed during the analytic phase. It satisfies our yearning to do less QC, to simplify laboratory testing, and to reduce costs and eliminate trouble-shooting and repeat analyses, all of which allow laboratory tests to be performed in testing sites where technical skills and laboratory experience may be lacking.

Where's the evidence?

Theranos is the latest example of how everything that can go wrong will go wrong in medical diagnostics and analytical testing.

In what is now known to be a fraud, the corporate leaders of Theranos claimed they had a diagnostic instrument that could run more tests, faster, cheaper, and easier than traditional core laboratory instruments. In reality, they had instruments that barely functioned, and they modified existing analyzers to perform tests

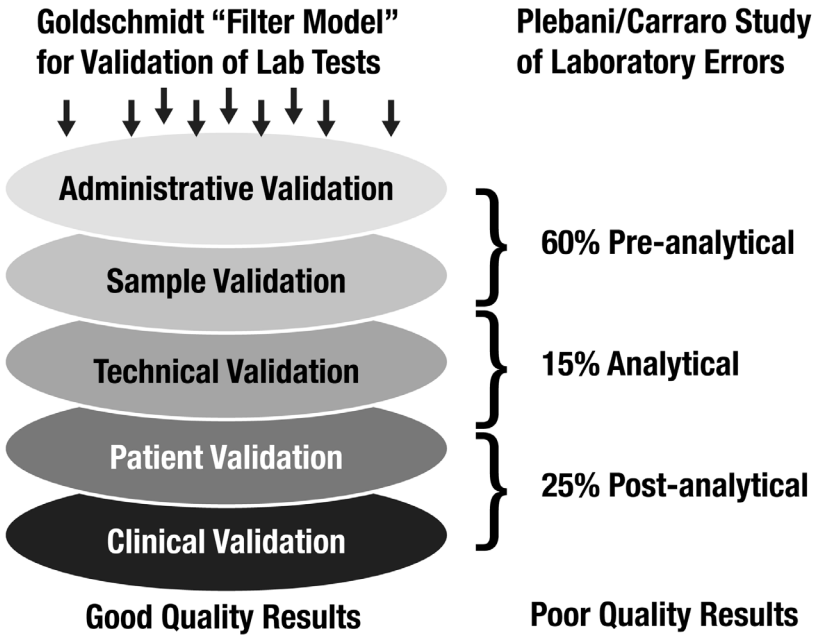
on specimens that were highly diluted. The gulf between the claims of performance and the actual performance were so wide, the company became a subject of a series of *Wall Street Journal* articles, a best-selling book, *Bad Blood*, a podcast, multiple documentaries, and a rumored feature film (still to come). Theranos went from being a Silicon Valley darling, a "unicorn" valued at $9 Billion(!) dollars, to going broke. The CEO, Elizabeth Holmes, went from having a personal net worth of $4.5 billion dollars, to being so broke she couldn't pay her own attorneys in one of the many lawsuits.

The fact that Theranos was able to fool investors, regulators, and patients **for 10 years** further proves the lack of robust standards and practices for quality management in medical laboratories. As Theranos's quality problems became evident, they belatedly tried to create a veneer of quality by adding a scientific board of advisors, which included past officers of the American Association for Clinical Chemistry, AACC. But adding a set of polished resumes didn't improve the central engineering failures. It was too late for that.

In this age of Evidence-Based laboratory medicine, where's the data to support these beliefs that we no longer need to worry about analytical quality? Consider one last myth – laboratories should focus and pre-analytic and post-analytic phases of the total testing process – because it rests on the other beliefs.

First, let's examine a more complete model for the total testing process, as provided by Goldschmidt et al[2] and shown in the accompanying figure. Called a "filter model," the figure illustrates a series of filters through which a laboratory test request, specimen, and sample must pass. In reality, these are mathematical filters, rather than physical filters as suggested in the figure, but as laboratory scientists rather than mathematicians, the approach is easy to understand from the illustration. This more detailed model describes 5 phases or filters for validating the total testing process.

- **Administrative validation** refers to steps beginning with the selection and ordering of the right test, collection of the right information to understand the context of the test, as well as validation of right patient conditions, right preparation, etc.



