Clinical Chemistry

Incorporating Clinical Considerations into Statistical Analyses of Markers: A Quiet Revolution in How We Think About Data

Article: Andrew J. Vickers.
Incorporating Clinical Considerations into Statistical Analyses of Markers: A Quiet Revolution in How We Think About Data.
Clin Chem 2016;62:671-672.
http://www.clinchem.org/content/62/5/671.extract

Guest: Dr. Andrew Vickers is an Attending Research Methodologist at Memorial Sloan Kettering Cancer Center in New York City.

Bob Barrett: This is a podcast from *Clinical Chemistry*, sponsored by the Department of Laboratory Medicine at Boston Children's Hospital. I'm Bob Barrett.

Biomarkers are used frequently in medical decision-making to determine the presence or absence of disease and individual's risk or prognosis or the response to therapy. Therefore, identifying new biomarkers and finding new applications for existing markers are important components of translational research. However, progress in this area is often limited by challenges in defining the sensitivity and specificity characteristics a biomarker needs to be clinically useful.

The May 2016 issue of *Clinical Chemistry* published a paper describing an innovative way to calculate such target specifications for biomarker performance. This method then enables investigators to evaluate the potential clinical impact of biomarkers in early phase studies. An editorial by Dr. Andrew Vickers accompanied that article and highlighted the need for incorporating clinical considerations into biostatistical approaches.

Dr. Vickers joins us in this podcast. He's an Attending Research Methodologist at Memorial Sloan Kettering Cancer Center, where he focuses on developing novel methods for assessing the clinical value of predictive tools.

So Dr. Vickers, you've suggested that there has been a shift in thinking as to how we evaluate tests. What did we used to do and why aren't we doing that anymore?

Dr. Andrew Vickers: The most traditional way of thinking about a diagnostic test was sensitivity and specificity, the most widely used concepts. I remember I learned them in biostatistics. I remember actually one of my first reasons to be suspicious of them was reading a paper by a doctor who reported a test and said, "Its sensitivity was this and specificity was that. But I can't work out what that means and therefore I'll have to do it in a different way."

Sensitivity and specificity really have very odd definitions. Sensitivity is the probability that a patient who has disease will have a positive test. Of course that's not what a doctor in clinical practice wants to know. It's not like, "Oh, so I've got a sick patient here. I'm going to try and guess that his test comes out positive." It's really a very peculiar concept. So that's one of the reasons why people turned away from it. It's because it didn't really reflect clinical practice.

The other one is -- because it doesn't actually tell you what to do. I'm a doctor. I'm reading some paper about some new test and I find out that the sensitivity is 78% and the specificity is 60% and then there's some other test where the sensitivity is 82% and the specificity is 50%. So, which of these two tests is better, if any, or should I be using either of these two tests? Is that sensitivity and specificity high enough?

I can tell you, I've actually posed this question numerous times to groups of statisticians. Okay, here are statisticians, the people who are telling us to measure sensitivity and specificity and I'll say, "Okay, here's some tests. Here's the sensitivity and specificity, which of these tests should we use?" And they can't answer the question.

So that's ultimately why we've moved away from some of these traditional biostatistical metrics, because they can't answer the question of whether use of the diagnostic test would do more good than harm.

Bob Barrett: Okay. So we're moving away from the old way, what is the new way of thinking about evaluating tests?

Dr. Andrew Vickers: The new way we think about tests really falls under rubric of decision analysis, and what decision analysis is about just saying, "What would happen if we took one decision, what are the consequences? What would happen if we took a different decision, what are the consequences of that?" Let's compare consequences and find out what would be the best thing to do.

This really started in business. If I invest in this strategy or I invest in that company, what are my likely returns in terms of profits and loss? We can think about diagnostic test in a similar way, if I give a patient a diagnostic test, the assumption is that if they're positive to the test, I'll give them a drug or give them a further workup in some way. If they're negative, I won't. So, we can actually think through the consequences.

I'll give you a very specific example of a particular problem we have in medicine right now, which is so many men get PSAs, prostate-specific antigen as a test for prostate cancer.

What we do know is that most men who have an elevated PSA don't have prostate cancer, but we biopsy them anyway. So there have been suggestions that we do an additional diagnostic test to determine which men with elevated PSA do we do the biopsy on.

The traditional way of doing it would say, "Well, here's a bunch of guys with the biopsy on all of them and we did a test first and here's its sensitivity and specificity." A different way of thinking about it would be saying, "If we this test, what would have happened? Here's the number of men we would have biopsied, here's the number of cancers we would have found. Here's the number of men wouldn't have biopsied, and here's the number of cancers we would have missed."

