Investigation of Out-of-Specification Results

Alex M. Hoinowski, Sol Motola, Richard J. Davis, and James V. McArdle*

Despite a multitude of educational resources, difficulties and inconsistent approaches to out-of-specification investigations persist. The Analytical Research and Development Steering Committee of the Pharmaceutical Research and Manufacturers of America (PhRMA) included the topic of conducting out-of-specification investigations in its annual workshop held in September 2000. The authors summarize the results of the workshop in this article.

PHOTODISC INC.

Alex M. Hoinowski, MS, is a senior consultant for The Quantic Group, Ltd., Livingston, NJ. **Sol Motola, PhD,** is the assistant vice-president of technical and regulatory affairs for Wyeth Ayerst Global Pharmaceuticals, St. Davids, PA. **Richard J. Davis** most recently was the senior vice-president for quality assurance and regulatory compliance for DuPont Pharmaceuticals Company, Wilmington, DE. **James V. McArdle, PhD,** is the vice-president of analytical development and quality for Isis Pharmaceuticals, Inc., 2292 Faraday Ave., Carlsbad, CA 92008, tel. 760.931.9200, fax 760.603.4655, jmcardle@isisph.com, www.isip.com.

*To whom all correspondence should be addressed.

Why are out-of-specification investigations still a leading cause of warning letters issued by FDA (1)? Many workshops, seminars, consultants, and guidances are available, yet difficulties and inconsistent approaches to out-of-specification investigations persist. Numerous working practices and various degrees of understanding of current expectations in this area likely have led to many different approaches to investigating out-of-specification results. For this reason, the Analytical Research and Development Steering Committee of the Pharmaceutical Research and Manufacturers of America (PhRMA) included the topic of conducting out-of-specification investigations in its annual workshop held in September 2000. Representatives from PhRMA member companies met to consider this topic, share current practices, and agree on an acceptable analytical practice (AAP) that represents good science and conforms to current regulatory guidelines and expectations. The output of that workshop is presented in this article.

Background

The guiding principles for out-of-specification investigations are based on a legal ruling by Judge Wolin in 1993 (2) and the draft FDA guidance that followed in September 1998 (3). Judge Wolin presided over a case brought by FDA enforcement action against a generic-drug manufacturer. Among the issues that FDA cited were averaging out-of-specification values with in-specification values to get a passing result; maintaining an ineffective program for process validation; discarding raw data; conducting multiple retests with no defined end point; performing inadequate failure investigations; and lacking method validation. Judge Wolin's responsibility was to listen to experts from both sides and craft an opinion that was based on the testimony and the law. In addition, he had to interpret the current regulations about good manufacturing practices.

The trial itself generated 2300 pages of testimony and 400 exhibits and declarations. Judge Wolin acknowledged that the pharmaceutical industry is a business with legitimate business motivations. He also acknowledged the very important task that FDA has as regulator of this industry. The judge found that the company was mired in uncertainty and conflicting opinions about good manufacturing practices and presented a troublesome attitude to the court.

His final decision reflects good common sense. The judge stated that the history of the product must be considered when evaluating the analytical result and making a determination

about the release of the product. Of course, analytical results that are wrong may arise. The exact cause of the error often is difficult to determine, and expecting that the cause of errors always will be determined is unrealistic. The inability to identify the cause of the error will affect retesting procedures, but it will not affect the inquiry required of the initial out-of-specification result.

Judge Wolin placed restrictions on the application of outlier tests. A firm may not frequently reject results on the basis of outlier tests, and outlier tests may not be applied to results of chemical testing. In fact, Judge Wolin cited USP as governing the use of outlier tests. The draft guidance does allow the use of outlier tests for providing perspective concerning the probability of obtaining the suspect result as part of the overall evaluation of the batch.