- **Sample validation** is concerned with obtaining the right specimen at the right time on the right patient, the right processing and transportation of the sample, and the right use of that sample for analytical measurements;
- **Technical validation** has to do with getting the right answer, which requires knowing the quality required for a test, validating the precision and accuracy of measurement process, designing the right QC procedure, and implementing the measurement and control procedures properly;
- **Patient validation** requires that right test result be correctly reported to the right patient record and considers whether that test result is consistent with knowledge about the patient, other test results on that patient, within the expected variation of the individual patient and the appropriate population group, as well as relationship to critical or alert values, and consistency with patient populations;

- **Clinical validation** is concerned with the patient receiving the right clinical treatment based on the laboratory test results and services. Clinical validation goes beyond what is normally considered to the validation of test results in the US.

This is a European model and demonstrates that patient and clinical validation have long been a critical part of the validation of laboratory tests. Patient and clinical validation are important professional responsibilities of MD and PhD level laboratory physicians and scientists. With increasing workload, they have developed computerized tools and programs to standardize and facilitate this "medical review" or "medical QC." The importance of the first step (administrative validation) becomes clear in the context of the information needed to complete this medical review and control.

Next, let's consider the most definitive study on the sources of laboratory errors [3]. Drs. Plebani and colleagues have studied laboratory errors for decades and are recognized as leaders in performing such studies. Their results document the distribution of errors shown in figure. Clearly there are more pre-analytic errors (60%) than post-analytic errors (25%) than analytic errors (15%). Many clinical laboratory scientists cite these figures to support the idea that pre-analytic and post-analytic errors are more important than analytic errors and often conclude that analytic quality is no longer an issue.

However, a more detailed reading of the study shows that from the total of 51,746 tests, there were 393 questionable results, 160 of which were confirmed as laboratory errors. Of these 160 errors, 46 caused inappropriate patient care, and 24 of those cases of inappropriate patient care were caused by analytical errors. That means that **over half the cases (52%) of inappropriate patient care are due to *analytical* errors**. Analytical errors are still the largest and most important source of errors that cause harm to patients!

We need to recognize that the core competency of a laboratory must be producing the correct test result. All sources of error in the testing process must be carefully managed and monitored, but it starts with analytical errors. If we can't, don't, or won't assure analytical quality of our test results, then we should not be in the business of providing laboratory testing services.

What's the point?

You need to think critically about quality and recognize that many of our current beliefs are not based on scientific evidence. These myths need to be exposed if the technical management of analytical testing processes is to be improved.

That's the purpose of this introduction! You need to assess many of the quality management practices that are accepted in laboratories. To begin, you need to understand how quality requirements can be defined, how method performance should be measured experimentally, how the experimental data can be analyzed with statistics to estimate analytical performance characteristics, and how a decision on the acceptability can be made.

Once the performance of a method has been judged to be acceptable (Basic Method Validation), you need to select a statistical QC procedure that can detect medically important errors (Basic Quality Management Systems), make routine measurements on the necessary number of controls, and interpret the control results using the appropriate decision criteria or control rules (Basic QC Practices).

References

1. Carreyou, Bad Blood: Secrets and Lies in a Silicon Valley Startup. Knopf (New York, NY) 2018.

1. Oosterhuis WP, Ulenkate HJLM, Goldschmidt HMJ. Evaluation of LabRespond, a new automated validation system for clinical laboratory test results. Clin Chem 2001;46:1811-1817.

2. Carraro P, Plebani M. Errors in a Stat Laboratory: Types and frequencies 10 years later. Clin Chem 2007;53:1338 - 1342.

Self-Assessment Questions:

- What myths of quality exist in laboratories today?
- What can be done to improve laboratory quality management practices?

3: What is the purpose of a method validation or verification study?

Dr. Westgard reveals the inner, hidden, deeper, secret meaning of method validation and verification. Knowledge of this "meaning" should make the method validation/verification process more rational and understandable.

Objectives:

- Understand what method validation and verification studies are supposed to study.
- Recognize the potential shortcoming of statistics in a method validation or verification study.
- Identify the different types of analytical errors that need to be assessed.

Lesson materials:

- **MVV – The inner, hidden, deeper, secret meaning,** by James O. Westgard, PhD

Things to do:

- Study the lesson.
- Review a method validation or verification report from the scientific literature.

Method Validation and Verification:
The Inner, Hidden, Deeper, Secret Meaning

James O. Westgard, PhD

When I was a freshman in college (this was quite a while ago, before computers but after quills), I had an English professor who taught me something I've never forgotten. He always asked, "What's the inner, hidden, deeper, secret meaning in what you're writing?" In other words, what are you really trying to accomplish? You'd better figure it out if you expect someone else to understand it. Sure, you can write down a string of words, but you've really got to be clear on what you want to accomplish, otherwise the true meaning will remain a secret.

