Validation of Computerized Liquid Chromatographic Systems

WILLIAM B. FURMAN and THOMAS P. LAYLOFF
U.S. Food and Drug Administration, Division of Drug Analysis,
1114 Market Street, Room 1002, St. Louis, MO 63101

RONALD F. TETZLAFF
U.S. Food and Drug Administration, 60 Eighth Street, NE, Atlanta, GA 30309

Abstract: Almost all vendors of liquid chromatographs offer dedicated computer-based controller/data-processing systems. These systems are integrated into the operation of the equipment and frequently use proprietary software systems for which complete documentation and program listings are not available to the user. In some instances, data in magnetic media are considered to be "raw data" for regulatory retention requirements. In addition, printed copies of final data used to make decisions and the procedures used to generate these data from raw data must be prepared and retained. In all cases, the systems must be validated to demonstrate that they are suitable for their intended uses. Approaches to validation of these systems are described and discussed.

Validation of computerized liquid chromatographic (LC) systems is of interest from scientific and regulatory viewpoints. A scientist wants assurance of reliable data and computed results. A regulator wants tangible evidence of validation, such as written standard operating procedures (SOPs), raw data, guidelines, and policy documents that may be reviewed during inspections of manufacturing facilities. We present and discuss two approaches to such validation -- modular and holistic -- and describe procedures to satisfy the needs of scientists and regulators. Our remarks are addressed to quantitative analysis of products intended for the marketplace, or components of such products, and exclude qualitative or preliminary investigational work.

Modular Validation

By modular validation, we mean individual calibration or validation of each distinct, important part of a computerized laboratory system; this general approach was discussed by Tetzlaff[1] in a review of U.S. Food and Drug Administration (FDA) requirements and expectations during inspections of such systems. Our concept of modular validation of computerized LC systems is outlined in Table 1, which is meant to be illustrative, not comprehensive. Although calibration of each module may be useful for troubleshooting purposes, such tests alone cannot guarantee the accuracy and precision of analytical results.

Detector Linearity

An attempt could be made to verify that the detector's output voltage is linear with concentration of a standard material. For a UV-VIS LC detector, this test would have to be performed at several wavelengths to determine boundaries of reliability; for example, at which upper and lower wavelengths does stray light start to reduce the concentration range in which the output is linear? Even if these measurements could be made, there would be little assurance they would guarantee linearity when the detector is used for analytical measurements on compounds different from those selected for the detector calibration. Similar arguments apply to other types of LC detectors (fluorometric, electrochemical, etc.).

Detector Wavelength Accuracy

We have often encountered variable-wavelength UV-VIS LC detectors whose indicated wavelength differs from the actual wavelength by considerable amounts, as much as 20 to 30 nm. Loss of wavelength calibration can occur from mechanical shocks during shipment, and we have observed such errors in detectors newly received from manufacturers. Wavelength inaccuracy is often not apparent to the person operating the detector and can lead to substantial errors in interpretation of data. The operator's manual usually gives a quick way to adjust the detector to meet the manufacturer's specification for wavelength accuracy; we urge all laboratories to perform this calibration.

Pump Flow Rate

An attempt could be made to measure flow rates at several pump settings with various backpressures. Such tests might be repeated over several days to evaluate the reproducibility of the pump's flow rate. Yet, such measurements are fleeting and of limited value for two reasons: Critical pump parts -- valves, especially -- deteriorate with time, and pump performance is often influenced by deposits of buffer salts on critical parts. Neither of these effects can be reliably duplicated in standardized tests.

---

[1] Tetzlaff, R. F. (1992) Pharm. Technol. 16 (5), 70, 72, 74-78, 80-83.

Injection Volume

Similarly, measuring within-day and between-day accuracy and precision of the amount of solution injected automatically or manually might be attempted. And again, such measurements could not be assumed to hold up over useful periods of time due to gradual deterioration in injector parts or due to differences in viscosity or surfactant levels among the solutions used to test the injector and those actually used during analyses.

Data Processor; Analog-to-Digital Converter

In a rigorous modular validation, an attempt would be made to measure the accuracy of the analog-to-digital converter's output (numbers) vs. the input signal (detector voltage). However, by connecting the voltage source, electromagnetic interference might be introduced from the test leads, or the load on the input of the analog-to-digital converter might change enough to affect its output accuracy. It is extremely difficult to design electrical test protocols that accurately simulate the conditions present within the instrument when the detector, not the test device, is connected to the analog-to-digital converter. This approach is beyond the capability reasonably to be expected of many laboratories.