Judge Wolin stated that the firm should have a retest policy and protocol and that the protocol must define the point at which testing stops and evaluation occurs. He did allow retests, but only when the laboratory investigations were completed and an analyst error documented. Retests also could be justified when a formal process investigation detected no process or nonprocess-related errors and a review of the laboratory test was inconclusive. He also stated that the firm could not conduct two retests and base a release on the average of three results, nor could the firm simply assume a sampling or preparation error as a justification for retesting and resampling. Judge Wolin directed that retests were to be done on the same sample, not a different sample, but retests might be conducted on a second aliquot from the same portion of sample or on a portion of the same larger sample previously collected for laboratory tests.

Invalid testing caused by known laboratory errors

An obvious error that invalidates the results may occur during testing. Such errors are recognized easily without a formal investigation. A known error caused by the analyst or by instrument failure does not require the performance of a formal laboratory investigation, whether or not data have been generated. Testing should be stopped, the supervisor notified, and the data invalidated. Examples of these types of errors include transcription errors, miscalculations, incomplete transfer of material, and incorrect settings of instrument parameters. The analyst should document the error on the raw data record, and the supervisor should initial and date the record. In the case of clerical errors such as transcription errors or miscalculations, the analyst should correct the error and continue with review of the data. In the case of laboratory errors such as an incorrect dilution or a detector set on the wrong wavelength, the analyst should invalidate the result. Testing may continue with no further consideration of the invalid result.

If a laboratory investigation is initiated before the discovery of the known error, the investigation should be completed as a means of documenting and addressing the error. The draft guidance further discusses the responsibilities of the analyst in this situation.

The formal laboratory investigation

The primary objective of the investigation of an out-of-specification result is to determine either that an assignable cause exists or that an assignable cause cannot be identified. An assignable cause is a documented and scientifically justified determination that the discrepant result can be traced to laboratory error. Note that atypical results (those that are unusual on the basis of experience, trending, or data review but still are within specification) also warrant an investigation and may require the implementation of suitable corrective actions. The rest of this discussion will focus on out-of-specification results, but many of the principles and recommended courses of action also will apply to the investigation of atypical results. An out-of-specification result is one that falls outside of the test's acceptance criteria in, for example, filed applications, approved marketing submissions, or official compendia.

Phase 1 of the investigation. As emphasized in the draft guidance, the analyst detecting an out-of-specification result must report that result promptly to his or her supervisor. The analyst and supervisor must then immediately conduct an investigation of the result. If no obvious error is revealed, an out-of-specification result must be reported to the quality assurance unit within a specified period of time, usually one or two business days. This report to the quality assurance unit constitutes the beginning of the formal investigation for most companies. However, some companies regard the notification of the supervisor by the analyst as the start of the formal investigation. The firm's standard operating procedure (SOP) should be unambiguous about the point at which the formal investigation begins.

The analyst and the supervisor conduct phase 1 of the laboratory investigation to determine whether or not the out-of-specification result is assignable to the testing laboratory. This part of the investigation should be completed in a short and predefined period of time, usually one or two business days, and documented in a report of the investigation. If this part of the laboratory investigation cannot be completed within the predefined period of time, then an interim report should be issued within the time limit.

A checklist of possible sources of error may aid this part of the investigation. Checklist items to use in reviewing the analyst's notebook include

- Was the correct amount of sample taken?
- Were the sample dilutions correctly performed?
- Were other reagents and solutions correctly prepared and used?
- Are the calculations correct?
- Were the correct volumetric flasks and pipettes used?

Similar checklists can be developed to identify sources of error in sample integrity and handling, instrument performance, method validation, analyst's technique, the test procedure, standards, and instrument output.

At this early stage, conducting limited testing of the samples that led to the out-of-specification result may be helpful. The draft guidance states that such immediate assessments "allow more credibility to be given to laboratory error theories." This limited testing usually will not suffice to determine the acceptability of a batch, but it should be captured in the report about the laboratory investigation.

If the cause of the out-of-specification result can be assigned to the testing laboratory, then the original sample is tested again.

If the result of the retest meets the acceptance criterion, then the original out-of-specification result is invalidated. The root cause or the assignable cause of the out-of-specification result is documented, and the effect of the finding on other data is determined. Corrective actions may be taken to prevent recurrence of this source of error. The investigation then is closed.