The real surprise came on my first job as a clinical chemist when I began evaluating the performance of a new multichannel chemistry analyzer. I studied all the existing scientific literature that provided guidelines for performing method validation studies, but it wasn't at all clear how to tell whether or not a new method was acceptable. No one was telling the secret! And that secret is of paramount importance to evaluate a method properly. Sure, you can collect some data, calculate some statistics, and provide some paper in a folder (or files in an electronic folder) to show a lab inspector, but is that really why you're doing this?

While I won't claim my English professor made me a better writer (nor can you blame him), he did make me a better scientist by helping me search for the deeper meaning and real purpose in what I do. What's the real purpose of method validation? What's the problem we're trying to solve? Does the present practice provide a correct solution? Is there a better way to do this? How do you know what's the right way to validate the performance of a method?

The Secret Revealed

Here's the inner, hidden, deeper, secret meaning of method validation – and you don't have to read any further to get the message – ERROR ASSESSMENT. You want to estimate how much error might be present in a test result produced by a method in your laboratory. With

this information, you then want to be sure that amount of error won't affect the interpretation of the test result and compromise patient care. If your observed errors are so large they can cause an incorrect interpretation, the method isn't acceptable. To be acceptable, the observed errors need to be small relative to changes that will cause a change in the interpretation of a test result.

A focus on analytical errors is the key to the whole method validation process. What kinds of analytical errors might occur with a laboratory method? What experiments can provide data about those errors? What's the best way to perform those experiments to assess the errors? How much data needs to be collected to obtain good estimates of errors? What statistics best estimate the size of those errors from the experimental data? What size errors are allowable without affecting the interpretation of a test and compromising patient care?

Method Validation is about error assessment – that's the secret!

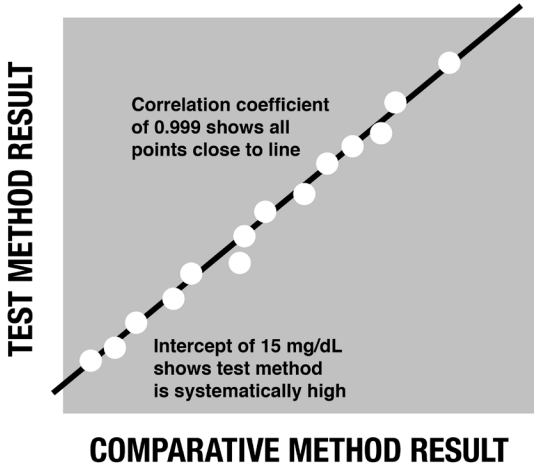
[By the way, here's the inner, hidden, deeper secret meaning of Method Verification: Did you get what you paid for? Does it live up to the label? Does performance match the claim?]

A Quick Proof

The correlation coefficient is a statistic that is almost always calculated and reported to describe the results from a comparison of methods study. A value of 1.000 indicates perfect correlation between the results of two methods. Other statistics (such as slope, intercept, and standard deviation of the residuals) can also be calculated from the same data to estimate the size of errors occurring between the methods. Which are more useful?

Consider the following situation. Here's a new cholesterol method where the results from a comparison of methods study give a correlation coefficient of 0.999, which is very close to ideal value of 1.000. Sounds pretty good, doesn't it? How close are the results between the two methods? Is the new method acceptable? Let me give you some additional information.

Here's the plot of the test results by the new method vs those from the comparative method. Note first that the correlation coefficient shows that the results are *close to the best line of fit* between the methods; it does not show that the test values are the *same* as the comparative values.



Results of a comparison study, where the new or "test" method values are plotted on y-axis and comparison values on x-axis.

The plot shows that almost all the new method values are systematically higher by 15 mg/dL. Does this information that there is a systematic error of 15 mg/dL help with your decision on the acceptability of the new method? It doesn't look so good anymore, does it? Being in error by 15 mg/dL may limit the usefulness of the test results produced by the new method.

As laboratory professionals, we intuitively understand errors and have a sense of how they might affect the interpretation of test results and the related care of patients. We don't have the same sense about statistics! That's why statistics should be used to estimate the errors that are meaningful to us[1] – that's a second important secret and we'll deal with it in detail later.

