A laboratory on the trail of troubling TSH results

Anne Ford

September 2014—It would be a nightmare for any laboratory professional: a misdiagnosed and mistreated patient owing to an aberrant test result.

Julia C. Drees, PhD, a scientific director for chemistry at TPMG Regional Reference Laboratory, Kaiser Permanente Northern California, found herself facing that situation two years ago. She and colleague Judy Stone, PhD, then a Kaiser scientific director who is now at UCSD, discovered that faulty TSH results from their laboratory had led to multiple patients being misdiagnosed, and some even treated inappropriately. And yes—as she told the audience in a talk at the American Association for Clinical Chemistry annual meeting in July—the clinical effects were significant.

For example, "one patient was a 19-year-old young woman who was a hypochondriac and suffered from depression, and the last thing she needs is for the doctor to say, 'Oh, there's something wrong with your thyroid,' give her medication and make her even sicker, even more tired, even less likely to get out of bed in the morning," Dr. Drees said in her talk, titled "Falsely Undetectable TSH in a Cohort of Euthyroid Patients." "It was really unfortunate."

Even more unfortunate, Dr. Drees knew, was that if she and her colleagues didn't get to the bottom of the TSH mystery, still more patients could be affected. "We were really baffled about what might be going on," she said. "We wanted to know: Are there other patients at risk, and can the lab help in this situation?"

Simple questions, but answering them required a months-long investigation. In the end, that intense work paid off, not only because Dr. Drees and her team were able to solve the mystery of the faulty results, but also because the strategies they employed along the way expanded their ability to reduce patient risk from aberrant results in general. They hope their investigation, the results of which were published in April, will benefit other institutions whose patients might be at risk (Drees JC, et al. *J Clin Endocrinol Metab*. 2014;99[4]: 1171–1179).

The quest began in 2012, when an endocrinologist colleague of Dr. Drees contacted the laboratory to report something troubling. One of her patients—a woman later determined to be clinically euthyroid— had had her TSH levels tested with the laboratory's normal screening assay, Siemens' Advia Centaur TSH-3 Ultra (also known as the TSH3-UL), five times over a nine-month period. The patient's TSH was consistently undetectable despite treatment with methimazole, an antithyroid drug. But when the endocrinologist ordered an alternative test, the Abbott Architect TSH assay, the result was >75 µIU/mL. Hindsight revealed that the methimazole had made the patient transiently hypothyroid. The drug was discontinued and the patient's hypothyroidism resolved quickly.

Dr. Drees

When confronted with discrepant results from an immunoassay, Dr. Drees explained, "We always do a normal interference study, which includes dilutions, and in this case we did not see a suddenly elevated result, and the heterophile antibody test was negative. While we're working on these interference studies, the endocrinologists talk to each other, and they identify three more discordant patients, as we called them, based on their discordant

TSH results, where one assay is undetectable and the other is normal or elevated."

Dr. Drees and Dr. Stone compared the timelines of each discordant patient's results and discovered a striking commonality. In her words: "[One patient had] a normal TSH screen before we implemented the TSH3-UL assay, and then all of a sudden this patient has undetectable TSH results. The next patient has lots of normal screens in the past; we implement the TSH3-UL assay, and now [that patient shows] undetectable TSH. This next patient has a history of hypothyroidism, so they're actually high, but under control. We implement, and then [the TSH is] undetectable. Another hypothyroid patient, under control, [has] easily detectable TSH, which is then suddenly undetectable after the assay's implementation. So this really helped us figure out that it was our TSH assay" that was at fault.

Unsure if their findings were limited to just these few patients, Dr. Drees and her team decided to perform a 30-day trial of reflexing all samples that matched the pattern seen in those four discordant patients. "They had undetectable TSH3-UL, and they had a free T4 that was not even close to the upper range, just sort of medium normal. These are the ones we chose to reflex from the TSH3-UL to the Architect TSH," she explained. "And so we're going along, and we get up to 108 in these 30 days, and the undetectable TSH3-UL is matching the Architect. They're all extremely low or undetectable. No problem. It's a fluke. We're fine.

"And then," she added, "we find one more [discordant] patient." In other words, other patients remained at risk of being misdiagnosed, inappropriately treated, or both. The nightmare wasn't over just yet. In fact, the laboratory ultimately identified 23 discordant patients (16 patients had false results reported; nine of those were treated based on the false results). "While this is going on, we're scrambling and baffled and also studiously, thoughtfully pursuing things," Dr. Drees said.

She sent samples to Siemens, which tested them on its Dimension Vista TSH and Dimension TSH platforms. She ran samples in her own laboratory using two other Siemens assays (the second-generation Advia Centaur TSH, referred to as the TSH2, and the Immulite 2000 TSH). She sent samples to ARUP Laboratories, which tested them with the Roche Diagnostics Modular E170 TSH assay, and she sent samples to a friend at the San Francisco VA Medical Center, who tested them with Beckman Coulter's DxI600 Fast hTSH assay.