Phase 2 of the investigation. The second phase of the laboratory investigation begins if an error assignable to the testing laboratory cannot be identified during phase 1. The rationale for the planned investigation should be documented. For example, the documentation should include the tests that will be conducted, justification for retesting or resampling if planned, the number of samples to be tested, and the criteria for evaluation of the data. The protocol for retesting or resampling should be included. The quality assurance unit should be involved in defining the retest protocol and in determining whether or not retesting is justified. The draft FDA guidance states that the firm's SOP should specify the number of retests to be performed on a sample. However, Judge Wolin's decision stated that the number of retests is case specific. The scientific basis for specifying the number of retests for all situations in an SOP is difficult to understand. Such a policy cannot take into account proper statistical design of the experiments that will be needed to obtain an outcome that carries adequate confidence. One paper that attempts to address this issue recently has appeared (4).

If the cause of the out-of-specification result is identified conclusively, then this phase of the investigation may be concluded. For example, a low assay might be explained by the appearance of a significant new impurity. If the investigation determines that the out-of-specification result cannot be assigned to the testing laboratory, then an investigation of the manufacturing process may be warranted. If that investigation does not reveal a manufacturing error, and sufficient analytical data exist from the retest to decide that the original out-of-specification result arose from an unassignable laboratory error, then one may conclude that the original out-of-specification result is not representative of the batch. Product history also should be considered at this point.

The results and conclusions of this phase of the investigation of the out-of-specification result should be documented in a laboratory report. An assessment of other areas or tests that might be affected should be presented. For example, other manufacturing operations or sites may be affected.

Phase 3 of the investigation. The final phase of the investigation involves a compilation of the investigation of the out-of-specification result and the investigation of the manufacturing process, if relevant. The primary objective of this phase of the investigation is to identify the most probable cause of the out-of-specification result. If it was determined that the out-of-specification result was caused by the testing laboratory, then perhaps equipment maintenance or additional training is indicated, for example. The effect of the out-of-specification result on other batches, ongoing stability studies, validated processes, and testing procedures should be determined. The final report also should specify individuals responsible for corrective actions to be taken to mitigate or prevent the recurrence of this type of out-of-specification result.

Phase 3 of the investigation should be completed within a predefined period of time, typically 30 calendar days from the discovery of the out-of-specification result. Justification should be presented when completion by the predefined deadline is not possible.

Additional concerns exist when out-of-specification results are generated regarding a commercially distributed product. Senior management in the quality assurance unit should be notified within a short period of time, usually 24 hours, in such an instance. A field-alert report must be issued to FDA within 72 hours of the discovery of an out-of-specification result if at that time the cause of the result cannot be attributed to the testing laboratory.

Responsibilities

The previous discussion makes clear that the analyst, the analyst's supervisor, and the quality assurance unit all have important responsibilities in the investigation of an out-of-specification result. A useful discussion of some of these responsibilities is provided in the draft FDA guidance. Some additional responsibilities of the analyst that may help to avoid the generation of out-of-specification results include

- verifying that all paperwork is in order and matches the label information
- verifying that the proper test method is used
- retaining the original test preparation until data and results have been checked
- discontinuing testing and notifying a supervisor as soon as possible if a problem is suspected.

Additional responsibilities of the supervisor include

- ensuring proper sample management within the laboratory
- ensuring each analyst is properly trained
- notifying the quality assurance unit when an out-of-specification result arises
- ensuring documentation and implementation of retesting protocols
- ensuring that the investigation is conducted and documented within the specified time frames.

Challenges and points for discussion

What is a reportable value, and do only reportable values trigger an out-of-specification investigation? The draft FDA guidance states that "a test might consist of replicates to arrive at a result. For instance, an HPLC assay result may be determined by an average of the peak responses from a number of consecutive, replicate injections of the same preparation (usually two or three). The assay result would be calculated using the peak response average." For example, consider an HPLC assay that has an acceptance criterion of ⩾98.0% for the average of four injections. If one obtains individual values of 98.0, 98.1, 98.2, and 97.9% by the validated method, is this an out-of-specification result? Nearly all meeting participants agreed this is not an out-of-specification result. The important point here is that the four individual values represent single data points, and the average of the four data points is the result of the test. The result of the test is compared with the acceptance criterion and judged to be within specification. No investigation is warranted.