Often, you can just look at that and say, "Hey! That's really great! For a thousand men, we would reduce the number of biopsies by 500 and we only would have missed two cancers. There's no way we would biopsy 500 men to find those two cancers, so, yeah, we're better off using the test."

So that's a very simple step one. Step two is actually to sort of put a value on it to say, "Okay, how much is avoiding a biopsy worth and how much is finding a cancer worth?" Then we can actually come up with one number, a net benefit, or profit if you like, and say, "What's our profit if we use the diagnostic test and what's the profit if we don't use the diagnostic test?" Then you just use the strategy with the highest net benefit.

Bob Barrett: So, does this have any implications about the way test results are reported back to the doctor?

Dr. Andrew Vickers: One of the problems that we've had in medicine for a long time is that we report tests back as positive or negative. That goes across medicine. So, we say somebody with a blood pressure of 145 has hypertension and somebody with a blood pressure of 135 doesn't. We say someone with a BMI of 30.01 is obese and somebody with a BMI of 29.95 is not obese.

Now, obviously, there is not much difference between a BMI of 29.95 and one of 30.01. The higher your BMI, the higher the risk that something bad is going to happen to you, like a heart attack. What we've done for many years is just dichotomize. Now, in terms of chemical chemistry, the classic report we get back is a list of blood markers, your hemoglobin, your neutrophil count, and so on. Then each of those is marked as being in range or out of range, so it's just a yes or a no, that you're -- and that thinking of about a test in terms of sensitivity and specificity.

A more contemporary way of thinking about it is saying, "Well, what we really want to know is the patient's risk of something bad happening." And to do that, to get that risk, we can't just look at individual markers and say, "Are they in range and out of range? Are they above a cut-point or below a cut-point?" We can actually combine information from numerous different markers. Plus also, information about, for example, the patient's age or the past medical history, and put it into risk prediction model.

Then, we can say -- if the patient is at high risk, then we want to think about doing something, and if they're at low risk, then we can reassure them and they may not need intervention. We're actually starting to see this more and more in laboratory tests. There are several laboratory tests that when you send away a blood sample, what the company sends back is actually the patient's risk of prostate cancer, or the patient's risk of death in prostate cancer, is a certain percentage. Now discuss that with the patient and decide what you want to do, rather than saying a level of PSA or free PSA, or if this marker is out of range or in range.

Bob Barrett: Finally, you've talked a lot about predicted probabilities, are you suggesting that doctors just give probabilities to a patient and let them make the decision?

Dr. Andrew Vickers: Yes. That's kind of what I implied in my previous answer. That's a very common critique of the new thoughts about how we do clinical chemistry and how we evaluate markers. That critique is -- I had someone in my office the other day, he has an eighth grade education and he just says, "What do you think I should do doctor?" And me, telling him he has a 9.6% chance of high-grade cancer is not going to be useful.

But that's really a straw man argument. I don't think anyone who's interested in the new statistical approaches to markers is saying that every single patient needs to be given a risk and then they make the decision just as -- it's not like every restaurant, go and get some menu and then orders their dish. What we do think though is that the doctor, once they've given the risk, at least has that option, and there are various ways in which a doctor can have a clinical consultation with the patient that involves a quantity of risk estimate.

For example, the doctor could see that the risk is high and say, "Look, the test results have come back, they're not good. I think we need to do a biopsy." Or could see they're low and say the opposite.

If they're somewhere in the middle, the doctor can evaluate the patient: what sort of patient is it? Is this the patient who likes to be making decisions and be in control? Is this a patient who understands probabilities and math and wants to hear that kind of thing? And then we'll have a conversation. Or is this someone who may not want that conversation and therefore I need to approach in a different way.

So, you might well imagine the same doctor telling one patient, "Look, your risk is 8%, what do you want to do about it?" Another doctor saying, "Well, your risk is somewhere in the middle so let's talk about it. Tell me how you feel about going through this procedure. How do you feel about the risk? How do you feel about the benefit? Then I will interpret what you're saying with reference to this risk that I've got back from the lab."

So it's not a dogmatic approach saying that, "Yeah, get the predicted probability, give that to the patient and have them make a decision." What it is, is saying that, "Having predicted probabilities can help the doctor make better decisions and in a sophisticated way with their patients." This is actually proven. There are several studies now showing that if you give doctors predicted probabilities, that they do act on them and they do actually change their behavior, and that they do that in a rational way. For example, they are more likely to do something with the patient who is at high risk and less likely to do something with the patient who is at low risk.

Bob Barrett: Dr. Andrew Vickers is an Attending Research Methodologist at Memorial Sloan Kettering Cancer Center in New York City. He's been our guest in this podcast from *Clinical Chemistry*. I'm Bob Barrett. Thanks for listening!