# Episode 52: Errors in Toxicology - When Science Gets It Wrong

This transcript was created by AI from audio and lightly edited for readability. Treat it with caution, and assume occasional transcription errors may remain.

**Transcript**

Hi everyone, welcome back to the Tox Lab. I'm Rebecca. I'm Rob. This week, we're going to be taking a look at when toxicology goes wrong and some of the miscarriages of justice that have resulted. Before we dive too hard into the science, I wanted to think a little bit about this idea of injustice.

Rebecca, therapy time. Got any memories of great injustices as a child? I mean, I think it's just classic. Your younger sibling gets away with so much more than I was ever allowed to get away with.

What about you? I definitely have a memory of, I can't even remember what happened whereby I was accused of breaking something that I didn't break. And that feeling as a child, where you're telling the truth and no one believes you is crushing. Yeah.

And I think as adults, it's not really something, I certainly don't really come across it very much. But every so often, the justice system gets it wrong. And when it happens to you as a child, and it's something silly, like something got broken and you got in trouble, in the grand scheme of things, it's really not that big a deal. But imagine being sent to prison for a chunk of your life, purely because the lab got it wrong.

And that's what we're gonna be looking at this week. So before we jump into today's episode, if you wanna stay up to date with our content, you can find us on Instagram at @the_tox_lab. We also have a page on LinkedIn and we're fairly active on there. And for those of you who don't use social media, but still wanna get in touch with episode suggestions or comments about our episodes, you can email us at thetoxlab@gmail.com.

So before we dive too much into today's paper, I want to tell the story of a particularly high profile case whereby the toxicology lab got it wrong. And as a result, there was a significant miscarriage of justice. So I wanna take us back to 1989. So quite a little while ago, Madonna's like a prayer was on the radio.

Solid song. And we're gonna be looking at the story of Patricia Stallings. And Patricia Stallings lived with her husband. They had a five month old son, Ryan, who developed signs of serious illness.

He was vomiting, showing signs of distress, was taken to the hospital. He had severe metabolic acidosis. And the hospital were concerned that this could have been antifreeze ingestion. They sent the sample to the lab.

The lab identified ethylene glycol, the component in antifreeze that's poisonous. And as a result, she was arrested. The child sadly passed away. She was charged with murder.

And she was sent to prison pending the trial. And while she was in prison pending trial, she actually bizarre twist of fate because this is actually ultimately what exonerated her. She was sent there whilst pregnant and gave birth to a second child whilst in prison. And this child also became very unwell, developed similar symptoms, metabolic acidosis.

And of course at this point, there was no way she could have poisoned the second child she was in prison. And this child was subsequently diagnosed with methylmalonic aciduria, which is an inborn error of metabolism, genetic disease that leads to altered biochemistry. And so the question was raised, could her first child have also had this same condition? And so the doctor that identified the methylmalonic aciduria in the second child actually retested the original samples and couldn't find any ethylene glycol.

Instead found propionic acid, which is structurally very similar to ethylene glycol. And it turned out that that initial hospital analysis was actually incorrect. There was an interfering peak from propionic acid and that led the hospital to misidentify the cause of illness in the first child as ethylene glycol. And as a result, Patricia Stallings was exonerated.

The case was thrown out, but she had already spent a significant chunk of time in prison all because a lab error misdiagnosed that first condition. And so this week, we're gonna be having a look at a paper that came out this year, didn't it? This paper has looked at significant number of cases where toxicology labs have got it wrong. We're gonna be looking at that paper and we're gonna be exploring what went wrong, why it went wrong, and hopefully go away thinking about what we can do to stop it happening in the future.

So Rebecca, where do we start? So this paper covers notable errors that have occurred in the toxicology world that have been collected by the two authors, Aaron Olsen and Charles Ramsey, over their combined 48 years worth of field experience. So they've split the different errors that have occurred in different categories. We're just gonna work through them one by one.