Of course, unexpected variation in individual values should trigger an investigation. Thus, if a test has an acceptance criterion of >90.0%, and one obtains individual values of 96.0% and 85.0%, almost all participants agreed that these values should not be averaged, and an out-of-specification investigation should be initiated. However, when informed that this hypothetical test had a relative standard deviation (RSD) of 15%, a small majority of meeting participants felt that an out-of-specification investigation need not begin.

Written method descriptions should be specific about reportable values, but situations still may exist in which investigation of individual out-of-specification data points may be necessary. An example of such an occurrence would be when the variability of individual data points that go into a reportable value is much greater than the expected variation for the method. Methods also should be specific about the treatment of individual data points that are within the expected variation of the method but outside of the acceptance criterion. For example, the true value of the purity of a hypothetical sample is 98.4%, and the assay method has a known variance. The lower limit of the acceptance criterion is 98.0%. The method should describe the procedures for averaging data points as well as describe statistical procedures for identifying highly atypical individual replicate data points. Continuing the previous example, consider the replicate values 97.9, 98.2, and 98.4%. The average is 98.2%, within the specified limit. One data point for the assay of the sample is 97.9%, which is below the limit. The RSD is 0.26%, not atypical of the inherent variability of the method. Because the RSD is typical, the data point that is below the limit is included in the average and no investigation is warranted. In other words, this result is to be expected as part of the normal analytical variation.

When is averaging acceptable? The draft FDA guidance states that if a series of assay results (not single data points) is part of a test procedure, then all test results should conform to the acceptance criterion to report an average of the results. Note that each of the individual results of this series may comprise single data points, and, as in the previous example, no requirement exists that each data point meet the acceptance criterion as long as no unexpected variation in the single data points exists. Thus, if three determinations of three sample preparations gave two results that met the acceptance criterion and one result that did not, nearly all meeting participants would not average these results but would begin an out-of-specification investigation. If the investigation revealed that the cause of the out-of-specification result could be assigned to the testing laboratory, then that finding would be used to invalidate the out-of-specification result, and that result would not be considered further. If the out-of-specification investigation did not reveal an assignable source of error or any discrepancies in the manufacturing record, then retesting would be considered.

Proper statistical consideration should be given to the determination of how many additional test results are needed to judge whether or not the out-of-specification result is an outlier (4). Should the out-of-specification result be determined to be an outlier, it is invalidated and not included in the average of all test results. If the out-of-specification result is shown not to be an outlier, then testing stops, and all of the data are used to guide decisions about the disposition of the batch. Consideration of the out-of-specification result also should encompass product history.

All meeting participants agreed that averaging is not appropriate for a test that is meant to measure variability such as a test for content uniformity. The draft FDA guidance is clear on this point.

Although the draft guidance offers advice about the notion of reportable values and the use of averaging, a number of meeting participants reported inconsistent treatment of these topics by regulators. For example, a number of meeting participants reported that recent exchanges with regulators led them to the conclusion that the regulators have taken the view that averaging cannot be used, period. In response, some meeting participants suggested determining HPLC assay values on the basis of a single sample preparation and a single injection, and a number of participants stated they already have adopted that practice.

What strategies may be used to establish acceptance criteria that will not lead to unwarranted out-of-specification results? First of all, acceptance criteria must ensure the safety of the patient. Individual impurities, degradation products, and residual solvents must be limited to those levels that provide an adequate margin of benefit-to-risk for the treatment under consideration. Obviously, analytical methods must be developed and validated to provide the appropriate level of confidence that impurity profiles are adequately determined. Acceptance criteria also must be established with an understanding and acceptance of statistical procedures and significant figures.