Data Processor and System Controller; Correctness of Program Code

If a laboratory has written its own program code or has purchased a custom-written program code for data processing or system control, the source code should be available for scrutiny and validation. Included here are any "macros" that may be written for an LC system; they must be documented and available for review. Minimum documentation of macros (and other custom-written routines) must include: (1) a statement of what the macro is supposed to do; (2) a written description of how it performs its task, including a clear explanation of each step, the formulas developed for custom-written calculations, and references to the vendor's user manual for vendor routines utilized; and (3) name(s) of the person(s) who prepared the macro, documented approval of the macro, date of its installation, date(s) of revision, and validation tests.

Let us now consider the programs supplied by the vendor of a chromatographic workstation -- the programs that process data and control the LC system. Most practicing chemists do not have the ability or inclination to perform systematic software review and validation of the vendor's programs, which may consist of thousands of lines of code. Further, most chemists use integrators or workstations that contain proprietary algorithms that the manufacturer will not share. The "logic" or the "applicability" of these internal programs cannot be verified because a source listing of the code is not available. Even if the manufacturer supplied the source code, it is unlikely that even the most competent programmer could discern all the faults in all the branches and loops of an unfamiliar program.

All equipment manufacturers should have validated their software before releasing it. Some vendors will provide useful information on the behavior and performance of their programs. Inquire about tests the vendor has run on the program code and request copies of the results.

Critique of the Modular Approach

Most chemists cannot evaluate the proprietary programs running in their computerized LC systems, so they require other procedures for system validation. Individual calibrations of

injectors, pumps, detectors, analog-to-digital converters, and the like are admirable but difficult. Although they may be useful for diagnosing problems with LC systems, they alone cannot guarantee reliable analytical results for the reasons given earlier.

Holistic Validation

In a holistic validation (Table 2), we use a series of tests to measure and evaluate the performance of the entire computerized LC system under the conditions of its intended use: chemical analyses. The minimum holistic validation is a two-step process: initial calibration and running calibration. Each laboratory should develop and include other tests to cover its own special requirements.

Initial Calibration--Linearity

Before any samples are run, the entire system -- injector, pump, column, detector, data-acquisition device, and data-output device -- must be calibrated for linearity (or for reproducible non-linearity) by injecting several standard solutions of different concentrations. Use at least four standard solutions to generate this linearity curve; 0 must not be used as a "data point." Initially, the linearity test must be run daily. If the linearity curve is proven reproducible over several days, the test interval may be lengthened: every other day, weekly, or as specified in SOPs. Documented evidence must be obtained to establish that the specified test interval keeps the LC system in control.

How shall the span of concentrations be chosen for these standard solutions? For routine work, in which the result is expected to fall within a known range, the standard curve must cover this range plus a bit more for safety. As an example, when analyzing individual tablets by the United States Pharmacopoeia (USP) procedures, results are normally expected to fall between 85 and 115% of the labeled amount. Linearity standards could be set up to cover 50 to 150% of label, for example. Each laboratory must prepare its own written requirements for calibration span of linearity standards and must not report analytical results calculated from data that fall outside this span.

In research work the result cannot always be anticipated. Examine the linearity of the system over several concentration decades to assure usable linearity or to set up guidelines for dilutions or concentrations that are within the acceptable range.

Initial Calibration--System Precision

At first, measure the system precision daily by making at least six replicate injections of a single standard solution and by calculating the standard deviation of the standard responses. Only after acceptable system precision with replicate injections of a standard solution is obtained, should sample analysis proceed. If the system precision is proven reproducible over several days, the test interval may be lengthened -- every other day, weekly, or as specified in the SOPs -- as long as documented evidence is available that adequate system precision is maintained over the chosen interval.

Legally recognized methods, such as those in the United States Pharmacopoeia[2] or Official Methods of Analysis,[3] often give requirements for system precision and procedures to measure it. If these guidelines are lacking, the "acceptable precision" of the process to be controlled must be compared to the precision of the LC system. Unless the latter is much better

[2] United States Pharmacopoeia (1990), 22nd Rev., U.S. Pharmacopeial Convention, Rockville, MD.
[3] Official Methods of Analysis (1990) 15th Ed., AOAC, Arlington, VA.

than the former, the LC system will be of little use. As a rough guide, currently at the Division of Drug Analysis we routinely obtain precision, expressed as a relative standard deviation, of 1% or less from a given LC system and operator.

Running Calibration

After satisfactory linearity and precision data are obtained and sample analysis has begun, how can quality control be sustained? A standard solution should be run at regular intervals -- every two hours or every five samples -- or at the interval given in the written SOP. Calibration intervals are arguable, but documented evidence must be provided to establish that the interval selected and specified keeps the LC system in control.

