

Data Integrity and Validation: An Interview with Cerulean's John Avellanet

We may not be hearing much about 21 CFR Part 11 lately, but that doesn't mean it is going away.

Data integrity is at the root of a number of 483s issued this year by FDA. In some cases, firms have not been able to demonstrate their IT department's ability to maintain documentation or to deal effectively with violations, said John Avellanet, Managing Director of the independent consulting firm Cerulean Associates LLC at a July 25 briefing sponsored by FDA Expert Briefings, Washington Information Source and FDA Info.

When considering whether your IT is truly addressing the need for data integrity, ask yourself the same questions you might ask about your driver's license, he advises. Are you really the weight you say you are?

Data management is particularly difficult when clinical trial data are involved, since so much of the documentation is handled manually and there are so many transfer points.

Differentiating Myths from Reality

Myths about Part 11

Irrelevant to compliance

Have to validate every system

FDA is going to get rid of

FDA is not enforcing

One size solutions don't exist

Assess risk to system

Reality about Part 11

Information integrity is crucial

Focus on information lifecycle

Part 11 is *not* the goal

GMP rules cite "validation"

Invest in economies of scale

Assess risk to information



Graphic courtesy of Cerulean Associates LLC. All rights reserved.

PharmaManufacturing.com asked Avellanet about two of the allegations brought up in the [Novartis legal complaint](#). We wondered whether other companies were experiencing any of the issues brought up in that document.

Here are his responses. (Please note that these comments reflect his general experience in the life sciences industry and have absolutely nothing to do with the case, the companies involved or their affiliates.)

PM — What are the essential steps required for validating drug safety data to ensure compliance with 21 CFR Part 11?

JA — Well, the short answer is there are roughly six core steps to ensure Part 11 compliance for clinical data:

1. Determine ahead of time how the data are to be collected and stored at the clinical site, transmitted to the pharma company, and then stored and archived there.
2. Test this process with some test data. Focus on loss of data integrity along the path. For example, if you have four patients, follow the process all the way through to pharma company archival to see if patient three's data is lost or corrupted along the way, that kind of thing (obviously not that simplistic of a problem).
3. Identify the risks at each of these points on the path, then identify potential controls.
4. Craft the procedures around these points and their controls.
5. Verify that all is working as intended.
6. Periodically audit and verify; making sure to report findings - good or bad - and then identify and implement (and monitor) any improvements or new controls.

It's really a pretty straightforward process, isn't it? Of course, human nature being what it is, we just can't seem to leave well enough alone.

PM — What mistakes have you seen that pharma companies typically make in this area?

JA — Here are the top three:

1. Lack of clarity and accountability for data integrity. To date, in every company I've dealt with, the finger points to "someone else" as accountable for data integrity. Records Management says it's an IT issue because IT holds the keys to the computer systems the data is stored on; IT says it's a Records Management issue because IT supports systems, not information; the scientists claim that they are responsible for it and so they want to burn it on CD and stick it in their desk drawer because they just know IT won't be able to find it down the road when the scientist might need it; management assumes that just like quality is accountable for the paper quality systems files (SOPs and the like), they must be accountable for the electronic stuff too, and so on and so on.
2. A division between R&D/preclinical data and the clinical and production data. From the FDA's perspective, this makes no sense (think Quality by Design). From a cynical IT perspective and from a big consulting firm perspective, this makes it easy. Either you deal in the R&D world or you don't. If you don't deal with R&D/preclinical, it's not your problem. One big pharma company is just beginning to grasp this nightmare; they hired a big consulting firm to handle all their data integrity issues world-wide only to find out that the contract specifically excludes anything not already in clinical or in production (the big consulting firm doesn't deal in R&D and preclinical). You can just imagine where this is heading.
3. Inability to translate and achieve mutual understanding between functional units. More and more, I'm convinced this is one of the root causes of the problem. IT can't understand Compliance, who can't understand the scientists, who can't understand Purchasing, who can't understand senior management - you get the picture. It's not that they all need to be in mutual agreement - they won't be - but they do need to figure out how to be in mutual understanding, all marching to the same beat of the drum.

And there are others, of course: egos, competitive pressures, etc. Part of these depend on where within the company you're looking. R&D often makes similar mistakes from company to company, but the same types of mistakes don't occur from R&D to manufacturing within the same company.

PM — What is wrong with "hard coding" randomization codes. What impact would this have on results, and is the practice ever appropriate in some cases?

JA — The short answer? It increases the risk of data manipulation and decreases proof of patient safety and product efficacy. At its core, automated randomization tests multiple different variables (patients, individual results, genetic makeup, etc.) to make comparisons and draw conclusions from. Because you are randomly assessing the whole group, when something you don't expect occurs, then you know that is statistically important.

As a simple example, if you look at all the patients and 15% had heart attacks 30 minutes after taking the drug, but no one else had any heart attacks at all, then - assuming all the patients were healthy - you could say your drug has a 15% chance of causing a heart attack within 30 minutes. You'd then want to do further studies to find out why, what is the risk after 30 minutes, etc.

Now, let's say that I got really good results from an initial clinical trial where I randomized my results, but discovered afterward that if I give my new drug to people who wear glasses, two months after the trial is over, they go blind. What would the data show if I went back to hard-code my randomization codes and limited them to those that excluded people who wore glasses? Results would be great and it would appear that all is good in the world. Even if it's not so cynically after-the-fact, once you hard-code random sampling, it's no longer random, and now a number of things happen that reduce the credibility of the results:

1. You could cherry-pick either test cases, patients or other variables to get the results you want.
2. Even if you don't cherry-pick, you end up with a set of comparisons, the results of which, if they are not the same, can't help you determine if that difference is important or not (was it due to the drug? some other variable, etc.?) because you only tested a set number. With randomization, you draw better correlations, analyses and conclusions.
3. You introduce two human risks: human error (when the guy is originally hard-coding) and then human temptation (such as when a company's internal reviewer is under pressure to make sure that new drug has good results, especially if his job depends on such a thing).
4. You also make it difficult for someone to be able to reproduce your results unless they not only use the same methods, but now the same variables (the same patient genetic make-up, etc.).
5. Finally, if someone does discover the issues, you can always claim an innocent human error.

Additional Resources

- For more on Cerulean Associates LLC and its services, visit <http://www.ceruleanllc.com>