



Article:

Ryleigh A. Moore, Joseph W. Rudolf, and Robert L. Schmidt.
Risk Analysis for Quality Control Part 2: Theoretical Foundations for Risk Analysis.
J Appl Lab Med 2023;8(1): 23-33. <https://doi.org/10.1093/jalm/jfac106>

Guest: Dr. Robert Schmidt is Medical Director of Health Systems and Head of Population Informatics at LabCorp.

Randye Kaye:

Hello and welcome to this edition of *JALM* Talk from *The Journal of Applied Laboratory Medicine*, a publication of the American Association for Clinical Chemistry. I'm your host, Randye Kaye. Laboratories devote considerable resources to performing quality control, or QC, to ensure that patient results are reliable. When designing QC procedures, laboratories have to balance costs and labor efforts with sensitivity for detecting errors. For example, narrow QC limits increase the ability to detect errors, but may lead to false positive alerts that increase labor and result turnaround times. On the other hand, wide QC limits increase the likelihood of missing errors.

The January 2023 issue of *JALM*, which is a special issue on Data Science and the Clinical Laboratory, includes a three-part article series on Risk Analysis for Quality Control. In these articles, the authors evaluate the performance of a commonly used risk analysis model, the Parvin model. They also developed their own alternative model that determines the cost to prevent an unacceptable reported result. Then they use data from two analytes to apply their model in practice.

Today, we're joined by the corresponding author for the article series, Dr. Robert Schmidt. Dr. Schmidt is a clinical pathologist who specializes in informatics and analytics. He has an extensive background in quality control, including a PhD in operations management. He's currently Medical Director of Health Systems and Head of Population Informatics at Labcorp. Dr. Schmidt, welcome. These papers describe methods for risk-based analysis. What is risk-based analysis and why is it important for clinical laboratories?

Robert Schmidt:

QC is designed to protect against failures and there's really two kinds of failures. We can have a false negative and that occurs when there's an error in our results that we don't detect or the other one is a false positive, when we can say there's an error when it's not there. We have risks for both of those. QC is designed to protect against these and it's interesting that research has found, there was a paper by Rosenbaum and Westgard some years ago, they found that most labs use a uniform QC policy. They use two standard

deviation limits and some labs use Westgard rules, but basically, it's a one-size-fits-all approach.

But when you think about it, the risk that we are faced with, with a false negative really depends on the assay. Would I rather have a false negative for troponin or a false negative for creatinine? I'm not sure I know the answer to that, but I'm pretty sure they're different. That makes you think that QC policies should be adjusted to reflect the underlying risk of the assay and that's what risk-based QC is all about.

Randye Kaye: Thank you. Let's just talk about the current approaches to risk-based analysis. What are the limitations and what's novel about what's being proposed in your articles?

Robert Schmidt: The one I'm most familiar with is the Parvin model and that's what underlies Bio-Rad's Mission: Control, if I'm correct about that. Basically, what Parvin did was let's just suppose that if an error occurs, it could be an error of any size. We don't know what it is. There's equal probability of any kind of error and then using statistical analysis said, how long would it take to detect one of these errors if they occurred?

For example, if you have a small error, those are hard to detect. You might produce many batches before you actually see it. Whereas if there's a large error, those are pretty easy to detect, so you might detect those right away or after say, one batch. The risk is, during these periods of failure to detect, that you're sending out batches with errors in them and you don't recognize it.

What Parvin did was figure out the probability of sending out what he calls unacceptable final results, results that are bad enough that they would exceed a total allowable error that you wouldn't detect it and it would go out. What we did that's different is that we tied this false negative risk with a false positive risk. The way we did that was looking at what's the probability that a shift occurs in the first place? Parvin didn't include that in his model and that's very fundamental to this because there's more risk the more often this happens. If the probability of something going haywire is high, there's more risk. If it's low, there's less risk.

We don't know that probability but it's intrinsic to what we do in thinking about QC. It's important to include that in a model. That's what's really different about what we did. And so, in doing that, we're able to connect the trade-off between false positive risk and false negative risk. I think that was missing from previous approaches.

Randye Kaye: OK, I see. Thank you. Let's talk about the impact of this approach. How does it advance the field?

Robert Schmidt: Well, I think it allows us to set optimal QC settings, at least in theory. I think the thing is, I think some people look at this model we had and then go, "Well, can you estimate these things?" Maybe, maybe not. I think we've shown in some papers that you can estimate, for example, the probability of a shift happening. But the thing is, is when you make any decision about QC, you're implicitly making some sort of judgment about these risks.

And so, our model makes this very explicit. Even if you don't know the answer to say some of the inputs that go in, you can do a sensitivity analysis, vary them and go, "What might happen if I were wrong? What would happen if this happened? What would happen if that happened?" We are working on developing a very simple computer interface that would enable people to experiment and look at the consequences of various kinds of choices here.

For example, one of the inputs to our model is the distribution of errors that happened. If something goes wrong, what happens? As I mentioned, Parvin made the assumption that any kind of error is equally likely. I don't know if that's a good assumption or not. I think most people I talk to would say that small errors are probably more likely than large errors. Can we estimate that? Can we, based from some sort of data, determine whether Parvin's assumption was right or this other one is right? Maybe, but it turns out also that the model is pretty insensitive to that, so it really doesn't matter very much whether you get that particular input right.

We're experimenting with those kinds of things right now. I'm thinking we'll have a fairly simple interface that people can use and apply this. The underlying mathematics is fairly complex and interestingly, it's the same math that underlies Google and we all use Google every day without even seeing a bit of the mathematics and that's what we're envisioning here. It's just some sort of simple interface that people can look at: What do I think. Is this a stable assay or not? Does it get upset frequently or does it take months before we need an adjustment? What are the relative costs of a false negative? Is it more like troponin? Is it more like creatinine?

You don't need exact answers to these things, but I think we have a simple interface. People can experiment with things and come to probably different decisions about setting QC than they would if they applied a one-size-fits-all approach.

Randy Kaye: It's definitely all about the user interface, so to speak, right?

Robert Schmidt: Yeah, I think so. To make this practical for use, we'll have to have something like that. The mathematics for this I think are fairly daunting to most people, but once again, like Google, there's a lot of mathematics that underlies it but we

don't see that. Same thing here that, we're envisioning an interface that will shield people from the mathematics but the results will be very simple, much like Google pops out a suggestion, this will pop out a suggestion, and I think it can become practical.

Randy Kaye: Sounds very promising. What's the current status of this model? Is it ready for use in the real world?

Robert Schmidt: Not yet. But we're making good headway. I think we found out that we can estimate some of the inputs fairly easily, which we didn't know when we started this. Some of the inputs don't matter very much. It turns out the model is pretty insensitive to some of them. So we're sort of working on those kinds of things at the moment. We're also working with a computer programmer to develop a web interface where people will be able to actually use this and get suggestions about the QC settings.

Randy Kaye: All right. Thank you so much. Thank you for joining me today, Dr. Schmidt.

Robert Schmidt: You're welcome.

Randy Kaye: That was Dr. Robert Schmidt describing the three-part article series, "Risk Analysis for Quality Control" from the January 2023 special issue of *JALM* on Data Science and the Clinical Laboratory. Thanks for tuning into this episode of *JALM* Talk. See you next time and don't forget to submit something for us to talk about.