Invalid rates workshop

Workshop Leaders: Sarah Kennett[1], Hermann Wätzig[2], Cari Sänger-van de Griend[3]

Workshop Design and Data Evaluation: Brian Nunnally[4], Kai-Jorrit Maul[2], Nadine M. Ritter[5], Richard Rustandi[6]

Scribes: Joop Waterval[7], Bernd Moritz[8], Jiann Kae Luo[9]

1 Food and Drug Administration, Silver Spring, MD, USA

2 Technical University Braunschweig, Braunschweig, Germany

3 Kantisto BV, Baarn, the Netherlands;
Uppsala University, Div. of Pharmaceutical Analytical Chemistry, Uppsala, Sweden;
ACROSS, University of Tasmania, Hobart, Australia

4 Biogen, Research Triangle Park, NC, USA

5 Global Biotech Experts, LLC, Germantown, MD, USA

6 Merck & Co., Inc., West Point, PA, USA

7 Merck, Sharp & Dohme, Oss, The Netherlands

8 F. Hoffmann-La Roche Ltd., Basel, Switzerland

9 Regeneron Pharmaceuticals, Inc., Tarrytown, NY, USA

Introduction

A central component of a good analytical method is its ability to perform reliably and consistently over time, providing results in a timely fashion in order to inform on decision making. The goal of method development is to produce appropriately reliable test methods. But how do we determine what is "reliable" for method performance? One measure would be the invalid rate of an assay. Everyone agrees a test with a high rate of invalid analytical results is undesirable. Reanalysis runs are costly in terms of reagents, material and time (by analysts, peer reviewers, QA reviewers, etc.), magnifying the impact of potential delays in issuing reportable results on test samples. However, while the invalid rate of an assay should be low, there is no harmonized definition of what constitutes an invalid assay, and so analytical scientists risk talking at cross purposes when discussing how to limit invalid rates.

In late 2015, CASSS took the first steps to elevate and formalize the discussion on invalid rates. In response to growing interest at CASSS meetings for an objective measure of the invalid rates for analytical methods used by many pharmaceutical industries, we undertook a survey to understand how invalid rates are defined and tracked at different companies. Since invalid rates are likely defined

differently for each analytical assay, we started with the example of capillary gel electrophoresis, or CE-SDS, a routine assay used in biopharmaceutical companies for purity analysis of drug substances.

The survey was divided into two parts. The first part inquired after the definition of invalid rates, the usefulness of tracking these numbers and the potential benefits in harmonizing definitions within the (bio)pharmaceutical industry. The second part concerned the use of CE-SDS and the tracking of invalid rates for this methodology. Participants provided the invalid rates within their organization (either measured or estimated) for CE-SDS analysis, as well as potential or identified root causes. Though the number of questions was kept low to limit the burden on survey participants, the responses still raised several interesting points.

Results of the survey were presented at the 2015 CE Pharm meeting, and attendees were invited to participate in a workshop where we focused on 3 questions central to this issue:

1) What should constitute an invalid assay?

2) What are the CE-SDS invalid rates and root causes seen at different companies?

3) What can be done to limit invalid rates?

What should constitute an invalid analysis?

A harmonized definition of invalid rates would provide a powerful tool to understand the deficiencies of an analytical technique. It was clear from the outset companies currently have separate definitions of what constitutes an invalid assay. While all the definitions center on the concept of "assay did not perform as expected", the subtle differences can cause great variations in the invalid rate. For example, a standard practice at many companies is to run a sequence of test articles flanked by reference standard (RS) controls (i.e. first and last injection of the sequence is the RS). If the flanking RS controls do not meet acceptance criteria, the test articles within the sequence are considered invalid, even if those test article runs were "successful". However, it was not clear whether such an event should be considered as one invalid or whether all of the test articles within that sequence would be counted in tallying up the invalid rate. As a way to get around this issue, some participants mentioned the concept of invalid analyses as those that don't pass predefined acceptance criteria. Each analyzed sample would then be considered independently and one invalid analysis would not impact the rest of the samples in the sequence. This approach may face its own set of complications. For example, by invalidating results that don't pass acceptance criteria (e.g. lower purity than expected), one biases the dataset towards only retesting samples that did not provide the expected results. Clearly, there is a need for more discussion on this subject to harmonize the best way for defining an invalid analysis.

The topic was raised again in a roundtable workshop at the recent AT Europa 2016 CASSS meeting in Vienna. The roundtable echoed some of the CE Pharm survey results, in that failure rate/invalid analyses is mainly needed to compare the reliability of analytical systems, which consists of instruments, standard operating procedures (SOPs), reagents, software, (trained) analysts, etc. The roundtable participants favored a simple definition e.g. failed runs as a percentage of the total number of runs. Such a definition would count the number of failed runs, and estimate the number of total runs. In this instance, "failed" would mean the data trace (e.g. electropherogram) clearly did not fit the expected pattern, despite passing system suitability tests (SSTs). The discrepancy should also not be sample related (e.g. not due to a stressed or degraded sample). As mentioned above, this definition may cause one failed run to spoil a whole or parts of a sample series, e.g. if an important SST or RS is missing, or if a series stops and samples are lost due to instability. Future discussions should focus on the challenge in navigating the pros and cons of various invalid analysis definitions.

