

Glucose PT criteria reset stirs standards debate

Anne Paxton

June 2016—It may not be an exact science, but resetting standards is a long-established means of improving quality of testing, and it can also be a way of adapting to improvements in quality that have already been realized. In the case of the CAP's recent tightening of proficiency testing criteria for hospital glucose testing, both purposes are at work. The new criteria reflect the fact that glucose meter performance has improved significantly, CAP Chemistry Resource Committee chair Gary L. Horowitz, MD, explains in the 2016 Program Update on Glucose Meter Performance.

But the change in Survey criteria has brought unexpected pushback from one of the leading hospital glucose meter manufacturers.

The CAP's Chemistry Resource Committee approved and implemented the new PT grading criteria for hospital glucose meter performance in early 2015. The cutoff for passing was changed from within 20 percent (or 12 mg/dL, whichever was greater) of the peer group mean to within 12.5 percent (or 12 mg/dL, whichever was greater) of the peer group mean. The CAP changed the criteria to match the new POCT12-A3 standard set by the Clinical and Laboratory Standards Institute in 2013 for performance of the meters.

As Dr. Horowitz notes in the program update, the reality is that current meters achieve values between 262 and 338 more than 98 percent of the time. With the mean peer group value for a proficiency specimen whole blood glucose being 300 mg/dL, the committee believed it made little sense to consider values as low as 225 and as high as 375 (a range of 150 mg/dL) to be acceptable grades. "If newer meters can, and do, perform much better than their predecessors, shouldn't our grading criteria reflect that reality?" he says.

However, Nova Biomedical believes the tightened PT criteria are not suitable for its Nova StatStrip Glucose Hospital Meter systems. In March 2016, Nova Biomedical addressed an information bulletin to customers, in which the company reported it had received a significant increase in complaints from StatStrip users following the 2015 CAP change in grading criteria for the Survey.

Nova contends that its internal testing has shown that the CLSI POCT12-A3 performance guideline, on which the CAP new Survey criteria are based, is appropriate for whole blood specimens but can have unintended consequences when used with artificial, manufactured PT glucose materials. Limitations of the PT materials, the company says, may cause properly working StatStrip meters to falsely fail or improperly working StatStrip meters to falsely pass, in a small number of cases.

After notifying the CAP of its concerns, Nova announced it was recommending that StatStrip customers who are having problems with proficiency testing consider switching from the CAP Survey to the whole blood glucose survey offered by the American Proficiency Institute. Chemistry Resource Committee members interviewed by CAP TODAY, on the other hand, believe that the impact of the new Survey criteria on pass rates, including the pass rates for Nova meters, is very minor, and they do not think switching proficiency tests is appropriate or advisable.

In both the laboratory and clinical arenas, there is occasional confusion about the differing standards for home meters and hospital meters, says Chemistry Resource Committee member David Sacks, MB, ChB, FRCPath, chief of the clinical chemistry service for the National Institutes of Health Department of Laboratory Medicine, adjunct professor of medicine at Georgetown University, and clinical professor of pathology at George Washington University. Standards for home meters are set by the International Standards Organization, while the CLSI sets standards for hospital meters.

The accepted home and hospital meter standards were essentially the same until 2013 when the CLSI tightened its

standards for hospital glucose instruments to the ± 12.5 percent or ± 12 mg/dL standard. “This was because the technology has improved considerably over the 20 years since the old criteria came out,” Dr. Sacks notes.

WBG CAP Surveys pass rates compared													
2014 Survey results, with “old” grading							2015 Survey results, with “new” grading						
	Specimen 1		Specimen 2		Specimen 3			Specimen 1		Specimen 2		Specimen 3	
	Mean value (mg/dL)	Pass rate	Mean value (mg/dL)	Pass rate	Mean value (mg/dL)	Pass rate		Mean value (mg/dL)	Pass rate	Mean value (mg/dL)	Pass rate	Mean value (mg/dL)	Pass rate
Manufacturer A	402.73	99.8%	156.31	99.7%	68.82	99.4%	Manufacturer A	411.02	99.0%	155.56	98.5%	64.56	99.2%
Manufacturer B	369.35	99.6%	207.11	95.8%	99.61	95.7%	Manufacturer B	379.52	99.7%	205.11	99.6%	94.70	99.5%
Manufacturer C	309.42	99.7%	167.33	99.7%	78.06	99.6%	Manufacturer C	304.86	97.9%	166.08	98.5%	73.55	99.2%
Overall		99.6%		98.9%		98.7%	Overall		98.9%		98.9%		99.2%

Adding to the mix, in January 2014, the Food and Drug Administration proposed guidelines for manufacturers of the meters. “The FDA guideline is not official yet, and we don’t know what its status is or when it will be released. They’ve received a lot of comments and may have modified the guideline but nothing is yet official,” Dr. Sacks says. The FDA draft guidelines for manufacturers have two different sets of analytical goals, one for hospitals and one for home use. “The hospital one is very strict: ± 10 percent or ± 7 mg/dL. So it’s much tighter than the CAP criteria.”