So the first category is traceability errors. So some examples of errors that have occurred include the use of inappropriate reference materials for developing methods, incorrectly produce reference materials with calibrate installations, and incorrect concentration calculations when developing methods. So let's just break it down a little bit, traceability errors. In other words, using a reference material which is not 100% traceable to what you're trying to measure.

I don't want to go too hard into meteorological traceability. I think that's probably a whole different podcast. But essentially when you're building your calibration curves using something that isn't appropriate for what you're trying to measure. So in one particular example, they had incorrectly produced their reference solution used to produce their calibration curve, which meant that when they calibrated their breath alcohol analyzers, they were calibrated.

So the results were 20 to 40% higher than they should have been. So everyone was reading as having a breath alcohol level higher than what was actually true. And I'm trying to think about how that could apply to our lab because I've got zero experience with breath alcohol. But actually this is perhaps like incorrectly diluting a reference material, right?

Or incorrectly weighing a reference material really early on in that process, which dilutions and weighing is something we do relatively often. Effectively that would be that, wouldn't it? A break in that traceability. So calibration errors have also occurred.

And often these have been where people have developed a single point calibration curve. So you have a blank and then one point for your calibration curve. And when you're then getting a result that's off the top of that calibration curve, some labs weren't doing anything to dilute the sample to get it back in within the measurable range. And we're just reporting it.

So assuming linearity outside of your calibration curve. One point calibration curves do have a place, although do they have a place in forensic toxicology? I'm not sure. It's tricky when your single point calibration curve doesn't cover a decent concentration range and you're going to get samples that are

repeatedly hitting off the top of that range.

And if you don't treat those appropriately, you're not necessarily getting reliable results. That's fair, yeah. The next category of errors involves discovery errors. So not showing details about failed calibrations or problems surrounding the rest of the batch with the defense, withholding of worksheets of failed calibrations of breath alcohol analyzers and generally withholding evidence.

One notable case occurred in April, 2023, where the Massachusetts Supreme Judicial Court ruled that approximately 27,000 drunk driving cases were eligible for new trials, all withdrawn guilty pleas due to the misconduct by the Office of Alcohol Testing as a result of withholding this information. Well, this is a huge impact, isn't it? 27,000 potential retrials or dismissals of previously convicted cases. I had to think about what this discovery error actually means for us, because obviously the UK legal system is quite different from the US one.

We have different rules regarding, we don't call it discovery, we call it disclosure. And it's tricky, isn't it? Because some of this was argued as not being required or rather it was argued that it shouldn't be disclosed. A lot of this stems from labs being quite protective, doesn't it, over their data?

And yeah, on the other hand, if a lab is sitting on data that could stand in the way of justice, then that information should be transparently shared. So another category of errors involves maintenance of analyzers. For example, the Data Master Breath Alcohol Analyzer had a recurring problem with a valve where it'd intimately get stuck in the wrong position, making it impossible for a subject to blow into the sample chamber and provide a sample of breath. The problem was there was no way for the operator to know that this was the case, which would lead to individuals being additionally charged for refusing to provide a breath sample.

Failure to provide a sample is often treated as a more serious offence than, certainly in the UK, as having a sample above the drink drive limit. If you're trying with all your might to provide a sample and the machine is simply not letting you provide a sample, that must be crushing. Yeah. Even if the person is above the limit, still they're trying to comply and provide a sample.

I think sort of more closer to home, maintenance is one of those things that we do and we don't really think about we do it, do we? Sometimes it's a chore, or have you done the maintenance on the mass spec this morning? It just shows how it is important in order to maintain a system that's validated and defensible, that if there is an issue, particularly if that issue is known to manufacturers and maintenance is part of the process to overcome that, it shows that actually this job that we see as a chore and a time sock actually could still be really important. Yeah.

And sometimes you can't really appreciate how important it is until months or years later. Another issue surrounding both instrumentation and software has occurred previously where devices have been approved for use often by a federal agency and then have subsequently been modified after this approval process, which brings into question how accurate and reliable that instrument is because it's not been checked and made sure that it's still in good working order after these modifications have occurred. And I've been trying to avoid putting my quality management hat on, but I'm gonna put my quality management hat on for a moment.