The goal is to document that the system is not drifting or has not experienced a catastrophe. The basic equation is as follows: The more often the system is calibrated, the less likely loss of quality measurements will occur. If the system does develop a problem, how soon will it be discovered? How many samples will have to be rerun since the last good calibration? Will those reruns cost more than the additional calibrations would have? Each laboratory has to balance this quality assurance equation in the most cost-effective way.

Definition and Storage of "Raw Data"

The term "raw data" is broadly defined in FDA's Good Laboratory Practices (21 CFR 58.3(k)).[4] We excerpt the following portions of interest:

Raw data means any laboratory ... records ... that are the result of original observations ... and are necessary for the reconstruction and evaluation of the report of [a] study ... Raw data may include ... computer printouts, magnetic media, ... and recorded data from automated instruments.

In the context of computerized LC systems, we further propose the following definition of "raw data": Of the data captured by the data processor, "raw data" are all those that can be saved and accessed later.

First consider the simplest data processor, an ordinary integrator that furnishes at least 3 items in a printed report: a chromatogram, a list of peak heights or areas vs retention time, and a table of "parameters" used to obtain the chromatogram and data (attenuation, noise rejection, slope sensitivity, construction of baseline, etc.). This simple printout constitutes all the "raw data" that can be saved and accessed later.

Now consider the other end of the scale -- a computerized workstation. The workstation can give a very abbreviated output -- a chromatogram, a list of peak heights or areas vs retention time, and a table of "parameters" -- just as with the simple integrator. But now not only this "finished presentation" is available but also additional raw data residing on the hard drive -- the detector readings and corresponding times from which the computer program calculated and printed the report. The workstation could capture much more data; UV-VIS spectra of each compound eluting from the column; and obtain spectra taken at the beginning, middle, and end of each emerging peak, etc. If the workstation is instructed to do this, much more raw data will be available on the hard drive for each run.

How long must raw data be saved? The minimum period is the time specified by the current FDA regulations applying to each project. When considering computerized workstations and digitized raw data (and with floppy disks or tape cartridges so cheap these days), it does not make sense to throw away any raw data, especially not the raw data that led to the results reported as final and correct.

---

[4] Code of Federal Regulations (1993) Title 21, Part 58, Section 3(k), pp. 247-248.

Consider the following scenario: a couple of small peaks are detected eluting ahead of the main compound. The current method does not mention these peaks, so they are ignored. A year later, the research group announces it has identified those two compounds as highly obnoxious materials, and management wants to know how to determine how much of these materials were in all those lots analyzed last year. Which is preferable: rerunning all the samples to integrate those little peaks, or having the computer retrieve the old files of raw data and integrate them?

So, our attitude toward saving raw data is very conservative; save all of it. If all that is available is a simple printed chromatogram and table of results from the integrator, save them. If the computerized workstation was instructed to get a chromatogram at a given wavelength but also to obtain and save the complete UV spectrum of each eluant, save all of those raw data as well.

If possible, equip each computerized workstation with a tape backup card and purchase a tape backup drive. The drive is the most expensive component, but one drive can serve many workstations. Adopt a consistent way of naming files that matches file names with sample numbers or research projects and that prevents overwriting of old files by new files. Then, at the interval specified in written procedures, save all raw data onto a tape cartridge. It is like having a time machine; the workstation can be put back the way it was on January 7, 1991, and again examine the raw data saved.

Definition and Storage of Data Used To Make Decisions

An analyst frequently works with the raw data and adjusts parameters to generate a suitable chromatogram and to obtain acceptable measurements of retention times and peak heights or areas. From these intermediate results the analyst (or the workstation) calculates the final results used for decision making. Examples of such decisions include whether to recalibrate the LC system, whether a lot passes or fails requirements, whether to adjust a manufacturing process, etc. These data and the conditions used to obtain them should also be saved as paper copies for the length of time specified by the current FDA regulations applying to each project.

Although the trend is toward "paperless laboratories," there are several strong arguments for retaining printed copies of all data used to make decisions. First, given a set of raw data, it is unlikely that two qualified independent operators of a computerized LC system will make exactly the same choices about how to treat the raw data to obtain finished results. The operators might choose different data-measurement parameters for smoothing data, noise-rejection levels, points to start and stop integration of peak areas, etc. Unless chromatograms, data-measurement parameters, final data (retention times, peak areas or heights, etc.), and methods of calculating final results are saved as paper copies, it may not be possible to fully reconstruct how the final decision was made. It is also wise to document any problems operators might have encountered in generating final data and results from raw data, such as integration difficulties or any other comments. Second, the LC system, including its workstation, will eventually become obsolete or defective. If the computer or its programs are lost, regenerating final data and results from raw data may be impossible. Third, review by supervisors, quality assurance staff, and FDA investigators will be greatly facilitated by available paper records that document all final decisions.