Dr. Sacks

There is no doubt that the FDA guidelines heightened the controversy. In fact, “a lot of people think the FDA draft guidelines are too strict,” Dr. Sacks says. “What they indicate is that 99 percent of values have to be classified within 10 percent, and 100 percent of values have to be within 20 percent of the reference at glucose concentrations ≥ 70 mg/dL. Which is probably not even attainable for most instruments in central labs, let alone the meters, which tend not to be as accurate.”

The FDA’s guidelines apply to manufacturers that are seeking FDA approval for their meters, so the CAP would not have to adjust to them, he notes. “But they almost inevitably would have some influence, because the meters will have to be more accurate, assuming the FDA keeps the criteria as strict as they have proposed. It is possible that the CLSI would convene a committee to reevaluate their guidelines based on the FDA’s recommendations. If CLSI changes its guidelines, then the College may too.”

The CAP’s tightened Survey criteria might seem on their face to represent a significant change, in that they appear to lower the passing score by 32.5 percent, Dr. Horowitz notes in the 2016 program update, but they were adopted only after a long review process, involving review of PT data for several prior years as well as trial grading with the proposed new criteria.

“The committee took two years, and the reason was that the College evaluated the data from several Surveys to determine what the pass rate among users would be with the new criteria as compared with the old criteria,” Dr. Sacks says. “That evaluation showed the difference in failure rate would be very, very small.” (See the comparison, above, of pass rates for the 2014 and 2015 WBG Surveys for the major peer groups, which represent about 90 percent of all participants.) “The College decided that clearly the meters have improved since 20 years ago, and so it was quite acceptable to move to the tighter criteria.”

Some of the technological improvements are proprietary to the manufacturers, Dr. Sacks explains. But he gives examples of others. “With the old meters, the strips were designed in such a way that you would put a little drop of

blood on the pad at the tip, then you'd have to wipe the blood off with a tissue. So sometimes you'd wipe too much blood off and sometimes not enough, and the amount of blood that went into the pad would vary. Also, if it took you a long time to get the strip into the instrument after you put the blood on it, the reaction would start before you even put it in the instrument," which would throw off the reaction timer.

With the newer meters, these problems have been eliminated, he says. "The drop of blood that's needed is much smaller. If there's not enough blood, the instrument won't perform the test, and the reaction doesn't start until the strip is in the instrument." In addition, there are lockouts if there is not enough blood, and much-improved information technology prevents the meter from giving a patient result unless the quality control is run before use.

"These are all just practical things related to how the test is done that have resulted in significant improvements in the accuracy of the meter." In fact, Dr. Sacks notes, "Very few of the CAP Survey failures are due to inaccuracy in measurement. Most of them are just related to actually completing the submission and sending it in, and transposition errors."

Nova Biomedical, Roche Diagnostics, and Abbott are the three largest manufacturers of hospital point-of-care glucose meters. In fact, they are five- to 10-fold larger than the others, so collectively they have the vast majority of the market, Dr. Sacks says. But to date, only Nova's StatStrip has been cleared by the FDA for use in all hospital and professional health care settings, including with critically ill patients.

This distinction originated several years ago when people decided to use hospital glucose meters for critical care patients, mainly in intensive care units, Dr. Sacks says. "The FDA never approved any meter for use in ICUs, so they told the manufacturers if they want to use their meters on critically ill patients, they had to meet quite stringent criteria. Unfortunately, this is an area of a lot of controversy. They never defined what 'critically ill' patients are. But Nova is the only meter that has met the FDA criteria so far."

He has heard anecdotally that Roche and Abbott are in various stages of applying for the same recognition and approval from the FDA, but notes that the FDA requires companies to generate a lot of patient data and submit them, so it's not known when the Roche and Abbott approvals might be granted.

Nova was a relatively small company compared with Roche, Abbott, and Johnson & Johnson, maker of LifeScan, he notes. Then, in 2013, LifeScan decided to withdraw from the hospital market. Around the same time, Roche brought out a new meter that did not yet have FDA approval, Dr. Sacks says.

"So basically there were a lot of hospitals that had been using LifeScan meters, LifeScan was not supporting hospital meters anymore, and a lot of hospitals introduced Nova meters because at that time the new Roche meter was not FDA approved. So Nova got a lot of users from that. And obviously, since the meter was acceptable in the ICU, I presume that motivated some people to get Nova meters too."