From this simple example, you can recognize the difficulty in interpreting a correlation coefficient, since it doesn't provide a useful estimate of analytical errors. Information about the *size* of analytical errors is more useful for judging the performance of a method[2]. The fact that the correlation coefficient is commonly calculated doesn't make it useful. It just shows that people don't know the secret of method validation!

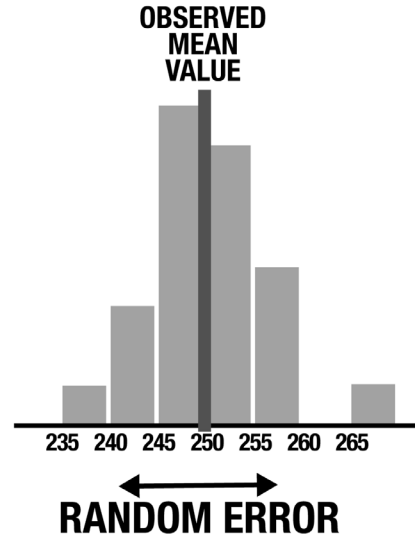
Analytical Errors

Let's focus on analytical errors and make sure we have a common understanding of the different kinds of errors that need to be estimated. Here's a list of terms you need to understand: random error (imprecision), systematic error (inaccuracy), constant error, proportional error, and total error.

Random error, RE, or imprecision

Random error is described as an error that can be either positive or negative, whose direction and exact magnitude cannot be predicted, as shown in the accompanying figure by the distribution of results for replicate measurements made on a single specimen.

Imprecision is usually quantitated by calculating the standard deviation (SD) from the results of a set of replicate measurements.

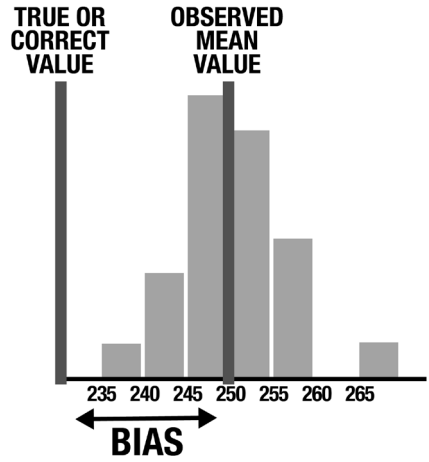


Random Error (RE) or Imprecision, *as shown by the distribution of test values.*

The SD increases as the concentration increases, therefore it is often useful to calculate the coefficient of variation (CV) to express the SD as a percentage of the mean concentration from the replication study. The maximum size of a random error is commonly expressed as a 2 SD or 3 SD estimate to help understand the potential size of the error that *might* occur.

Systematic error, SE, or inaccuracy

A systematic error is an error that is always in one direction, as shown in the accompanying figure where a systematic shift displaces the mean of the distribution from its original value. In contrast to random errors that may be either negative or positive and whose direction cannot be predicted, systematic errors are in one direction and cause all the test results to be either high or low.

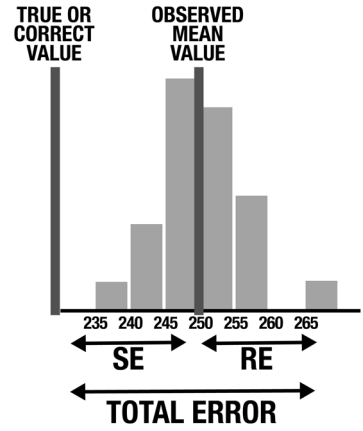


Systematic Error (SE) or Inaccuracy, *as shown by shift or bias between mean value and correct value.*

How high or how low can be described by the bias, which is calculated as the average difference, or the difference between averages of the values by the "test" method and a "comparative" method in a comparison of methods experiment. Alternatively, the expected systematic difference may be predicted from the equation of the line that best fits the graphical display of test method values on the y-axis vs comparative method values on the x-axis. SE may stay the same over a range of concentrations, in which case it can also be called **constant error**, or it may change as concentration changes, in which case it can be called **proportional error**.

Total Error, TE

Total error is the net or combined effect of random and systematic errors, as shown in the accompanying figure. It represents a "worst-case" situation, or just how far wrong a test result might be due to both random and systematic errors. Because laboratories typically only make a single measurement for each test, that measurement can be in error by the expected SE, or bias, plus 2 or 3 SD, depending on how you estimate the effect of RE.