The Roche and Beckman Coulter assays all detected TSH in the samples Dr. Drees supplied, as did Siemens' TSH2. The other Siemens assays did not. "This is pretty clear: that we have some Siemens assays that are not detecting it," said Dr. Drees. "What's interesting about these assays is that they were all developed by different companies that Siemens subsequently acquired….There are lots of other studies that identify patients with low TSH concentrations, TSH gene mutations, and thyroid dysfunction, but we believe this to be the first report of biologically active TSH that is undetected by widely used, FDA-approved TSH assays."

Something else was becoming clear: The problem was very likely related to the patients' ethnicity. "Other than two patients who identified as Persian and Middle Eastern, they all considered themselves South Asian," Dr. Drees said. "Our hypothesis was that the discordant patients share a TSH variant that is not recognized by some antibodies in these assays. If you have a monoclonal antibody, it only recognizes one specific epitope on your antigen; monoclonal antibodies are exquisitely selective in that way. If you have monoclonal antibody A that only recognizes [a particular] epitope, and that epitope is slightly altered, that antibody is not going to recognize it, whereas another assay might use monoclonal antibody C, and it has no problem." Was that what was going on here, she wondered?

A helpful clue came courtesy of researchers at Siemens. "They did a clever thing: They mixed and matched antibodies and made experimental hybrid assays. And these experimental hybrid assays did a pretty good job of detecting TSH in these controls, and then when we gave them a couple of discordant patient samples, as we expected, the TSH2 detected it just fine, whereas the TSH3-UL could not detect it." What does this have in common with the other assay that can't detect it? "It has in common the TSH3-UL solid-phase antibody. So they concluded that the TSH solid-phase antibody is failing to detect TSH in the discordant patient serum. And they demonstrated that the monoclonal antibodies in the Immulite assay shared an epitope with the solid-phase

antibody from the TSH3-UL, as did the detection monoclonal antibodies used in the two Dimension assays. So if they share an epitope, that also supports our hypothesis that there's a common mutation altering the epitope, and that these hyperselective monoclonal antibodies can no longer recognize it."

DNA sequencing was the obvious next step. Dr. Drees had the serum from remnant samples from 12 discordant patients and four control patients sequenced for the alpha and beta subunits of TSH. Indeed, all of the discordant patients were found to share a point mutation in an alanine to guanine in the TSH beta subunit. "This mutation changes the codon into a new amino acid," she explained. "Whereas it was once arginine, it is now glycine, and that's clearly a large change from a long, positively charged amino acid that was probably surface exposed on the molecule, since it wants to be in contact with the aqueous, as opposed to the hydrophobic, interior, and changing to an amino acid with no side chain whatsoever. So that appears to be significant. Interestingly, there was no mutation in the TSH alpha subunit, which you would expect."

Siemens worked closely with Dr. Drees and her colleagues in investigating this rare TSH variant and notified its customers and appropriate regulatory authorities of the discovery. Siemens tells CAP TODAY: "The falsely low TSH values were observed in a small group of patients of South Asian descent over a 30-month period at a rate of occurrence of $0.6 \times 10\text{-}7$. As with any immuno-recognition measurement of a peptide, extremely rare genetic variants may exhibit varying degrees of detection. It is recommended that abnormal or clinically discordant TSH values are co-interpreted in conjunction with thyroxine and T3, patient history, and clinical signs and symptoms."

At last, Dr. Drees and her colleagues knew exactly why their discordant patients had such misleading results. They didn't know, though, exactly how many patients had been affected and to what degree. Given that Kaiser Permanente Northern California Regional Laboratory serves 3 million patients from 21 medical centers and daily performs about 4,000 TSHs, carrying out a retrospective review would prove to be as complicated as identifying the source of the error in the first place.
Fortunately, the period for the review was crystal clear. The laboratory had implemented the TSH3-UL assay on March 15, 2010; reflexing to the Architect began July 1, 2012. ("Proactively, we continued this reflex of all very low TSH3-ULs to the Architect," Dr. Drees explained. "Our TSH3-UL is on a nice automation line, and it's still a great assay for the vast, vast, vast majority of people, so there was no reason for us to take it offline. And this [reflexing] would help us prevent future false results and prospectively identify patients so they didn't fall into this trap.")

During that period, the laboratory reported 3 million TSH results for 1.6 million unique patients. "Impossibly large numbers," Dr. Drees called them. "But we did have two things on our side. The first is a very extensive electronic medical records system with the all-important lab results, but also pharmacy data on all of the drugs they've been prescribed, the diagnosis codes for every encounter, and, of course, demographics. We also had two very talented data analysts."