The ISO standards are quite clear when it comes to validated systems that are subsequently modified, effectively potentially no longer validated and therefore needs revalidation. I'm trying to think if there's any examples of that in the lab and we don't tend to modify our devices very much, but do we occasionally modify our methods? It does make you think, doesn't it, about whether there are small modifications that occur in processes, for example, that maybe do have a bigger knock-on effect than we necessarily think about?

I think it's until you hear about stories where cases are being thrown out or results are coming into question that you then realize the impact of that modification and what that could mean further down the line. What sort of modifications were done? Do they go into it? The example they give in the paper is to do with the Intoxilizer 8000, where there were modifications to software, the pressure transducer, the computer board.

Software, I think, is a big one, actually, that people don't always think about. You get a software upgrade on the computer that controls your mass spec and hopefully the manufacturers have done their due diligence and the input outputs were the same as the old software, but sometimes that isn't the case. That moves quite nicely into source code errors. So there have been examples where there have been more lenient control ranges and analyzers.

So there have been wider tolerance limits for results. So things have passed based on what the analyzer is saying that wouldn't have passed otherwise. Other source code errors include rejecting a valid sample size of breadth because of a programming error, and then various errors in the code within the instrument, meaning that it failed to meet standards used by federal agencies. So that is obviously gonna have a massive impact on...

Any case that had been, that had relied on that device up to that point, rounding errors too can be another one within software. Sometimes software rounds things in a weird way. And if you're working to a defined cutoff, that can be the difference between guilty and not guilty. So interfering substances in methodology, which we mentioned at the beginning with the Patricia Stallings case, can be a massive problem.

Other examples have included labs where they've not been able to distinguish between Delta-A and delta-9-THC, and the law is quite clear that it's delta-9-THC that is illegal. Yeah, this was the error that really got me thinking the most actually, again, with my quality management hat on, because what we like most of all in toxicology is methods that don't change. We love validating a method, and then we go, great, this method hasn't changed. We've been using this method for 20 years.

The equipment's the same, the process is the same, everything is the same. And it's possible to have validated a delta-9-THC method 20 years ago, have an absolutely beautiful method that met all of your validation criteria. And what has changed in this case isn't the method, but the world. Because when you validated that method, Delta-8 wasn't on anybody's radar, it wasn't a consideration.

You've now got an isomer of Delta-9 cropping up in casework, and you're now misreporting on your beautiful assay that was once great. It does depend on the context of what you're doing, doesn't it? But as you say, if you're doing a drug drive case and Delta-9 is illegal and Delta-8 isn't, that's quite significant. Yeah, absolutely.

And there have been other examples, including where isopropanol that has been produced from conversion from acetone has been mistaken for ethanol in alcohol breath tests. So there was a case where an individual with type 1 diabetes who was going through a diabetic ketoacidosis episode was incorrectly testing positive for ethanol when in fact it was a complication of his diabetes. Yeah, volatile alcohol cross-reactivity actually is something I've come across as well. Methods not able to tell the difference between acetone and isopropanol and things like that, which can lead to incorrect diagnosis.

So another class of errors that I found quite interesting when reading this paper came about due to lab contamination. So the Washington State Toxicology Lab moved into a space that had previously been used to synthesize methamphetamine for training purposes. And this resulted in nine cases testing positive on a screen for methamphetamine between October 2018 and June 2019 that could not be

confirmed on follow-up testing. Environmental testing of the lab was then carried out which revealed meth residue on the ceilings, walls, air vents, and floors that exceeded state occupancy limits.

And what was worse is this error wasn't disclosed for over a year and was kind of a throwaway comment on a different case that was made. You've now got two errors there, haven't you? The first is maybe inappropriate use of premises following use of it for another reason. And I've come across situations like this as well, actually where bulk analysis is being done in one part of a lab and an air conditioning unit is moving that air around containing cocaine dust to the lab that's carrying out urine toxicology.