***Total Error (TE)**, includes both systematic error (SE) and random error (RE).*

Why do we combine the errors? While we in the laboratory like to think about imprecision and inaccuracy as separate errors, the physician and the patient experience the total effect of the two, or the total error. Total error provides a customer or consumer-oriented measure of test performance, which makes it the most important parameter for judging the acceptability of analytical errors.

[Note: "total error" here means total *analytical* error, it is not the total of analytical, pre-analytical, and post-analytical error. Some seek this total total error, but it remains nearly impossible to formulate this sum.]

Trends and Directions

Efforts to provide worldwide standards of laboratory practice are changing the terms and concepts we use. ISO, the International Standards Organization, provides specific guidelines for healthcare laboratories in its document **15189 – Medical Laboratories – Particular requirements for quality and competence** [3]. To give you a flavor of the ISO approach, let's see what it recommends:

> "When describing the performance of procedures and the reliability of their results, ISO terminology should be used. Results should be universally comparable and this requires metrological traceability, the concomitant uncertainty indicating reliability should be obtained in a universal and transparent fashion, and should be combinable." [4]

The preferred ISO concepts and terminology are "trueness" and "uncertainty." ***Trueness*** is used to describe the "closeness of agreement between the mean obtained from a large serious of measurements and a true value." This is equivalent to the terms bias and systematic error in this chapter. ***Uncertainty of measurement*** is used by ISO to describe a "parameter, associated with the result of a measurement that characterizes the dispersion of the values that could reasonably be attributed to the measurand," where measurand refers to the particular analyte or test. Uncertainty describes a range of values that correspond to a given test result, e.g., a test result of 200 may have a "standard uncertainty" (SD, CV) of 4 units or 2%, indicating that a value of 200 represents an "expanded uncertainty" of 192 to 208 units (95% or 2SD confidence interval). This concept sounds and looks similar to precision, but the estimate of uncertainty may also incorporate components in addition to the random error of the method, e.g., the uncertainty associated with calibrators, uncertainty in the estimate of any bias, etc. There are ongoing discussions about the relevance of the concepts of Total Analytical Error vs Measurement Uncertainty, but for now the error concepts provide more practical tools for applications in the medical laboratory [5].

These concepts of trueness and uncertainty come from the world of metrology, where customers are provided with products having assigned target values along with the uncertainty that expresses the correctness or "doubt" in the target value. The ISO approach expects customers to know the meaning of uncertainty. **In general, they don't.** The world of laboratory medicine is different. Physician customers and patient consumers are not aware of the science of measurements and the uncertainty in test results. It would be better if laboratories managed their analytical methods to verify the attainment of the intended clinical quality of results, but in the absence of doing so, it will become necessary to inform the

customer of the actual "doubt" of the reported results. Measurement uncertainty may be part of your future.

What's the point?

You must understand the "why" of method validation in order to understand "how" method validation should be accomplished! The "why" defines the purpose, which is to determine the amount of error that might occur with a method. The "how" defines the experimental protocols and data-analysis procedures that provide estimates of the errors. Method validation is all about errors!

Laboratory regulations in the US require that method performance for any new method be "verified" prior to reporting patient test results. **Under the CLIA Final Rule, laboratories must verify the reportable range, precision, accuracy, and reference intervals for all *non-waived* methods implemented.** For methods that are developed in-house or modified by the laboratory, the additional characteristics of analytical sensitivity (detection limit) and analytical specificity (interference, recovery) must also be verified. More extensive reference range studies are also appropriate.

The responsibility for method verification or validation resides with the laboratory. While manufacturers will often run studies and collect data during the installation of new analytical systems, the laboratory is still accountable to see that adequate data have been collected and that these data show that the new methods provide acceptable performance in the laboratory.

References

1. Westgard JO, Hunt MR. Use and interpretation of common statistical tests in method comparison studies. Clin Chem 1973;19:49-57.

2. Westgard JO, Carey RN, Wold S. Criteria for judging precision and accuracy in method development and evaluation. Clin Chem 1974;20:825-33.

3. ISO/FDIS 15189 Medical laboratories – Particular requirements for quality and competence. 2012. International Organization for Standards, Geneva Switzerland.

4. Dybkaer R. Setting quality specifications for the future with newer approaches to defining uncertainty in laboratory medicine. Scand J Clin Lab Invest 1999;37:579-584.

