# Why Accuracy-Based Programs are Important to Ensure Quality Laboratory Testing

February 17, 2021

**Julie McDowell:**

The CAP's Accuracy-Based Programs use genuine human specimens that exhibit essentially no matrix effects. Such samples can be used to compare different methods to each other and to reference methods. In doing so, pathologists and laboratory professionals can assess whether the method use in their laboratories is accurate, not just that your performance matches your peer group. In this CAPcast, Dr. Gary Horowitz, chair of the CAP's Accuracy-Based Programs Committee explains why it's important to incorporate accuracy-based testing into your quality process. Dr. Horowitz, how are accuracy-based programs different from conventional proficiency testing surveys?

**Dr. Gary Horowitz:**

Well, let's start with conventional proficiency testing surveys. One important way to ensure your lab is running assays well is to compare your results to other labs. Ideally, proficiency testing providers would send genuine human specimens to labs for comparison, but it's really difficult to find specimens with the range of concentrations needed to assess all the relevant concentrations of all the tests. As a result, PT providers resort to using materials that simulate human specimens to which they add materials to obtain the range of concentrations needed. It turns out that these specimens do not react exactly like genuine human specimens, a phenomenon called matrix effects. In short, two different methods for a given test, say Vitamin D might get exactly the same results on human specimens, but very different results on the proficiency material. For that reason, performance is graded by comparison to labs using the same methods. Peer groups. Conventional surveys can tell you quite well whether you're running your method comparably to other labs running the same method. This is not only required by accreditation agencies. It's also very reassuring to know that you match your peers.

**Julie McDowell:**

Why do we need accuracy based programs at all?

**Dr. Gary Horowitz:**

Well, accuracy based surveys use genuine human specimens that exhibit essentially no matrix effects. The CAP obtains these specimens in several ways. For example, for some tests, testosterone, or cortisol, or A1C, it's relatively easy to get specimens from individuals with a range of concentrations. Men versus women, morning versus evening specimens, individuals whose diabetes is controlled to varying degrees. In other cases like creatinine, we've proved that we can add creatinine to normal serum to get high concentrations without introducing matrix effects. These matrix free specimens are referred to as commutable. Such samples can be used to compare different methods to each other and to reference methods. In so doing, we can assess whether the method used in your laboratory is accurate, not just that your performance matches your peer group.

**Julie McDowell:**

You've touched on this a little, but why else is testing using accuracy based programs so important?

**Dr. Gary Horowitz:**

For some tests, it's important to know not only that you're running them well in comparison to your peers, but also that your results match the reference method. Examples include cholesterol, hemoglobin A1C, testosterone, creatinine, indeed, any method where doctors are making decisions based on cut points established with reference methods. If your results are on the mean for your peer group, it means you're running the method well. But if the method itself has a bias, then doctors may be making mistakes using the data your method provides.

**Julie McDowell:**

Now, does this really happen? Can you provide some examples?

**Dr. Gary Horowitz:**

Yes, it actually does. In general, we should emphasize, things work pretty well, but there are several examples where problems have been detected by the use of accuracy based surveys. These include hemoglobin A1C, creatinine, cystatin C. In these cases, the problem is not with the individual laboratory, but rather with the manufacturer of the assay. And actually, for lab developed tests like most LC-MS assays, the problem would be with the individual laboratory because the laboratories have developed the test and they haven't gone through FDA clearance testing. The good news is that when a problem is detected, the CAP shares that data with the manufacturers and the manufacturers will improve the assays accordingly. Past examples include cholesterol, creatinine, A1C, cystatin C. For example, prior to establishing the accuracy of creatinine methods, some methods had biases that translated into clinically significant differences in estimated GFR calculations. And another example with respect to A1C, over the course of several years, methods whose performance was of dubious quality are no longer on the market, and the number of different methods providing accurate results has increased manyfold. There are many publications describing these accomplishments.

**Julie McDowell:**

Finally, Dr. Horowitz, can you describe harmonization versus accuracy?

**Dr. Gary Horowitz:**

Sure, it's a little difficult to understand the first few times you hear about it, but it's pretty straightforward once you think about it. In the absence of a reference method, it's impossible to know what the true value is, which one needs to establish accuracy. However, what you can determine with commutable matrix free specimens is whether different methods get the same results. If all the results are the same, harmonized, that's a good thing, and once a referred method and reference materials are developed, we can then assess accuracy. If different methods are not harmonized, again, using commutable specimens, we would hope that the reference intervals are accordingly different.

For example, if a commutable specimen is measured as having a TSH of 5.5 by one method, whose reference interval goes up to 5.0, this would be compatible with hypothyroidism. If that same specimen has a TSH of 4.2 by another method, one would hope that the upper limit of that method would be 3.8 or so. If the upper limit of the reference interval was also 5.0, that specimen would appear to represent a normal value. Which result is correct? Is the patient hypothyroid or is the patient euthyroid? In short, conventional proficiency testing surveys provide a great deal of important information, but accuracy-based surveys can, at a minimum, provide insight into whether various test methods are harmonized. That is get comparable results on genuine human specimens, and particularly in the case of tests used with national or international guidelines, whether various test methods are accurate.

**Julie McDowell:**

Thank you, Dr. Horowitz. For more information about the CAP's accuracy-based programs, please visit estore.cap.org and enter accuracy based programs in the search function.

Thank you for listening to this CAPcast. Be sure to listen to our other CAPcasts from the CAP on our SoundCloud channel by downloading the SoundCloud app on your mobile device. And we're also on Apple Podcast and the Stitcher app. To find this podcast, search for the word cap cast on these apps. Once you find our podcast, be sure to click the subscribe button so you don't miss new CAPcast episodes.