



Article: David Ng, Frank A. Polito, and Mark A. Cervinski
*Optimization of a Moving Averages Program Using a Simulated Annealing
Algorithm: The Goal is to Monitor the Process Not the Patients.*
Clin Chem 2016;62:1361-1371.
<http://clinchem.aaccjnl.org/content/62/10/1361>

Guest: Dr. Mark Cervinski is Director of Clinical Chemistry and Point-Of-Care Testing at Dartmouth-Hitchcock Medical Center and Assistant Professor of Pathology at the Geisel School of Medicine at Dartmouth.

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.

Ensuring test result accuracy and repeatability are fundamental components of clinical laboratory operations. One way to do this is through periodic assessment of known quality control samples. However, some laboratories also determine the moving average of patient results to warn of potential problems. This system compliments traditional quality control procedures and correct parameter selection is key to its effectiveness.

The October 2016 issue of *Clinical Chemistry* published an article describing a novel way to optimize and evaluate moving averages algorithms in a clinical laboratory. The article's senior author, Dr. Mark Cervinski, is Director of Clinical Chemistry and Point-of-Care Testing at Dartmouth-Hitchcock Medical Center, and Assistant Professor of Pathology at the Dartmouth Geisel School of Medicine. Dr. Cervinski joins us for this podcast.

So doctor, can you describe for us the concept of moving averages?

Dr. Mark Cervinski: Sure. Moving averages is in short a method to use the average or mean value for a particular analyte or test to monitor assay performance in real time. With moving averages, the mean value for a set number of patient results is recalculated with each new patient result. When first starting a moving averages protocol, there's a lag period equal to the number of patient samples you wish to average. But once that initial number of samples has been analyzed, the mean is recalculated with each additional patient sample tested.

In essence, you can think of moving averages as a window of a fixed width. Every result within that window is average to generate a data point. With each new patient sample, the window moves down by one sample, and the average for the values within that window is recalculated.

Moving averages is a continuous quality assurance tool that complements traditional quality control procedures.

Bob Barrett: When setting up moving averages, what does a laboratory need to do?

Dr. Mark Cervinski: Well, setting up moving averages does have some challenges to it. At a minimum, the user needs to select the number of patient samples to average, how big of an error or shift you wish to detect, and the biggest challenge is determining the concentration at which to truncate outliers out of the mean calculation.

This last step is the most challenging for analytes that are not normally distributed or roughly normally distributed. We found that in the absence of computerized optimization, that placing the truncation limits at four times the standard deviation of historical patient population distribution for roughly normally distributed analytes worked well, but it was an entirely different story for skewed distribution.

The early protocols we set up were heavily influenced by previous literature as well as the training offered by our middleware vendor.

Moving averages is not a new concept at all, but the way we decided to optimize our protocols and select the number of patients to average and the truncation limits was novel.

Bob Barrett: So how you selected these parameters appear to be the key points to your article. How did you choose the number of patient results to average and which patients to exclude?

Dr. Mark Cervinski: That's somewhat of a long story. As I previously mentioned, our early protocols were based on previous literature and/or training in the use of the moving averages software. The number of patients to average for each analyte in the early protocols was determined by calculating the ratio of the patient population standard deviation to the analytic imprecision standard deviation of the assay and applying it or that ratio to a previously published nomogram. This strategy worked for us for tests such as the electrolytes, bicarb, a few others, or at least we believed it did, as we didn't have any objective evidence at the time.

Around that same time, David Ng, one of the coauthors on our manuscript, was a resident rotating through the chemistry laboratory. David had experience using various programming languages, and we had him simulate the moving average protocols in MATLAB and induce systematic error into historical data to calculate the average number of patient samples affected until an error was detected.

The results of David's calculations were eye opening. While we were able to detect error on a few protocols, it was clear that we were not doing as well on other assays. We made some changes based on David's findings, beginning with separating inpatient and outpatient populations when it became clear that there were substantial differences between the populations. However, the real improvement came when David proposed a radically different strategy for optimizing the protocols.

David proposed that we let MATLAB select the best parameters to detect systematic error through the use of a simulated annealing algorithm. I admit that I had not heard of the algorithm before, but David made it sound quite appealing.

Bob Barrett: Well, let's talk about that. What is a simulated annealing algorithm and explain how that works?

Dr. Mark Cervinski: Sure. Without going into too much detail, the simulated annealing algorithm works by randomly varying parameters with a goal of minimizing a cost function. In short, the algorithm randomly selected the number of patient samples to average as well as the high and low truncation limits with the goal of minimizing the average number of patients that would be affected until an error is detected.

Although the algorithm function is available in MATLAB, there was additional work that needed to be done to apply to our specific task. We of course needed data and a plan on how to simulate error conditions for each assay we wished to optimize. Thankfully, we had been accumulating historical data prior to David's rotation, and we had 400 days of patient data from our laboratory with which we could work.

We then had to decide on where to induce the error and how many times over the course of 400 days to induce the error. After determining all of these parameters, David developed additional scripts that allowed us to use the algorithm. To test the algorithm, we devised a way to calculate the average number of patient samples affected until the error would be detected (or ANPed). Using this metric, we could then compare the error detection ability of our initial protocols to the optimized protocols.

Bob Barrett: And did you see improvements over your initial protocols following the optimization?

Dr. Mark Cervinski: Yeah, absolutely. In all cases, the simulated annealing algorithm resulted in protocols that could more rapidly detect error. The improvements, however, were variable and some protocols were vastly improved, whereas others,

such as potassium which was already good prior to the optimization, showed little improvement.

Bob Barrett: Since you've been using moving averages now for a few years, have you been able to detect errors using the protocols developed for this paper?

Dr. Mark Cervinski: Yes. We've been using these optimized protocols for nearly three years at this point, and there have been a number of times over the last year that the moving averages protocols have detected errors well in advance of the next QC event. We share one such example for a sodium ion-selective electrode in the article.

Bob Barrett: Well, finally, doctor, is there anything else about this strategy that you'd like to share with our listeners? What should we know?

Dr. Mark Cervinski: One other phenomenon that we described in our article is that the mean values for inpatient and ambulatory patients differ substantially for some analyte such as calcium, albumin, and total protein. These differences are noticeable enough that early on, we noted many mornings that we would detect false shifts in the mean value for this test. This phenomenon was clearly due to shifts in the population distribution and not shifts in assay performance.

To correct for this challenge, we split our population into ambulatory and inpatient groups, then we optimized each group separately. We found that separating the populations improved error detection, and we now have three moving average protocols for many analytes, one for all patients, one for ambulatory, and one for inpatients.

We view this as a novel idea that could readily be adopted by many labs without the need for optimization via computer modeling of the process.

Bob Barrett: Dr. Mark Cervinski is Director of Clinical Chemistry and Point-Of-Care Testing at Dartmouth-Hitchcock Medical Center and Assistant Professor of Pathology at the Geisel School of Medicine at Dartmouth. He's been our guest in this podcast from *Clinical Chemistry*.

I'm Bob Barrett. Thanks for listening!