Benjamin Franklin, early American inventor, entrepreneur, author, politician, and diplomat, was known for his pithy expressions. One of his most famous, oft-quoted sayings is, “An ounce [28 g] of prevention is worth a pound [454 g] of cure.” Less well known is that Franklin’s sage advice about prevention was given in the context of his organizing the first fire-fighting company in Philadelphia (1). At that time, fires posed an ever-present risk for the residents of Philadelphia and other cities. Meticulous handling of candles, lanterns, embers, and ashes helped avoid property loss, financial ruin, injury, and death. The safety of the populace was inextricably tied to the fire-handling practices of every individual.

Today, clinical laboratories are essential components of healthcare systems, contributing invaluable information for the diagnosis, treatment, and avoidance of disease. Decades of research and development have yielded a wide array of sophisticated assays and analyzers. Simultaneously, societal trends and improvements in disease management have fueled an increase in the prevalence of chronic disease, which is driving a clinical demand for long-term consistency in laboratory test results. Assay suppliers attempt to minimize lot-to-lot variation in reagents via manufacturing processes and by conducting lot-release testing to avoid distributing unsuitable lots. Clinical laboratories perform lot-to-lot validation testing to verify a manufacturer’s performance claims and assure the ongoing reliability of testing. Yet, practicing laboratorians know that far too much time is spent dealing with inconsistent results.

In this issue of Clinical Chemistry, Algeciras-Schimnich and colleagues report multiple failures of lot-to-lot validation procedures to detect significant between-lot differences in an insulin-like growth factor 1 (IGF-1) assay over a 5-year period in 2 laboratories, the Mayo Clinic and the University of Virginia (2). During this period, the Mayo Clinic used 32 reagent lots, and the University of Virginia used 16 lots. With every lot change, each laboratory conducted lot-to-lot validation studies according to its standard operating procedures without identifying any reagent lot as significantly different from its predecessor. QC and proficiency testing also failed to identify lot-to-lot inconsistency. Clinicians in both institutions, however, began contacting the laboratories with increasing frequency about IGF-1 results that appeared spuriously high, yet the laboratories were unable to confirm the clinicians’ suspicions in lot-to-lot validation studies, QC, or proficiency-testing results. Eventually, the results of an expanded investigation that included a retrospective, longitudinal analysis of lot-to-lot validation results and lot-specific patients’ means, medians, and proportions of results exceeding reference interval limits clearly showed significant lot-to-lot inconsistency.

Algeciras-Schimnich and colleagues are to be commended for their openness and transparency. Their relevant, compelling article suggests several important points: (a) Manufacturers’ lot-release procedures are insufficiently stringent to maintain lot-to-lot consistency, and even highly reputable laboratories’ lot-to-lot validation procedures are inadequately powered to detect clinically significant differences. (b) Results of lot-to-lot validation studies need to be evaluated for cumulative effects; otherwise, significant undetected drift may occur one insignificant step at a time. (c) QC and proficiency testing are unreliable for detecting lot-to-lot variations. (d) Manufacturers’ lot-release procedures and clinical laboratories’ lot-to-lot validation procedures favor the acceptance of new lots, which forces clinicians into the role of quality safety net rather than confident consumer. Finally, (e) failure of validation studies to detect lot-to-lot differences often has important clinical and financial consequences.

The experience of Algeciras-Schimnich and colleagues is not unique. For example, van Rossum and colleagues reported a conversion to a new cobalamin assay with strikingly similar outcomes (3). A significant downward shift in patients’ results failed to be detected by the manufacturer’s assay-release studies, the laboratory’s validation studies, ongoing QC, or proficiency testing. Clinicians, however, perceived an increased frequency of results below the reference in-
terval. A retrospective analysis revealed a 6-fold increase in low results, confirming the clinicians' suspicions.

Published reports of undetected lot-to-lot inconsistencies undoubtedly represent just the tip of the iceberg. Consequently, we can assume that lot-to-lot inconsistencies frequently either go undetected or incur delays in detection and resolution, producing human and financial costs of unknown but substantial scope. The question is why. Are we laboratorians simply tolerant of inadequate lot-release and lot-to-lot validation procedures? Are we unwilling to bear the cost of or have we inadequate expertise in conducting more-rigorous studies? Are we inadequately connected with other laboratories to leverage our collective resources, data, and analytical skills? Do we have inadequate knowledge or capacity to monitor assays prospectively? Do we not recognize that data sharing may facilitate earlier detection of assay issues and help drive improvements and innovation, or do we simply lack mechanisms for sharing validation-failure data or ongoing monitoring data? Are we content to rely on clinicians to tell us when there is an assay problem?

