Episode 36: AI in Toxicology - Potential and Pitfalls

Transcript

Hi everyone, welcome back to the Tox Lab. I'm Rebecca. I'm Rob. This week we thought we'd have a look at AI and toxicology and some of its possible implications. AI is everywhere, isn't it Rebecca?

Yeah. It's become a big buzzword. Lots and lots of companies building AI solutions to things. Possibly AI solutions to things that don't need AI solutions some of the time.

On the other hand, we've seen huge steps forward, haven't we, in the last 18 months. Massive improvements in AI models like ChatGPT. I've been messing around with generative AI for the last couple of years. And the improvements we've seen in art quality from the original DALL-E image generator to where we are now really is quite significant.

So we thought we'd have a look at AI and how it applies to toxicology. There was a recent letter in the Journal of Analytical Toxicology, wasn't there? Talking about high res, accurate mass, mass spec data and data mining. And this really got us thinking a bit about the role of AI in that process and the role of AI in toxicology in general.

Before we jump into today's episode, if you wanna stay up to date on our content or get in contact with episode suggestions, you can find us on Instagram at @the_tox_lab. We also have a page on LinkedIn. And for those of you who don't use social media, you can email us at thetoxlab@gmail.com. So how does high resolution, accurate mass, mass spectrometry fit into this wider discussion of AI?

Up until relatively recently, most labs have been doing targeted toxicology analysis, whether that's using GC, GCMS, LCMS, in which you develop a method for whatever drug you wanna look for, and then you go and look for it. In the last few years, we've seen an increase of high resolution, accurate mass spectrometers. And these work in a fundamentally different way. They're capable of acquiring all of the data in inverted commas.

They can acquire huge amounts of untargeted data, and then you can effectively ask questions of your data rather than specifically targeting your method to a particular compound. You're asking your data, is this particular compound in there? So even though high res mass spec can acquire so much extra data and is effectively untargeted, most people are still using it in a targeted way, but they just have more targets that they're looking for because there is just so much data to sort through. So effectively, we're still doing targeted work.

You still need to know what you're looking for in your data set, even though the data has been acquired in an untargeted way. Now, I don't wanna get into the detail too much about how high res accurate mass spectrometers actually work. Here's an analogy to explain high res mass spec versus traditional LCMS. You're given a smoothie.

Okay. LCMS says it's a strawberry and banana smoothie. High res accurate mass can zoom in and find all of the molecular parts of the strawberry, all of the molecular parts of the banana, all of the molecular parts of that dash of kale that ended up in there because the blender hadn't been rinsed out from the last smoothie. Ultimately, you come to the same conclusion.

It was a strawberry and banana smoothie, but the amount of data that was generated quite a bit higher. Either way, fundamentally, the way we're using high res instruments most of the time is

essentially targeted, right? We acquire all of this data and then we drill down into it and ask questions of the data. Is fentanyl in this data set?

Is fentanyl metabolite in this data set? We're actually only looking at a fraction of the data that's been acquired. And so there's a large amount of background data that isn't sorted through. And this is a point that was brought up in the letter, which I will link in the description of the episode below if you wanna go check it out.

Because you've got so much extra data to look through, it would take someone hours and hours to go through it all and pick apart every PQC and try and look through this mess of untargeted stuff. And the software that we currently have isn't really built for looking for all of these unknowns. And that is a really big problem because you could be missing something that because you don't know it's there and you're not looking for it could be important in your case. So this is what got me thinking about the power of AI and large language models because we only look at a fraction of our data that we acquire now.

And we generate absolutely gigabytes. I worked out that a single run on our high res mass spec produced as much data as 15 years of our triple quad running day in, day out. Which is insane when you think about it. And whilst the high res machine has been looking for a lot more drugs, ultimately they've been answering the same sort of questions.

So this is where I think AI could really play a role. You could, in theory, feed in all of this data. AI essentially is very good at pattern recognition. If there's any AI scientists listening, I appreciate there's a lot more to it than that.

But at its heart, that is what AI is good at. And so potentially being able to compound predict, identify features, identify common themes amongst cases, it could potentially use literature mining tools to try and assign identities to these. We've already got in silico fragmentation predictions. So potentially it could even match fragment spectra to candidates.