5. Westgard JO. Error methods more practical, but uncertainty methods may still be preferred. Clin Chem 2018;64:636-638.

Self-Assessment Questions:

- What are the two major types of analytical errors?
- What is meant by "total error"?
- How is total error related to the basic types of errors?
- How does your literature report describe the errors of the method?
- What statistics are used in the literature report?
- How do the report's conclusions relate to the errors of the method?

7: How are the experimental data analyzed?

Before explaining the details of specific experiments, Dr. Westgard gives an overview of the data analyses that are useful and appropriate for the different studies. The approach here is to consider data analysis "tools," rather than statistics and equations. These tools are readily available in the form of calculators, electronic spreadsheets, and computer programs. Online calculators are introduced to provide easy-to-use tools for use with this book.

Objectives:

- Minimize your fear of statistics.
- Identify the tools and techniques needed for data analysis.
- Match the tools with the experiments and errors to be estimated.
- Recognize the capability of available calculation tools.

Lesson materials:

- **MVV – The data analysis tool kit,** by James O. Westgard, PhD
- **The method validation data analysis tool kit,** http://www.westgard.com/mvtools.html

Things to do:

- Study the lesson.
- Practice using the online calculators with the sample data.
- Review the statistics presented in a published validation report.

Method Validation and Verification:
The Data Analysis Tool Kit

James O. Westgard, PhD

This chapter is actually about statistics, but I didn't put "statistics" in the title because too many people get turned off as soon as they see that word. Others become uncomfortable when they see the equations for the calculations. By now – three sentences into this chapter – you may be wondering if you can just skip the chapter and avoid the topic. The answer is NO. You need statistics to make sense of the data collected in method validation experiments.

Tools, not equations!

To reduce the mental roadblocks in understanding statistics, **there aren't any equations in this chapter!** Instead, we're going to assume the calculations can be easily performed with the informatics available today. Your main job will be to recognize *what* calculations are useful for different sets of data.

When I lecture on this topic, I begin by showing the class a bunch of tools, such as a hammer, wrench, saw, and screwdriver. Office tools (such as a stapler, scissors, paper, and pen) would also provide good examples, but you're too comfortable with those tools. I want you to learn that you can use tools, even if you're not so comfortable with them. So, let's consider the hammer, wrench, saw, and screwdriver.

- Which tool would be most useful for hanging a picture on the wall?
- Which tool would you use to tighten the bows on your sunglasses?
- Which tool do you want to take along at Christmas time when you go into the forest to get your tree?
- Which tool do you hope to have along if your car has a flat tire?

You don't have to be an engineer, mechanic, or carpenter to recognize which tool fits these jobs. Everyone makes use of these tools to do certain basic jobs. While there are more complicated applications that take more skill and knowledge – and sometimes more specialized tools – everyone is capable to making practical use of the common tools.

Another key takeaway is to recognize that all tools have limitations. You wouldn't try to inflate a tire with a hammer because it's the wrong tool for the job. Knowing the limitations of a tool is also key to effective use.

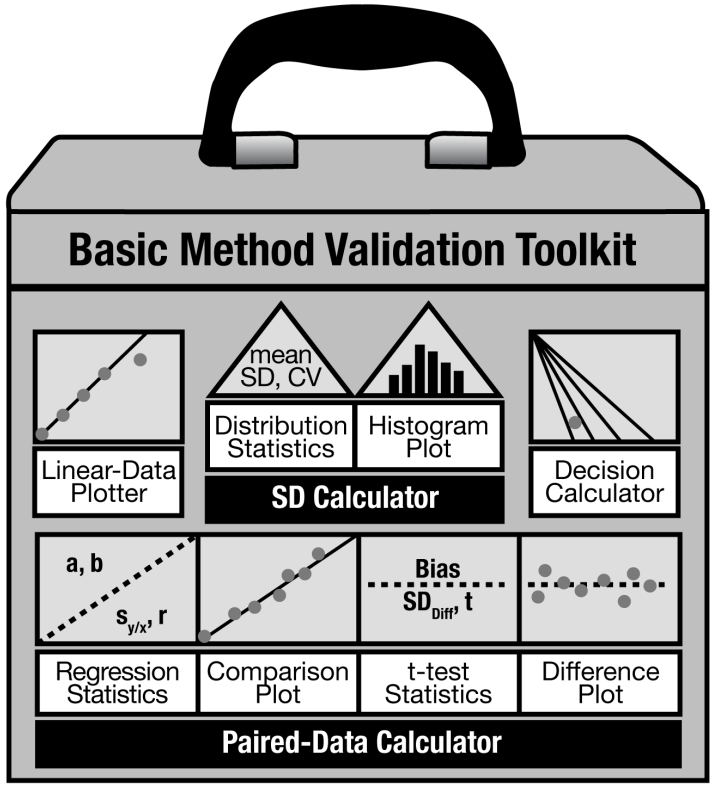
It's the same with statistics!