But you've also then got the fact that this error wasn't disclosed, which is a bigger issue in some respects. Reporting errors was another big one. So one example is the Minnesota Tri-County Regional Forensic Lab made errors in impaired driving cases from January 2010 to July 2010 where they had failed to convert their units from grams per 100 millilitres to grams per 67 millilitres for urine alcohol testing. So all the results were a third higher than it should have been which resulted in 11 incorrect impaired driving charges.

And this is a basic maths error, isn't it? But it should have been picked up by case review or... Yeah, it definitely shouldn't have gone on for six months. And it's impressive that only 11 charges were incorrect.

Yeah, that's true. Another reporting error affected a pregnant woman by the name of Eileen Bauer who ate a linguine salad made with a dressing containing poppy seeds and shortly after dinner went into labor. A routine admission drug screen flagged positive opiates despite passing all previous screens and the presence of morphine ingredient in her sample led to the Lawrence County Children and Youth Services obtaining a court order and removing the charge from her care without confirmatory testing.

It turned out that in this case, the opiate cast off threshold was set to 300 nanograms per millilitre rather than the federal workplace standard of 2000 nanograms per millilitre. And this resulted in the infant remaining in foster care for 75 days before being reunited with mum. Really sadly, this isn't the only case that this has occurred in. And this was really a two-fold error here, wasn't it?

Because firstly, the wrong cutoff was used for the screening test but also no confirmatory testing was carried out. Yeah, which always seemed shocking to me that you wouldn't want to confirm such an important result as that. And not confirming results has been other errors that have reported people reporting immuno-assay results without confirming on a mass spec. And those of us who work in SOTS College you know that there are many interfering substances that can cross reality with an immuno-assay giving a false positive result which is why we do confirmatory testing in the first place.

Yeah, absolutely. There have also been errors involving chain of custody. So there was one case that they mentioned in the paper where someone had tested positive for cocaine on a urine drug screen despite their hair testing being negative. After some more investigation it turned out the urine sample contained two DNA contributors showing that it had been tampered with after collection to produce a positive result.

And we don't know where that additional contribution came from only that at some point in that samples collection process there was another contribution made. So the final category of errors that they've covered in this paper is fraud. So example, this include certification records not being signed by the individual certifying the analyser. There was a notable case in Illinois in 2006 where individuals were manipulating the date and time settings on instruments to backdate certification tests to meet the statutory requirement that an instrument needs to be certified every 62 days.

There were forged calibrations to make it appear as if calibration checks had been performed when they hadn't. So these are some quite serious errors. And I think that on a scale of fraud, again, this is similar to what we were saying about how maintenance is seen as a chore, isn't it? That, oh, it's just maintenance.

It doesn't matter if it's not been done. I'll just sign it to say that it has. Actually, the knock on implications could be, A, well, the equipment could go wrong as we saw in the maintenance errors. But actually, even if the equipment doesn't go wrong, the fact that fraud was involved anywhere in that process, you've potentially completely invalidated any results generated.

So even what looks like on its surface in inverted commas, small fraud could still have very big implications in the legal system. Yeah, and especially when it comes to forging calibration checks to say that you've done your checks and that it's all passed and fine, you could just be producing completely nonsense results and have no idea that that's the case. Yeah, for sure. The final case I want to dive into involves Randox.

And there were concerns over manipulation of results, which affected over 10,000 criminal cases across 42 police forces between 2013 and 2017. 50 criminal prosecutions were dismissed due to compromised evidence, and these covered cases involving road traffic offenses, violent crimes, sexual offenses, and unexplained deaths. And this was such a huge shake up in the UK, wasn't it? UK forensic science sector.

And it was a few analysts manipulating chromatograms to effectively forge QC results to make a run pass. And I remember at the time when this all kicked off, you are left thinking, could that ever happen in my lab? That would never happen in my lab. But I think you've got to take a bit more responsibility than that as a lab leader, haven't you?