I think it could have a real role in potentially identifying novel compounds. Could flag up where new peaks are appearing in a cluster of cases or things like that. And AI could be even more useful in this sort of situation. If you could feed in case details and it looks at the case as a whole and identifies peaks related to that.

So if you think that it's a case related to a novel opioid and it can then look for those and identify potential peaks that could be related to an opioid, that could be really useful. Yeah, for sure. I think one of the other things that has really captured my imagination is I'm a biochemist by background. There's a lot of endogenous compounds present in a sample that aren't drugs, but are there metabolic clues potentially hidden in the background data that could give you clues towards a toxic drone or a cause of death?

I think there's so much data that we are sitting on that we don't look at because it's hard. Yeah. And actually AI potentially can really help in the right circumstances actually really help revolutionize that. So how would this work in practice then in theory in this magical world, we've got access to this large language model designed for this purpose.

You can imagine a world whereby you upload your high-res data, you input your case details, it sorts it in the context of all the other cases it's been trained on, potentially just gives you a list of things it's found, things it considers relevant. And there has been some work using large language models, hasn't there, looking at autopsy data. I found a paper looking at that whereby features found at post-mortem were input into an LLM and it was able to predict with 95% accuracy what the pathologist would certify as a cause of death. That's pretty cool.

Pretty cool unless you're a pathologist. Well, yes, but. But even then somebody had to carry out the post-mortem, so. While all this sounds amazing, the problem with these sorts of models is that only as good as the data you train it on.

And the problem with the data we look at most often, which is post-mortem tox data, it's really, really messy. And there will be lots of things in there that aren't relevant and you're not always going to be able to pin down what drug was most important in terms of what caused the death, just because often you get so many drugs that are at high enough concentrations that could be relevant. For sure. I mean, we're sorting through, I'm going to regret using this term, a chemical soup.

Oh, no. Of just thousands of compounds. And so you're right. Actually, your AI is only as good as the data you've trained it on.

And if you've trained it on messy data, your outputs are probably going to be messy. And even though it has the ability to sort through all of this data, there is a chance that you could potentially introduce a subtle bias and it starts ignoring things that might actually be relevant to your case. And so you think that it's kind of taking an unbiased view of all this untargeted data that is excluding half the things that you actually might find interesting. Yeah, that's completely fair.

Bias in AI is a big concern at the moment, isn't it? Typically, when we talk about bias in AI, I think people are talking about political bias in large language models. But that logic actually applies to any other use of an LLM as well, producing biased outputs. I think there's another big drawback with this idea in that is that it becomes complete black box technology, doesn't it?

Data gets uploaded, answer comes out. We're not really sure how that answer was derived. I could see this being useful in the clinical world. Clinical toxicology, somebody comes into the emergency department, gives a sample, it gets analysed, it gets uploaded and answer comes out.

You could clinically validate that protocol as having a useful diagnostic capability and you wouldn't need to know how that has worked. Where things get messy is in the forensic world, whereby you can't produce a forensic report that explains how you have reached your conclusions. Because, well, computer said so. Yeah, it's not a great answer, is it, in court?

Being cross-examined because the computer said so. It's true. There are some moves towards explainable AI. Now, again, caveat, I am not an AI scientist, so I'm possibly going to trip over all of this and get it all wrong.

But there is at least one I have come across called SHAP or Shapley Additive Explanations, which is some mathematical terms, I think. But essentially, it is a model that takes in your data and then can produce a report that explains what weighting it has applied to each feature it has used to reach its output.

So, for example, it might say likely cause of death, fentanyl poisoning, and then say, I've reached this conclusion by 70% of my work because fentanyl was present, 10% because there was no fentanyl there, and how each of the different components had provided weight to that conclusion, which could provide some level of explainable report, which could potentially, eventually, have a role to play in a court setting. That is fair. I think another challenge with the front of the world, there's so much nuance that I'm not convinced an AI can take all that and apply that to each case that it sees.

