Clinical Chemistry

Automating the Detection of IV Fluid Contamination Using Unsupervised Machine Learning

Article:
Nicholas C Spies, Zita Hubler, Vahid Azimi, Ray Zhang, Ronald Jackups, Jr., Ann M Gronowski, Christopher W Farnsworth, Mark A Zaydman.
Automating the Detection of IV Fluid Contamination Using Unsupervised Machine Learning
Clin Chem 2024; 70(2): 444–52. https://doi.org/10.1093/clinchem/hvad207

Guests: Dr. Nick Spies from Washington University in St. Louis School of Medicine and Dr. Mark Zaydman from the Division of Laboratory and Genomic Medicine of Washington University in St. Louis.

Bob Barrett: This is a podcast from *Clinical Chemistry*, a production of the Association for Diagnostics & Laboratory Medicine. I'm Bob Barrett. Clinical laboratory test results play a central role in guiding patient care decisions, but what happens when errors in sample collection introduce IV fluid into a patient's blood samples? Are results generated from contaminated specimens accurate? Analytically, yes. Do the results reflect the true concentration in the patient's blood and should they be used to inform clinical decisions? No and no. Clinical laboratorians recognize that pre-analytical errors, including IV fluid contamination, represent a risk to patients, but traditional tools to detect this type of error are woefully inadequate. Clearly, an innovative approach is needed, but what will this new tool look like and how can it be integrated into the routine lab workflow?

A new research article appearing in the February 2024 issue of *Clinical Chemistry* describes the use of unsupervised machine learning to detect IV fluid contamination, an approach that addresses many of the shortcomings of traditional tools. In this podcast, we are excited to chat with the article's lead and senior authors. Dr. Nick Spies is the clinical pathology chief resident at Washington University School of Medicine in St. Louis where he focuses on informatics approaches to pre-analytical quality improvement.

Dr. Mark Zaydman is an assistant professor of pathology and immunology in the Division of Laboratory and Genomic Medicine of Washington University in St. Louis. His primary clinical role is in the section of pathology informatics where he focuses on utilizing data analytics to improve the quality, efficiency, and value of laboratory services.

So, Dr. Spies, we will start with you. Let's get very basic, just what is IV fluid contamination? How frequently does it occur, and well, why do we care?

Nick Spies: Yeah. So I think fluid contamination refers to the presence of exogenous substances such as medications or crystalloids

within a specimen. This matters because contamination can interfere with laboratory testing either by directly altering the concentration of a measurand or by interfering with the analysis of others. These errors may lead to significant patient harm. As an example, a patient with altered mental status presented to the emergency department and was started on an infusion of normal saline with 5% dextrose. Their first basic metabolic panel came back with a glucose of about 1,000 mg/dL and so their clinical team administered insulin.

That patient soon developed a seizure and a point-of-care capillary glucose actually came back near the lower limit of detection. It's pretty likely that the specimen was actually contaminated by a dextrose-containing fluid, which led to a markedly elevated but erroneous initial glucose result and some downstream problems. As far as how frequently it occurs, that's actually a really good question and we don't have a great answer for it.

There isn't a real gold standard method for detecting IV fluid contamination, nor is there a universal definition for the type or severity of contamination that we would consider clinically significant. However, the IFCC does have a quality metric for these events and it defines IV fluid contamination as the frequency of results that end up being canceled due to severe contamination. In 2019 they approximated that a median of about 2 in 10,000 specimens are canceled due to contamination.

Bob Barrett: How do we detect this contamination, and when we do, what comes next?

Nick Spies: The current standard for detecting IV fluid contamination relies on systems of verification logic that are programmed into laboratory information systems. This often prompts a manual review by the technologist or medical director and often consists of things like feasibility limits or delta checks. These methods have their limitations, which may include a potential lack of sensitivity, the expense of consuming the valuable technician and medical director's time, and a delay in reporting the results.

There is no real standard approach to what action should be taken once contamination is expected. Speaking for what we do at our institution, the procedure is to call the care team to discuss the possibility of contamination whenever we do kind of land on contamination as being suspicious. We often ask the care teams was the specimen drawn from a line that was infusing exogenous fluids? Were there any major patient status changes before the result was collecting? Or is this result consistent with the clinical picture that you're seeing?

If that conversation ends in a consensus being reached that the specimen was likely to be contaminated, we'll cancel the order and ask that a new specimen be redrawn if it's necessary. Obviously, this process is relatively laborious and it does interrupt the workflow of both the laboratory and the clinical teams, so an automatable solution would be pretty preferable.

Bob Barrett: Doctor, what is unsupervised machine learning and why do you think that it might help detect IV fluid contamination?

Nick Spies: Yeah. We can start by defining machine learning more broadly. So, machine learning is the use of statistical algorithms that can improve their performance on a task without needing explicit instructions on how to do so.

Speak in English, a computer solves a problem without the human telling it exactly how. This often takes the form of supervised machine learning where the algorithm is provided a specific set of labeled training examples, and then it learns to assign those similar labels to new examples it's never seen before. A major challenge in supervised machine learning is producing large amounts of high-quality labeled training data. In unsupervised machine learning, the algorithm simply learns patterns in the unlabeled training data.

Some examples of the unsupervised paradigm that might be familiar to these listeners are the various clustering algorithms like hierarchical clustering and principal component analysis. We thought to use an unsupervised machine learning algorithm here because for IV fluid contamination, we often measure many various analytes all together in a single panel like the basic metabolic or comprehensive metabolic panels, and we expected that these contamination events would actually cause more correlated changes along many of the analytes which make it relatively conducive to machine learning algorithms.

The added benefit of unsupervised machine learning also helps with the challenge of not having a high quality and scalable access to gold standard labels for IV fluid contamination.