The last question is particularly troublesome. Although clinicians should never be discouraged from giving feedback, undue reliance on clinician reporting in laboratory quality management is what the US Federal Aviation Administration calls “regulating by counting tombstones” (4). Fortunately, laboratorians can do much to improve the current situation. I see the article by Algeciras-Schimnich and colleagues (2) as a multifaceted call to action.

Institute Rigorous Validation Testing by Pooling Resources

Viewing assay validation as “a skilled, time consuming, and costly undertaking,” the English National Blood Service (NBS) formed a national kit-evaluation group for blood donor infectious-disease assays (5). After verifying that manufacturers’ claims meet the specific criteria of the kit-evaluation group, the NBS reference laboratory performs extensive initial validation testing. Unsuitable lots are rejected at this point. Lots successfully passing this initial testing are distributed to testing laboratories for “delivery acceptance” testing. During a 9-year period, initial validation testing at the reference laboratory rejected 7 lots, and 5 failures occurred in delivery-acceptance testing (6). As a national entity, the NBS has the critical mass to centralize some aspects of validation testing and to aggregate data from distributed testing, thereby improving the cost-effectiveness and robustness of validation studies. A similar opportunity exists for integrated healthcare systems or collaborative arrangements between independent laboratories.

Include Longitudinal Assessment of Lot-to-Lot Validation Results

Although each lot-to-lot validation study by Algeciras-Schimnich and colleagues failed to show significant differences, a longitudinal view of lot-specific patients’ means and lot-to-lot regression slopes showed significant drift over time. Additionally, drift could have been identified earlier by quantifying the cumulative effects of lot-to-lot changes, as recommended in testing the activated partial thromboplastin time for heparin management (7).

Implement Prospective Monitoring

The NBS instituted a prospective monitoring system across all laboratories to assure ongoing high sensitivity for infectious disease agents while also maintaining high specificity (6). Indicators of poor assay performance include initial and repeat reactive rates, discrepancies between duplicate test results, QC failures or trends, instrument problems, wastage, quality incidents, discrepancies with subsequent donor testing, and more. Aggregating data from all testing laboratories increases the sensitivity for early detection of inherent or emerging problems. Five assays have been withdrawn on the basis of findings from prospective monitoring (5, 6). Additional approaches for prospective monitoring include the use of patients’ means or medians in average-of-normals algorithms (8), the proportion of patients’ results outside of reference intervals (2), δ checks, specimen and tube type verifications, and absurdity checks (9). These methods are not applicable to all assays or patient populations, but properly selected tools are useful in quality management (9).

Increase Transparency and Share Data Publicly

The archetype for transparency is commercial aviation. Public reporting of incidents and accident investigations is a critical element of regulatory oversight and a substantial force for technological and procedural improvements (4, 10, 11). In public health, data sharing is vital for identifying and controlling emerging infectious diseases (12, 13). Similar benefits would likely accrue if clinical laboratories were transparent about validation failures and postvalidation issues, but this degree of openness would raise concerns about proprietary information and liability risks if findings were made public.
Combine Efforts of Manufacturers and Laboratories

Algeciras-Schimnich and colleagues call for manufacturers to provide better postrelease monitoring, including the automated gathering of test results from manufacturer-interfaced analyzers, to provide earlier identification of lot-to-lot inconsistencies not detected by manufacturers’ and laboratories’ validation studies (2). Manufacturers and clinical laboratories should also explore opportunities to couple lot-release and lot-to-lot validation into a single effort that incorporates real-world variables.

Promote Research in Validation Methods

Universally accepted guidelines for lot-to-lot validation procedures or acceptance criteria do not exist. Theoretical power calculations often fail to hold true in practice (2). Lot-validation procedures based on evidence from the real world would be a boon to clinical laboratories and manufacturers. “Epidemiological” investigations of clinical, financial, and operational consequences of validation failures could provide hard data to drive methodological, technological, and procedural innovations. Academic institutions, professional societies, and industry would make a valuable contribution by promoting and supporting these endeavors.

W. Edwards Deming stated, “Quality is everyone’s responsibility” (14). In general, this statement rings true for laboratory quality in the same sense that fire safety in Benjamin Franklin’s Philadelphia depended on every individual. In achieving lot-to-lot consistency, Deming’s statement may well be interpreted as a collective responsibility in which the joining of the resources, data, expertise, and efforts could multiply individuals’ 28 g of prevention into metric tons of cure.

Author Contributions: All authors confirmed they have contributed to the intellectual content of this paper and have met the following 3 requirements: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; and (c) final approval of the published article.

Authors’ Disclosures or Potential Conflicts of Interest: No authors declared any potential conflicts of interest.

References