Recommended tools for data analysis

Statistics are just tools for evaluating experimental results, i.e., data, and summarizing all that data in just a few numbers that provide information about the data. Remember that the objective of each experiment is to estimate the amount of error from the data collected. The key with the statistics is to know which ones will provide useful information about the errors of interest in the different experiments.

First, we want to know the imprecision or random error from the 20 or more data points collected in a replication experiment. Then we need to define the usable analytical range (or reportable range) of the method so that the experiments can be properly planned and valid data can be collected. The reportable range is usually defined as the range where the analytical response of the method is linear with respect to the concentration of the analyte being measured. Finally, we need to make a judgment on the performance of the method on the basis of the errors that have been observed. The statistics are used to make reliable estimates of the errors from the data that have been collected.

Here's a picture of the tool kit you need to analyze the data from basic method validation experiments. The tool kit includes several calculators and plotters:



- **Linear-data plotter** to display the observed method response versus the relative or assigned concentrations for a series of solutions or specimens;
- **SD calculator** to determine distribution statistics (mean, SD, CV) and to display a histogram of the distribution;
- **Paired-data calculator** to determine regression statistics (slope or a, y-intercept or b, standard deviation about the regression line or $s_{y/x}$, and correlation coefficient, r), display the data in

17: How do you validate or verify a qualitative method?

For tests that aren't entirely quantitative, validation and verification is not as well described. There are fewer guidelines and little in the way of published examples. This chapter reviews how to approach validation and verification for a qualitative or semi-quantitative method. In the age of COVID19, these techniques are especially relevant.

Objectives:

- Express precision as the uncertainty at the cutoff.
- Express accuracy as clinical agreement with a gold standard or comparative method
- Calculate Percent Positive Agreement (PPA) and Percent Negative Agreement (PNA)
- Understand the impact of Prevalence with Positive Predictive Value (PPV) and Negative Predictive Value (PNV)

Lesson materials:

- **MVV – Evaluation of Qualitative Tests**
  by James O. Westgard, PhD
- Westgard 2x2 Contingency Calculator
- Westgard 2 Test Comparator Tool

Things to do:

- Study the materials.

Method Validation and Verification:
Evaluation of Qualitative Tests

James O. Westgard, PhD, Patricia E Garrett, PhD, Sten Westgard MS

As we were preparing this 4th edition, the COVID-19 pandemic hit and the need for validation of qualitative tests became a high priority. Under the conditions of Emergency Use Authorization (EUA), the FDA registered many new tests without requiring the extensive documentation of performance required for the 510k approval process. Laboratories approved by CLIA for performing moderate and high complexity tests were eligible to implement manufacturers' EUA tests, but were still required to perform some minimal validation studies, plus analyze positive and negative QC samples with each analytical run of patient samples. The FDA also required that laboratories confirm the first 5 positive and first 5 negative patient results by comparison with a previously approved EUA method. CLIA was slow to issue any specific guidance for laboratories, thus the main sources of guidance were from the FDA [1] and from the CLSI EP12-A2 [2] document that provides general guidance for evaluating the performance of qualitative tests.

In this book, our applications and discussions have focused on quantitative methods that have a continuous measurement scale, i.e., a cholesterol test can have a result that is any number from 0 to say 400. The key test performance characteristics are precision and accuracy. Precision is related to random error and typically expressed as a standard deviation (SD). Accuracy is related to systematic error and is described by trueness (bias) and Total Analytical Error (both random and systematic errors).

Qualitative tests provide binary results, yes/no answers, or positive/negative results. A pregnancy test is a good example of a binary output. The result is the patient is pregnant, or not, with no possibility of being just a little bit pregnant. Blood bank screening tests also provide positive/negative results, but often use a cutoff to convert an internal continuous response to a binary result. The existence of an internal continuous response means there are some possibilities to apply some of the experiments used for quantitative test methods, thus for Enzyme Linked ImmunoSorbent Assays (ELI-

Index

A

B

C

D

E

Q

R

S

T

U

V

W