AI Use Disclosure Note:

This transcript was generated using Claude Fable v5.1 (Medium effort) on October 6, 2026 by Marc-André Bélair. See Appendix A for prompt.

Text in square brackets has been added to provide clarification or descriptions that may not have been included in the original presentation.

AI for Public Health — Session 1: Introduction to AI in Public Health (transcript)

Marc-André Bélair:

Good morning and good afternoon, everyone, depending on where you find yourself today. Welcome to this inaugural session of the Artificial Intelligence for Public Health seminar series, Introduction to AI in Public Health. My name is Marc André Bélair, and I'm an epidemiologist and trainer at the training and development unit at the Regulatory, Operations, and Emergency Management Branch here at the Public Health Agency of Canada. I'm based in Vancouver, BC, and I'm really honored to be your host today. Une note pour mes collègues francophones. La session d'aujourd'hui se déroulera uniquement en anglais, mais la vidéo de la session sera disponible en ligne avec des sous-titres en français sous peu. Par contre, je vous invite à poser vos questions en français et j'aurai le plaisir de les traduire pour vous.

Before we get started, I wanted to take just a quick moment to acknowledge that I have the honor and privilege of living, working, and playing on the lands and waters in the unceded traditional territories of the xʷməθkʷəy̓əm (Musqueam), Sḵwx̱wú7mesh (Squamish) and səlilwətaɬ (Tsleil-Waututh) nations. I'm grateful as a settler and colonizer for the stewardship that they've had over this wonderful region. As we explore this topic of artificial intelligence in public health, I encourage everyone to consider the biases that can be present in training data that we use for AI models. And as we approach Truth and Reconciliation Day next week, take a moment to engage with and reflect on the first peoples of this land and their stories and history.

So as part of our work today, we would like to get to know you a little bit better. And so I invite you all to join us on Slido here. So you can use your phones to scan the QR code or a browser to go to Slido.com and use the hashtag AI4PH1. And we'd like to get to know you a little bit better, to see what's your experience using AI in public health is.

While you're filling that information out, I'll just give you a little bit of an update on or a little appearance of what our agenda is today. First, after you're done hearing me talk, you'll hear some introductory remarks from Zoe Kahn, the Director General of the Center for Emergency Preparedness of the Regulatory, Operations and Emergency Management Branch here at the Public Health Agency of Canada. And following that, I'Rosella, a look at– we'll take a look at some of the results of the survey together, and then we'll speak to our

main event, Dr. Laura Rosella who will introduce us to the core concepts of AI in public health. And following Dr. Rosella's presentation, I'm pleased to moderate a Q&A session. In order to ask your questions, you'll be directed back to the Slido, where we will open the Q&As for you to be able to submit your questions. I encourage everyone to look at the questions that are posted as well, and upvote any that you find interesting or that you'd like to answer. And lastly, I just want to acknowledge, too, that our team is working really hard to ensure that upcoming sessions in this series are able to accommodate as many participants as possible to ensure the best possible learning experience for you.

So while you are all filling out your surveys, I see a number of votes coming in, which is great. I'm happy to see that. I'd like to invite Zoe Kahn, our Director General here at the Center for Emergency Preparedness, to make a few introductory remarks to set the stage for today's session. Zoe.

Zoe Kahn:

Thanks, Marc-André, and good morning and good afternoon, depending where you are. Thank you for joining us. I would like to start by acknowledging that here in Ottawa, I am honored to be on the traditional unceded territory of the Algonquin Anishinaabe people who have lived on this land since time immemorial. So on behalf of the Public Health Agency (of Canada) and the Office of Public Health Field Services and Training I'm very pleased to welcome you to this new seminar series "From Pilot to Practice: Applying AI in Public Health", which we are delivering in partnership with AI for Public Health and PHAC's AI hub. At PHAC, we have hosted AI focus lectures before, but what makes this series different is its practical focus. It's about showcasing real world applications of AI and using these examples to talk about what it takes to apply AI in real public health settings today. The goal of the series then is twofold. First, it's to build understanding and enthusiasm about how AI can help address complex public health challenges. And then secondly, is to prepare leaders and practitioners to set AI projects up for success within their own organizations.