The winnowing began with a decision to look only at the patients who had had two very low TSH3-UL results during the relevant period, given that "if they only had one, the doctor would follow up, and we'd find them," Dr. Drees said. Their starting number of suspected discordant patients was almost 5,000. "Anybody who ever had anything more than .05, we said, 'OK, that doesn't fit the pattern,' so we're down to 2,000 here. If they ever had an elevated T3, T4, it's more likely to be legitimate hyperthyroidism than this crazy thing we were seeing." These criteria brought the suspected number of patients down to around 600. Dr. Drees consulted Kaiser Permanente endocrinologist Rick Dlott, MD, who advised on additional criteria. "He told us, 'You can look for pan-hypopituitarism and thyroid cancer. Those are two diagnoses that would legitimately give you low TSH and low or normal thyroid hormones.' And so we're down to 500. And then we said, 'Well, if they were prescribed methimazole before we went live with this assay, they were probably diagnosed with hyperthyroidism based on a good TSH assay, or one that was unaffected by this issue, at least.' So we're down to 478 discordant patients."

But the narrowing-down process wasn't over yet. "We had weekly reports. So now we're taking this report of suspected patients, and every week, we would look and say, 'OK, do they have a TSH3-UL that has jumped up now?' And we cut it out if they did. Did they still have a low one, but it matched the Architect? Good. And then did

they have kind of borderline TSH3-UL but a free T4 that indicated it could be hyperthyroidism? 'OK, we'll cut you out, too.'"

During this process, she recalled, colleagues experienced with previous recalls would ask, "'When will the numbers be reasonable so we can just sit down and do some chart review? We'll all get in a room, we'll have our 10 computers, and we'll all just sort of sort through it and figure it out.' Well, what we realized was that we had just done chart review. There wasn't that much more we could look for, except reading the interpretive comments from some thyroid scans. We had essentially done chart review by looking for the diagnosis codes, looking for the various lab results, looking for the medications."

That said, some patients did merit a traditional chart review—the patients who had been given new prescriptions for methimazole based on their TSH3-UL results. "They were being diagnosed for hyperthyroidism based on this assay that we know has this problem, so we figured they were the highest risk," Dr. Drees said. "And sure enough, when we look at these 18 patients, two of them really matched the pattern, and when we contacted the physician and the patient came back in, we checked them on both assays, and they were two discordant patients." Treatment was always discontinued immediately when the results were communicated to the physicians, Dr. Drees said. Ten of the 18 high-risk patients were ruled out, "because they had very obvious hot nodules on thyroid scans."

In the end, the patients who needed to be called in for a new TSH test numbered just 262. "So we sent them a letter saying there was a lab error, which killed me, [because] it wasn't our fault," Dr. Drees said. Those patients' primary care physicians were also notified. "We did not want to wait for these patients to just wander back through our doors."

In addition, "we put in bulk TSH orders on those patients," she said. "Maybe the patient didn't get the letter. Maybe the physician forgot to follow up. But the next time this patient comes in for a cholesterol, the TSH order is going to be there. We're going to get it." At this point, fewer than 100 of those patients are outstanding. Dr. Drees tells CAP TODAY: "I expect to continue to find the occasional new discordant patient, but I am confident that the systems our lab has in place will catch them before any misdiagnosis or treatment can occur." In fact, thanks to the reflex and review systems, no falsely undetectable TSH results have been reported since July 2012, and none of the discordant patients identified since the end of 2012 were treated.

The moral of the story? "No lab test is infallible"—not even a gold-standard test such as TSH, Dr. Drees said. Certainly that point has been driven home at the Kaiser regional laboratory. But what Dr. Drees and her team also have gained is a new appreciation of their own abilities to reduce the risk to patients from erroneous results.

"Because we are reflexing all of these undetectable TSHs, we're starting to find a lot of different limitations that are sort of inherent in immunoassays," she said. "I'd say that immunoassays are the most likely test to give aberrant patient results." For example, "I found one very interesting case of macro-TSH, which is very rare. We found three cases of fluorescein interference, where if a person gets injected with fluorescein, because they're having an angiogram of their eye, and then 20 minutes later they wander down to the lab to get their cholesterol and their TSH tested, they get a falsely undetectable TSH3-UL result." The fluorescein interference had never been reported and led Siemens to notify its customers of this rare interference, Dr. Drees said.

Then, too, the chemistry team has created an internal email list to help triage potentially troubling results. "We are lucky to have seven scientific directors in chemistry. If one of us is out, somebody else will be able to answer it, and we sort of take turns answering these questions," from medical technologists and others, she explained. "We have an algorithm for the TSH, but invariably there's something that the technologists still need a little help on.

"And then we get questions from physicians," she continued. "They'll request additional specialized testing, including investigating hook effects or heterophile antibody interferences. And then primary care physicians are being asked to monitor their pain patients for their opiates, and this is not an easy thing for them to interpret. So with every opiate confirmation result, if they click on 'see comments,' they get a table showing the drugs and

metabolites you would expect to see. ‘If they don’t match this, or if you have further questions, contact our email distribution list.’ And so we get a few emails a week on that.”

They are leaving nothing to chance. “And we have all these tools to be able to do that,” she says, “which feels really good.”
[hr]

Anne Ford is a writer in Evanston, Ill.