Responsibilities

Although regulatory liability for failure to calibrate equipment is shared by management within a firm, the primary responsibility for instrument calibration resides with the chemist or technician who produces the data. He or she may receive assistance from repair personnel, but the ultimate responsibility stays with the person generating the data. Management's responsibility includes guidance on what is acceptable; an individual chemist or technician must not have free rein to decide what constitutes good data and what does not.

Why is responsibility placed with the person acquiring the data and with the supervisors to follow SOPs and not with repair personnel or with the manufacturer of the equipment? Let us first dispense with "manufacturer responsibility." If the company who made our scientific apparatus could be sued because we cannot control the quality of the data it gave us, we would soon be forced to build our own liquid chromatographs and write our own computer programs to run them. The manufacturers would quickly abandon us for greener, less litigious pastures. Besides, no self-respecting laboratory manager would abrogate his or her responsibility in this manner.

Next, repair personnel do not produce the data and provide the final results to management. How can we hold repair personnel responsible for calibration of the LC system? Even if we did, there might be subtle differences between the test conditions used by repair personnel and the conditions existing when the LC system is being used to analyze samples, as detailed earlier under Modular Validation.

Documentation

For an institution to make quality measurements, the very highest management levels must support this commitment. If such support is present, an institution will manifest this commitment in several ways, among them written SOPs that clearly state what management expects and how the staff fulfills these expectations. As a minimum, a laboratory should have a written SOP for each element of validation it elects to implement (see Tables 1 and 2 for example elements). Management must approve these procedures in writing; the procedures must be in the hands of the operating staff; and the institution's quality assurance unit must check to see that the procedures are being followed.

Granted, reviewing and validating proprietary computer programs purchased from commercial vendors is usually not possible. But it is best to have documentation on the applicability and limitations of such programs. As a minimum, the user manuals would suffice that were supplied by the firm who sold the proprietary programs. These manuals, plus additional documentation on program behavior and performance tests if supplied by the manufacturer, should be referenced in the written SOP. State that these programs are to be operated as intended and within the limitations specified in the manufacturer's manuals and documentation.

As program revisions are received from the equipment manufacturer or from other program vendors, their installations should be documented: date installed, which systems were upgraded, etc. References to new user manuals and new manufacturer's documentation on program behavior and performance tests should also be updated in the laboratory's written operating procedures. If a laboratory has several similar instruments, it is desirable to use the same revision of the program on all of them. If there are "compatibility problems" among revisions, using the same revision throughout an installation may become a necessity.

Regardless of the data processor employed, it is important to describe raw data and to include a definition of raw data in the operating procedures for the specific types of testing conducted.

Summary

Even if all elements of a modular validation were passed, there would be no direct assurance that final analytical results are correct. We urge all laboratories to test variable-wavelength LC detectors for wavelength accuracy. Otherwise, we suggest the holistic approach to calibration and validation of computerized LC systems. Because reviewing and validating proprietary programs is usually unfeasible, other methods must be used to verify them. Chemists should use chemical methods, as outlined, to achieve these goals.

Acknowledgments

References

Table 1. Modular validation of computerized LC systems

Detector (UV-VIS)
  Is output voltage linear with concentration?
  For variable-wavelength detectors, are wavelength settings accurate?
Pump
  Are flow rates accurate and reproducible?
Injector
  Are injected volumes accurate and reproducible?
Data Processor (Integrator, Computer)
  Is analog-to-digital converter accurate?
  Is program code correct? Does it give correct calculated results (plots, peak heights or
   areas, assays, etc.)?
  Are the macros coded correctly? Validated? Documented?
Controller (Computer)
  Is program code correct?
System Documentation
  Is system documentation in place?

Table 2. Holistic validation of computerized LC systems

Initial Calibration--Linearity
      Use at least four standard solutions.
      Concentration range of standards must span anticipated results, plus safety margin.
      Run standards daily[5] before starting sample analysis.
Initial Calibration--System Precision
      Calculate precision daily[5] from at least six replicate injections of a standard solution
         before starting sample analysis.
Running Calibration
      Run a standard solution at specified time intervals or after a specified number of sample
         solutions.
Data Processor (Integrator, Computer)
      Are the macros coded correctly? Validated? Documented?
System Documentation
      Is System Documentation in Place?

---

[5] Or at intervals specified in SOP.