What are the CE-SDS invalid rates and root causes seen at different companies?

While the definition of invalid assays may vary at different companies, it was widely recognized that this was an important attribute to track. Tracking tends to be more formalized in GMP and QC groups, while research labs typically rely on the subject matter experts to recount their past experience. Figure 1 represents the first attempt to provide an industry wide overview of CE-SDS invalid rates. Despite the lack of a harmonized definition of invalid rates, 13 of the 19 survey respondents showed the CE-SDS method had invalid rates of 7% or less. However, the other 6 participants reported distinctly higher invalid rates. This data highlights the need to discuss and implement best practices for lowering invalid rates, as well as harmonize definitions so the performance of these best practices can be tracked.

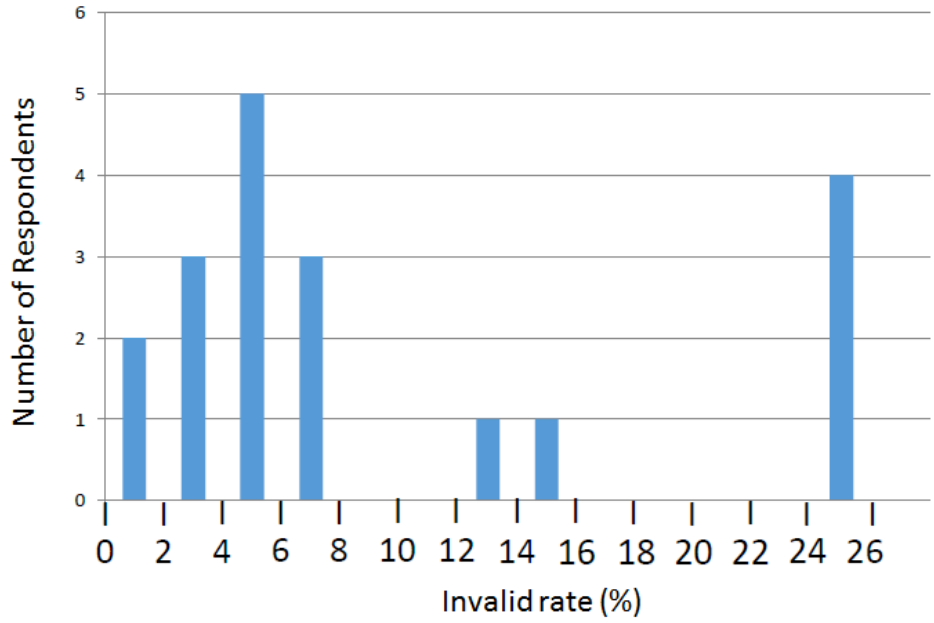

Figure 1: CE-Pharm survey results on the magnitude of CE-SDS invalid rates. Mean = 10.14%, median = 6.75%, standard deviation = 8.99% (n=19). The definition of the invalid rate may be different between different respondents, as discussed in the text.

Regardless of how invalid analyses are defined, they result from some failure in the assay, which could be traced to certain root causes. Survey participants were asked to define these root causes and in this, there were many more commonalities between the different participants. Instrumentation was a common theme and many attendees cited hardware failures, several times attributed to the use of the previous generations of instruments. It was clear vendors had received and responded to this feedback, since an equally common theme among participants was that hardware issues are much less frequent in the current generation of CE instruments.

Likewise, quality control of commercial reagents and consumables was mentioned as a contributing factor. This issue is not unique to CE. For example, batch to batch variability in chromatography columns has sometimes resulted in companies stockpiling column lots that have superior performance. This is not a practical solution for smaller companies, who may not have the resources to either purchase a

stockpile, or to discard the unused stockpile upon the end of shelf life. Recent antibody standards generated by USP and NIST prompted participants to ask vendors to consider incorporating these into their reagent QC testing.

What can be done to limit invalid rates?

While instruments and reagents play a role in the overall invalid rates, they may not be the major factors. Figure 1 indicates a wide range of invalid rates, with a 10-fold difference between the lowest and highest invalid rates. Clearly, there are some unique factors limiting the reliability of CE-SDS in some organizations. Three approaches were discussed as potential solutions to high invalid rates: 1) a detailed training program by in-house experts to reduce invalid analyses due to analyst error; 2) a detailed use and maintenance SOP to reduce invalid analyses due to improper maintenance, or because nuances in the assay method were not successfully captured; 3) work with the vendor in cases where a root cause cannot be easily identified or corrected. Participants also suggested future surveys should try to link invalid rates to root causes, so we can begin to identify the frequency of certain root causes, as well as identify best practices from certain companies in reducing invalid analyses.

Conclusions

The investigation of invalid rates was identified as an important general subject by CASSS in 2015. The focus of the initial study was CE-SDS, but best practices for limiting invalid rates could be transferable to other situations. For instance:

- The need for comprehensive training and SOPs
- Careful consideration of instrumental aspects
- Comparison with techniques which have very low failure rate (such as HPLC), in order to identify potential avenues for improvement.