We know that surveys of field epidemiologists about their use of AI at work– use of AI at work– the training they most want is practical. What use cases does AI have in public health? What AI tools are best for specific tasks? They want to understand how to use the technology ethically and to learn from others about the pitfalls, mistakes, pros, cons, do's, and don'ts of applying AI in our public health context.

So, based on this feedback, our series is designed to move deliberately from theory to practice, from understanding core concepts and methods to hearing directly from teams who have implemented AI projects in the field. We know that many beyond field staff share these same training needs, so this series was designed for a broad and critical audience, including public health professionals in all levels of government, PHAC field staff who are embedded across provincial and territorial systems and trainees, researchers, and collaborators shaping the next generation of public health innovation. Through this pilot to practice approach, over the next few months, you will hear about applied AI methods from AI for health experts; you'll hear about real pilot projects developed through the AI for Public Health Internship Program that illustrate the full journey from identifying the

problem to building and implementing solutions. And importantly, you'll hear about the real world constraints and the lessons learned. So that includes data challenges, governance, privacy, ethics, and organizational readiness.

You'll also hear that AI must be used ethically in ways that address existing health disparities rather than deepening them. Topics of equity and responsible use will be woven throughout every session as fundamental considerations for applying AI in public health work.

For OPHFST and AI for Public Health, our collaborative work goes beyond a speaker series. Both groups have mandates to build capacity across the public health workforce, and we hope to continue our collaboration beyond the speaker series to bring you new training opportunities in the future. But through this series, we also want to help you connect to each other to find practical opportunities to collaborate. So during each session, a link to Slido will be posted, which will allow you to create a post to reach out to others on this webinar, maybe to find someone with the expertise to solve an AI problem you're currently experiencing, or to find collaborators for projects that you're putting together, or even to signal other opportunities to work or learn from each other. At the end of each session, we'll circulate these posts to session registrants so that you can reach out to others and make these connections as makes sense to you. So in closing, I'd like to thank you for being part of this conversation.

We hope this series equips you with the inspiration and the practical insights needed to bring AI into your work in ways that strengthen public health practice across Canada. And with that, I will turn back to Marc-André to continue with our session. Thank you so much.

Marc-André Bélair:

Thank you so much, Zoe. I really appreciate you taking the time to set the stage for us today. So really happy to do that. Okay. So I hope you've all had a chance to fill out the questionnaire. I'm going to switch screens here quickly, and we're going to take a moment to take a look at those results together. For those of you who are signed on Teams, the links that we reference today can be found in the Q&A tab.

If you click the Q&A section here at the top, you'll notice there is a post from Heather in there, that has the details on the interactive polls, the making connections form that was just referenced, as well as a couple of other things. If you might need to click to expand the post to see all of the content, but it's there, and we'll make sure to circulate the links. They are available up on the AI for PH website. And in other places. Okay. So without much further ado, let's take a quick peek at some of these slide results– –these results. So getting to know you. Let's see who we are, who's with us today. So 40% of– We have about 268 people who've responded. 40% are epidemiologists, 24% are other.

So some data analysts, program and policy analysts, medical officers, public health nurses, health inspectors, and Emergency Management and Preparedness folks. So a pretty great cross-section of individuals, which is lovely to see. These numbers will change a little bit because folks are still answering, which is great. Functional role, the plurality of you are researchers, scientists, or academics, or senior advisors. There's also some frontline

practitioners with us, and a couple of students and some decision makers. Yeah, it's just a great, great mix.

70% of you report that you are not early career researchers, but we do have a good 30% of you, which are, which is fascinating.

And primarily, 93% of our audience is working primarily in Canada, with 7% outside of Canada. This is an open text block. We will have to look at that later individually. And the best statement that describes your current use of AI. So 42% of you report using AI occasionally for specific tasks, with 28% of you regularly using it to improve your work, exploring it: 15% and 7% reporting integrated in some of your recurring workflows, 5% reporting no use of AI tools, and three percent of you saying you're leading or championing AI adoption within your team or organization. So some fairly good familiarity with the tools out there, which is lovely to see.