Because it becomes a cultural thing. It's not, do I have bad staff? It's sometimes, am I putting too much pressure on the staff to get the results out? Is the culture to turnaround time driven?

Do the staff feel empowered enough to say, I'm really sorry, Rob, this batch has failed. We're not going to get those reports out today like we wanted to do so. And I think that's so much harder to actually look inwards and go, actually, do we have that culture? Do we have that resource built into the system to account for the fact that sometimes batches fail and we need to not report?

Definitely. And it's hard sometimes when you're under that sort of pressure to take a step back and reevaluate what those results mean. Like it's not just numbers on a report or peaks on a screen. It's whether someone's going to get charged with something or not, or whether you're going to report the presence of a drug that could then lead to a child being removed from its family environment.

There are really big stakes at play within the world of toxicology. And I think sometimes just taking a step back and remembering that that is the case is really important. Yeah, 100%. The outcomes of some of what we do can be so significant, can't they?

And we do need to really consider that. This whole paper is really worth the read, isn't it? It was not an easy read. In fact, it was quite a hard read at times when you look at all the ways in which things can go wrong.

But the general conclusion of the paper is that we need to be transparent and we need to be open to the fact that things do go wrong. And when it does, not try and hide it. That's the thing, we're all human. We make mistakes.

That is a part of life. And being able to hold up our hands and go, guys, this hasn't worked. Let's start again from scratch. I think it's really important.

Yeah, definitely. Toxicology is a weird discipline to try and work within. My background is healthcare, general biochemistry within a hospital setting. And I still do a little bit of general chemistry.

And I know how often I get a phone call from a doctor out on the ward who says, this blood test doesn't look right for this patient. And the first thing we go is, yeah, sometimes these tests don't work right. Let's work with you. You've got the patient in front of you.

Let's see if we can figure out what's gone on. You don't have that option in toxicology. If we report the presence of a drug and somebody says, I don't care what your result says, I haven't taken cocaine. The assumption is that the person who's given the sample is lying, not that the test is wrong.

So it is especially important. I know that it's not quite comparable because mass spec versus an immunoassay and so on. There are obviously technical differences in how we generate those results, but the implications are very different, aren't they? And I think having worked in quality management and seen just all the ways a big lab can go wrong from small errors to very large errors, I think you are left wondering a little bit, okay, I've validated my systems.

I think I've got all my ducks in a row and I think everything's functional. Are there things going on that I don't know about that I haven't yet discovered? Hmm, that is very true. I think a lot of these cases in this paper, they weren't something happened and the error was picked up.

These were errors that had persisted for a very long period of time, either undetected or worse swept under the rug. And it was only when somebody external came in and actually identified it. But also that's the other sort of side of this, isn't it? That we do have external people come in all the time.

We have our labs accredited to ISO standards. Labs get assessed against forensic codes of conduct. That didn't necessarily stop some of these errors happening. Just because a lab has been approved as working to ISO 17025 or 15189 doesn't actually mean stuff isn't happening under the hood that isn't picked up.

So I think a lot of the takeaways from this really were, actually there's a lot of ways in which things can go wrong. And some of them I think we can go away from here and go, okay, I'm gonna think about my calibration traceability and things like that. Errors relating to fraud and data manipulation, that's much harder to address. That's a cultural thing if you're in a position of lab leadership.

And I think for me, the big takeaway from this is the Delta 8, Delta 9 and the fact that a method can be perfect and suddenly no longer fit for purpose because the world has changed. And I think that is why as toxicologists, we need to be not just living in our bubble of the tests that we validated, but actually being aware of what is going on around us. Because that could apply to lots of things, couldn't it? Not just cannabis and Delta 9, but potentially lots of new drugs that are emerging on the market might start cross-reacting with some of our other drugs.

We do need to really consider that. Well, thank you for joining us this week. I do hope that it's been an insightful listen if it's not necessarily been an enjoyable listen. And I hope that it's inspired people to move forward and as a discipline, we can do better in the future.

Hope you will have an amazing week and we'll see you next time, bye.