Acceptance criteria also must ensure the efficacy of the product. All relevant data must be included in the process of establishing acceptance criteria. The development history of the product should be taken into consideration, as should any experience in manufacturing. A number of meeting participants reported recent exchanges with regulators who were willing to consider only a subset of development history or manufacturing experience. Ironically, some regulators labeled as "outliers" certain data that firms regarded as reflecting the established range of process capability. As is easily predicted, this narrow approach to establishing acceptance criteria led to frequent out-of-specification results and subsequent investigations. Some of these firms then felt under threat of being cited for conducting manufacturing operations that are not under control, when the real issue was having improperly established acceptance criteria. Meeting participants also reflected the notion that some regulators believe that unless some percentage of products fails, acceptance criteria are not set tightly enough. Most meeting participants found this idea to be unsettling and without merit.

All meeting participants readily endorsed the principle that acceptance criteria must ensure both the safety of the patient and the efficacy of the product. Acceptance criteria also must reflect all relevant data that contribute to understanding process capability.

What are some other acceptable strategies for limiting out-of-specification results? The proper use of system suitability samples can establish additional confidence in analytical results.

Historically, many firms' system suitability requirements tended to be broad. Perhaps this practice grew from the desire of the quality control laboratory to limit the occurrence of system suitability failures. Meaningful system suitability criteria should be determined during development and finalized with the quality control laboratory. Only necessary criteria should be established for the quality control laboratory. A failed system suitability test may lead to an investigation, but this investigation will be different from an out-of-specification investigation. A failed system suitability result leads to a halt in the test if no data have been collected or the collected data have been invalidated. An out-of-specification investigation of data generated following a system suitability failure should not be required.

Another strategy for limiting out-of-specification results is the proper qualification of equipment and analysts. Resources invested in these activities will realize a return in reduced retesting and investigation. An understanding of the critical aspects of a given method and an understanding of the critical calibration requirements that an instrument must meet to perform that method are essential. All firms have active training programs for their analysts. However, some firms believe that general training in a technique such as HPLC is sufficient to qualify an analyst to run most HPLC methods, but other firms provide specific training for each HPLC method that an analyst needs.

Careful and timely transfer of analytical technology between the development laboratory and the quality control laboratory is another valuable tool for limiting the generation of out-of-specification results. An understanding of interlaboratory variability is key to this effort. Training and qualification, as discussed in the previous paragraph, also are important elements of technology transfer. A sound program of technology transfer will anticipate potential out-of-specification situations before the marketing application is filed.

The use of internal or in-house limits (also called action limits) on some parameters may be effective in limiting the number of out-of-specification results. Internal limits may be effective when used consistently in conjunction with SOPs. The SOP defines the investigation required for not having met an internal limit, but that investigation and its documentation may be different from those required for an out-of-specification result. A significant number of meeting participants expressed concern about this approach. The concern arises from recent comments made by field staff of regulatory agencies wherein no distinction was allowed between in-house limits on quality attributes and acceptance criteria concerning a specification. None of the meeting participants could understand this stance, and none could find a basis for it in statute or guidance. In fact, most recognized that the ability to detect, investigate, and correct declining quality attributes before those attributes fall out of specification is good manufacturing practice that should be encouraged.

Another strategy for limiting the occurrence of out-of-specification results is to trend data. Data trending can be of two types: out-of-specification findings and lot-to-lot data. Most meeting participants agreed that the benefits of trending are not being realized and that information available from trending data was not always being used to actively manage the laboratory. For example, trending of out-of-specification findings may reveal that certain instruments or even certain analysts are associated more often with out-of-specification results than are others. Trending of lot-to-lot data may reveal declining quality and allow steps to be taken before an out-of-specification result is generated. Trending of data is often left to the laboratory supervisor, but most participants felt that the quality assurance unit should have some role in this function.

Smart SOPs

The SOP that governs out-of-specification investigations should define what is an assignable cause of an out-of-specification result and help the analyst and the supervisor uncover laboratory error. The SOP also should define the options for what to do in the event that no assignable cause is discovered, describe finite reporting times for the various phases of the investigation, and describe the internal documentation that will demonstrate that the SOP was followed. The SOP should emphasize how essential it is that all data generated during testing and during the subsequent investigations be retained.