In Nova's March 2016 bulletin, the company expressed the belief that the API survey's grading criteria are more appropriate than the CAP's stricter grading criteria for use with artificial matrix materials. But in actuality, none of the manufacturers' meters were designed for use with the artificial matrix materials that the CAP uses for proficiency testing, Dr. Sacks notes. "The biggest deficiency of the College's PT for glucose meters is that the material is artificial. We've spent a lot of time trying to get whole blood. But the problem with whole blood is that glucose is metabolized by blood cells, so the sample is unstable."

"In fact, we've tried several companies in different parts of the world and tested materials from different companies, including one in the Netherlands that uses blood from cows. But the College has been unable to find materials that work in CAP Surveys." This has not shaken laboratories' confidence in the CAP Surveys in any way, he emphasizes. "Anecdotally, I can tell you almost every hospital I know chooses CAP proficiency material for pretty much everything that CAP has available," Dr. Sacks says.

Dr. Sacks does not think Nova's stance on the new CAP Survey criteria reflects actual performance levels of Nova

meters on the Survey and may be an overreaction on the part of Nova. “Obviously some users who pass at 20 percent fail at the new 12.5 percent cutoff. But based on the 2015 grading comparisons, with a breakdown of different companies, I can tell you almost all Nova participants—and there are hundreds of them—they almost all pass.” He has seen a drop in the pass rate generally of about one percent. “But more than 98 percent of the participants pass. The fraction of participants who fail is very, very, very small.”

“Most lab directors and other laboratorians are well aware of what is going on in the field. The Chemistry Resource Committee made an appropriate and right decision for the PT Survey as far as quality considering the actions and decisions of the FDA and the CMS in the last couple of years,” says committee member Stephen E. Kahn, PhD, vice chair of clinical services and professor of pathology at Loyola University Medical Center, Maywood, Ill., and a past president of the AACC. The guideline put out by CLSI—the basis for the new CAP PT criteria—“recommended appropriate analytical goals for expected glucose meter performance in hospitals.”



Dr. Kahn

Commotion and concern has surrounded application of the guidance where meters were being used for critically ill patients, Dr. Kahn agrees. “The draft guidance that came out in early 2014 from the FDA was really directed to the manufacturers, and the requirements were even more stringent than the CLSI guidelines in terms of expected accuracy and precision. Then CMS came out later that year and sort of underscored the need for strict requirements, but a few months later, CMS backed off a little and said further study was called for.”

The CLSI guideline (POCT12-A3) is more stringent than the accuracy and precision requirements that have been followed for several years, he notes. But the CAP—and, Dr. Kahn believes, all the major laboratory organizations—considers POCT12-A3 a reasonable approach.

“The Chemistry Resource Committee is consistently thorough, deliberate, and thoughtful whenever it considers a decision about any aspect of PT whether it’s for a new Survey, new analyte, or new material,” he maintains. “I have not seen anything but a thoughtful, serious consideration of the issues from this committee, and always with an effort to balance what we do with quality in PT testing while also considering practical, real-world issues from the manufacturer’s perspective as well as from the customer perspective and that of participating labs.”

“I was surprised at the strong pushback we heard about and that despite the official response from CAP a few months ago, Nova decided to send that letter out to their customers,” Dr. Kahn says. “The Chemistry Resource Committee strives to use commutable materials whenever it is feasible. There are certain Surveys like WB glucose where we have yet to identify commutable materials although grading is still by peer group.” He is not aware of any reason why Nova performance should be affected in a different way than other manufacturers’ performance.

It’s possible to have an approach to proficiency testing that passes virtually everybody, Dr. Kahn points out. “But that’s never the purpose of PT. Unless a small percentage, or a very small percentage at least, doesn’t pass each and every time, then what good are the PT criteria? You’re doing nothing to focus on comparably poor performing methods. Plus you’re doing nothing to focus on potential labs or institution sites where there might be some type of quality problem, whether it’s analytical performance or operator based, or whatever the case may be.”

Dr. Kahn thought the CAP PT program to be strong and robust long before he joined the Chemistry Resource Committee, and his experience on the committee has only reinforced his impression of how meticulous the CAP’s approach to offering proficiency testing is. Regarding Nova’s information bulletin about the PT criteria, Dr. Kahn

says, "Should we [Loyola] use Nova in the future, we will continue to look very closely at our PT performance. But we have been participating in the CAP Whole Blood Glucose Survey for several years and we intend to continue for the foreseeable future. We don't see any reason to change."

As the CAP Survey goes forward, he believes, it will continue to keep its standards relevant and useful in maintaining the high quality of the meters' performance. "The most important consideration, first and foremost, is to try to do whatever you can to optimize patient safety and try to minimize the likelihood of medical error, particularly with such a commonly performed test as glucose meter testing."

[hr]

Anne Paxton is a writer and attorney in Seattle.