Bob Barrett: So has anyone used this approach to detect IV fluid contamination before, and if so, how does this approach differ?

Nick Spies: Yeah. So there have been many prior examples of different types of approaches for detecting IV fluid contamination, some with machine learning and some without. Dr. Jason Baron over at Roche actually reported the use of a supervised machine learning approach to build a set of decision trees for detecting dextrose contamination that worked quite well, and

the more recent work from Dr. Tom Durant's lab over at Yale has derived a pretty sophisticated set of multianalyte delta checks for detecting normal saline, lactated ringers, and 5% dextrose.

We hope to build on a lot of this prior work by attempting to find a single solution that can be generalizable across many fluid types that would not require a recent prior for delta checks and would not also require a very extensive and expertly labeled training data set.

Bob Barrett: So let's talk about your study. What did you do there and what were some of the findings?

Nick Spies: So in our study, we employed an unsupervised machine learning algorithm called the Universal Manifold Approximation and Projection Algorithm, or UMAP. What UMAP does is it simplifies highly dimensional data into a lower dimension representation. We applied UMAP to a large clinical dataset that consisted of about 25 million basic metabolic panel results over the span of six years from Barnes-Jewish Laboratory in St. Louis, Missouri.

The result of this transformation was pretty intriguing. We observed that most clinical results embedded onto the central density of this lower dimension representation, a manifold, consistent with relatively normal results. However, there were a series of projections that kind of extended outward from the central density. We hypothesized that these extensions outward might represent the contaminated specimens and so we simulated IV fluid contamination and then apply that same UMAP model to the simulated contamination. The resulting plot for the simulated data shows that those projections were even more prominent with the more severe contamination actually living further out on the projections than the mild contamination or the real patient data.

Our next step was to turn these plots into more objective metrics that could be used for classification. So we developed something that we call the enrichment score which compares the density of simulated contamination to that of the real patient results for each coordinate on the manifold.

Finally, we demonstrated the enrichment score could be used to predict whether a new patient result was contaminated and validated the performance of that score using a prospective data set with consensus of expert review scores as the reference method. A particularly interesting finding from this work is that a pretty substantial proportion of the contaminated results from our expertly reviewed prospective validation study were actually missed by our current approach, but detected by the UMAP model. This certainly

raised our concern that we are frequently reporting contaminated results into the electronic medical record.

Bob Barrett: So how would you plan to use a tool like this in the clinic?

Nick Spies: That's a really important issue and that certainly needs quite a bit more exploration. Our current systems for detecting and managing these errors may lack sensitivity, but increasing the proportion of results that we flag or cancel may not necessarily be the best approach when considering all the other complexities surrounding patient care at the hospital. When we speak to our clinical colleagues, one of the main pain points that they often highlight is a delay in turnaround time and one of the major benefits of an automatable solution would be reducing the operation burden that these events do place on our routine workflows including that turnaround time.

We could imagine a future state where results that are predicted as contaminated by a machine learning model are auto verified into the medical record and resulted as something like "see comment" or we provide an interpretive comment that provides the measured results with a disclaimer stating that these results are likely to be contaminated by IV fluids and should be interpreted with caution. However, using a prediction from a machine learning model as criteria to auto-verify results certainly represents a stark divergence from the current paradigm and we would need a pretty substantial amount of validation both of the machine learning model and of the implementation science behind how we would plan to use it before such a solution could be safely and confidently implemented.

Bob Barrett: Thank you, Dr. Spies. It sounds like some fascinating work. Dr. Zaydman, you've been lurking in the background during all this. What are some of the next steps for you in this work?

Mark Zaydman: Thanks for that question, Bob. Happy to lurk, but happy to join in. Ultimately, our goal is really to design tools that are going to be useful in practice, and so building on this work, we're going to work towards a future state where we can apply such an algorithm clinically, so we can benefit our patients by improving detection of IV fluid contamination, and we can also improve efficiency in the lab by automating these very expensive manual review processes. So, we really see this current study as an initial step only in this direction and we're excited to build on it towards our goal of clinical translation.

Steps that come to mind that we are actively working on, we would like to explore using supervised machine learning to further improve the accuracy of our prediction models. We believe that a hybrid approach, that is an approach that

combines both unsupervised and supervised machine learning methods, would really be ideal, and it's going to capture the unique and distinct benefits of each method. We want something that's going to achieve high degree of accuracy, which usually means supervised machine learning, but we want to avoid excessive burden of labeling training data, and we also want to maintain robustness to contamination events that are not well represented in our training datasets, and that's where unsupervised machine learning is attractive.

We need more, as Dr. Spies pointed out, real-world clinical validation studies. We don't yet understand how these algorithms are going to perform in the real world. What will be the impact on patient care? What will be the impact on laboratory operations? And so, we need some sort of prospective, likely silent, validation studies, so that we can get an unbiased estimate of real-world performance.

Finally, we would like to develop tools that not only help us here at Barnes-Jewish in Washington University, but rather algorithms that will generalize to other institutions. The question of generalization is a tricky one in machine learning, and this is going to require extensive collaborative external validation studies. There are key questions there. Are these models going port and block? Are we going to have to do additional training, fine-tuning, et cetera?

So putting this all together, we have a goal in mind. There's a lot of work to do, but we're excited to have made this first step and to continue with the next steps.

Bob Barrett: That was Dr. Mark Zaydman and Dr. Nick Spies from Washington University in St. Louis School of Medicine. They wrote a research article evaluating the use of unsupervised machine learning to detect IV fluid contamination in the February 2024 issue of *Clinical Chemistry*, and they've been our guests in this podcast on that topic. I'm Bob Barrett. Thanks for listening.