The SOP should include careful consideration of measurements that may be exempt from out-of-specification investigations. For example, data generated during training or generated under a technology transfer protocol may be excluded if the suspect data were generated by a person not yet qualified to perform the method. Similarly, data generated on intentionally stressed samples should not have to be investigated as out of specification. The SOP also may distinguish between out-of-specification investigations and out-of-trend investigations. These investigations may differ, but an SOP should define the procedures.

A majority of meeting participants favored having separate SOPs for development and commercial manufacturing. Maintaining one SOP for both development and commercial manufacturing was favored by some because the two procedures will have a common investigation process. However, most felt that separate SOPs are preferred because, although the investigation process is the same, various actions may result from the investigation. During development the analytical laboratory has a major responsibility to understand the appearance of new impurities or degradation products or a change in product performance such as dissolution. The analytical development laboratory also has the responsibility to improve continuously the suite of methods used during testing. Improvements in analytical methods may lead to the apparent discovery of new impurities that upon investigation prove to have been present in early batches. In addition, more flexibility exists during development than during commercial manufacturing to improve acceptance criteria as more is learned about the drug substance or the drug product.

In general, the analytical development laboratory has a major responsibility to contribute to knowledge of the development candidate to enable future reliable commercial manufacturing. The commercial analytical laboratory must find an answer to the question of whether the product should be released for the commercial market. In addition, the firm has important responsibilities should an out-of-specification result be obtained on a product already distributed for sale.

Summary

The essentials for conducting an investigation of an out-of-specification result in a compliant, reliable, and timely manner include having a well-organized and understood process; having well-informed and well-trained analysts, supervisors, and manufacturing personnel; and maintaining clear and concise documentation. The firm's procedure for investigating out-of-specification results should describe the various phases of the investigation and the recommended timing for the completion of the phases. The types of errors that may arise and how to deal with them should be defined. Checklists may help determine the presence or absence of certain types of errors. The procedure also should provide for the thorough documentation of the investigation.

The best practice is to minimize the generation of out-of-specification results. Careful description of methods and reportable values, including appropriate system suitability parameters, can help to avoid out-of-specification results. Qualification of equipment and analysts and rigorous transfer of analytical technology will help to ensure that a laboratory is fully capable of performing an analytical method.

Acknowledgments

We thank Drs. William Gore, Galen Radebaugh, and Jean Wyvratt for their expert assistance in leading the breakout sessions and capturing all the valuable input from the meeting participants. We thank Dr. Jeffrey D. Hofer for his helpful discussion of statistical issues.

References

1. F-D-C Reports, Inc., *The Gold Sheet,* **34** (3), 1–22 (2000).
2. "FDA's Summary of Judge Wolin's Interpretation of GMP Issues Contained in the Court's Ruling in *USA vs. Barr Laboratories*," (GMP Institute), www.gmp1st.com/barrsum.htm.
3. "Guidance for Industry, Investigating Out of Specification (OOS) Test Results for Pharmaceutical Production" (Food and Drug Administration, 4 September 1998), www.fda.gov/cder/guidance/1212dft.pdf.
4. J.D. Hofer and E.C. Rickard, "A Guide to Determining Sample Sizes and Outlier Criteria in Analytical Investigations," *Pharm. Technol.* **25** (3), 52–66 (2001). **PT**

FYI

50th-edition Chem Show

The International Exposition Company (IEC) has announced that the 50th edition of the biennial Chemical Processing Industry trade show (the Chem Show) will be held 18–20 November 2003 in New York, New York.

The IEC reports that the show's special events, exhibits, promotions, and conference programming will commemorate industry milestones from the past 100 years.

For more information, contact the Chem Show, 15 Franklin St., Westport, CT 06880, tel. 203.221.9232, fax 203.221.9260, info@chemshow.com, www.chemshow.com.