It might be a really good starting point for being able to sift through data and come up with what it thinks is likely, but you're still going to need that human input to kind of fine-tune that interpretation. That was going to be my next question. Are we all going to be out of a job in five years? I don't think so.

I can definitely see it being a really useful tool. I can see it streamlining our work. Oh, for sure. Can I see it replacing us?

At least not in its current form, but artists were saying that when the first DALL-E model came out and these really, really, really janky images started appearing on the internet. Horses with two heads. And now look where we are in the generative AI world. So never say never.

I think where AI in front of toxicology could come in most useful is looking for those unknowns in your massive, untargeted dataset. So often you get a case and you can't see anything obvious. You're spending hours and hours trawling through the background data, trying to work out as a peak for a potential drug that you might think is significant. And it just feels like bashing your head against a brick wall.

So if AI could help identify peaks and potentially ID them better than our current software can, that I think is its real asset. Yeah, that's completely fair. And we have had some success with that recently, haven't we? We had a rogue peak that appeared in our data and it turned up in most cases.

So we thought it's probably not going to be relevant. We managed to track down its formula based on its accurate mass or a likely formula, but we couldn't figure out what it was. And then ChatGPT actually did figure it out, didn't it? It was a, we think it was a plasticizer, a phthalate-free plasticizer that some tube manufacturers had effectively reformulated their plastic production to stop using phthalates.

And so this new plasticizer was suddenly present in all of our high-res data. And it was a success story actually for the AI, figuring out what we couldn't. It was able to scan through PubMed relatively easily and go, actually, I think it could be this. Yeah, no, it definitely has shown promise but things like that.

I think that the trust in interpretation is going to be one of the biggest barriers. Yeah. I think, how much trust do you put into a toxicology report? Some of that is down to the person that produced that report, right?

Yeah. And you take the human element out of that. And I think that trust potentially goes out the window. We already live in a world whereby scientists aren't trusted, whereby experts are considered to be sometimes part of a wider conspiracy theory.

I don't know if you remove the human element even more, does that get even worse? Potentially. I think it's quite hard for people to trust something they don't fully understand and by taking out that human element that can fully explain what's going on, you're not going to get that. On the other hand, humans are flawed and make mistakes.

But AI also makes mistakes. Well, yeah, I was about to say machines don't but then I stopped myself for exactly that reason. You're absolutely right. That's a big point as well, isn't it?

Hallucinations in AI. That was always really freaky. Yeah, yeah, yeah, for sure. I think that's the other thought actually here is if you do move towards AI interpretation and it gets it wrong, who's faults that?

Hmm, that is a really good point. And I think that actually is possibly part of a wider concern. It's possible to have an independent toxicologist with access to textbooks and the internet to interpret data and reach a conclusion. If you're using AI, you need probably to use an existing LLM from a large billion dollar company that's already got the language model running in the background.

And so we're suddenly trusting all of our data to large corporations and trusting them to produce an unbiased answer back to you. So I think until we can get maybe smaller LLMs running at a local level, and it is possible. You can run LLMs on a laptop, believe it or not. Most people don't.

Actually, potentially that's another stumbling point effectively, you're relying on outsourcing your AI work to a handful of companies. So I do think it's gonna be a case of watch this space to see what happens in terms of where AI goes in the toxicology space. I could definitely see space for it. I can definitely see the use of AI as a tool to help us do our job better.

I hope I'm not being naive, but I don't think we're gonna be replaced at least anytime soon, although. Who knows? Yeah, who knows, who knows for sure. I think maybe we'll revisit this topic in six months and we'll see whether the world has changed because it is very fast moving.

It'd be really good to see a better solution than the software we currently have for searching through our massive data sets. Yeah, that's true for sure. And ideally one that doesn't involve us uploading our data to chat GPT. Yeah.

If anyone listening has had some success with AI helping them interpret data or anything like that, I'd really like to hear from them because I do think it's a field where currently there's probably a handful of enthusiasts doing really, really good work behind closed doors. And it'd be really, really good to see what is happening and hear the successes. I do hope you've enjoyed listening this week. As Rebecca said, do follow us on socials to stay up to date with our content or give us a like or a share or tell your friends that would be amazing.

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