So 48% of you feeling slightly confident with your understanding of AI and how it can be used in public health, which is a great reason for you to be here today. 25% reporting moderately confident and 20% not confident. So a good, good distribution of folks here. And biggest barrier to using AI in your work. 63% of you report policy, privacy, or governance concerns, some trust in AI outputs. Lack of skills and training, organizational readiness, data quality or access, and five percent saying others. So a lot of concerns there, which is really helpful to see. Okay. So I am going to stop screen sharing right now.

Thank you so much for indulging me the chance to review this screen. Keep this window open on your phone or keep it handy on your desktop, because we will use this for the Q's and A's. Oh, I will turn on the Q&A function while Dr. Rosella is speaking. So you'll be able to post your questions for the Q&A session, and then I will– and please upvote each other's questions. You should be able to do that. And we will tackle that after Dr. Rosella's presentation. And I will make sure to share the link again just before we start the Q&A, in case you've lost it. You can also find it in the inside the Teams Q&A as well. So enough of me.

You're all here for our star attraction here today. So I do have the honor and privilege of introducing Dr. Laura Rosella, who's our speaker for today's session. Dr. Laura Rosella is a professor at the Dalla Lana School of Public Health at the University of Toronto and the Canada Research Chair in Population Health Analytics. Her research focuses on population health analytics, predictive modeling, and the use of data science and machine learning to improve public health and decision making on health system planning. Dr. Rosella leads numerous collaborations with governments, public health organizations, and health systems across Canada to translate advanced analytics into practical solutions for real world public health challenges. She's the Chief Scientist and Stephen Family Research Chair in Community Health at the Institute for Better Health at Trillium Health Partners, and the founder of the Population Health Analytics Lab at the Dalla Lana School of Public Health. She's the nominated principal investigator of Artificial Intelligence for Public Health or AI4PH, a national health research training platform funded by the Canadian Institutes of Health Research. Dr. Rosella, thank you so much for joining us today, and over to you.

Dr. Laura Rosella:

Well, thank you so much, Marc-André, for that really kind introduction and for inviting me to be here. So I'm very excited about this whole speaker series, actually, but starting with today, and I'm so thrilled to see so many of you turn up, and I see some familiar names of colleagues in the chat. So it's really nice to be here with you all. Just making sure before I kick in, you can see my screen.

Marc-André Bélair:

(Marc-André: I can see it loud and clear, yes.)

Dr. Laura Rosella:

Okay, great. So I'm going to jump right ahead. So I am joining you from the University of Toronto. So here in Toronto, I'd like to acknowledge the land from which the University of Toronto operates. For thousands of years, it has been the traditional land of the Wendat Nation, the Seneca, and the Mississaugas of the Credit. And today this place is still the home to many indigenous people from across Turtle Island.

And we are grateful to have the opportunity to work on this land. And as Marc-André said so nicely in the beginning, our topic today is AI for public health. And so I think it's a great point for us to ask of us what this means in the context of data, who owns data, whose knowledge informs our work, who has authority over the data, and how do communities and particularly indigenous communities, participate in these discussions that we're having. I also want to acknowledge there are a tremendous number of indigenous scholars and leaders that are really driving the AI conversation from the indigenous perspective, and it's a wonderful opportunity for everyone to actually engage and listen to what's happening in those spaces.

So today, I have the privilege of being the first one speaking in the series, so I am going to do some foundation and stage-setting for the rest of the things that you will see in the rest of the speakers coming up, um, which I think are fantastic. I'm gonna talk about really the concepts because we can't really talk about the practical parts unless we actually understand conceptually what's underneath it, very briefly. Then most importantly, which I know people are really interested in how is AI being used and how is it potentially adding value? We'll have a broader discussion on impacts. I want to signal that we could have a whole hour and a half on any of those sections in terms of health, equity, environmental impacts, but I will touch on a few. And then lastly, what does it take to implement, which I think is the spirit behind this series, practically, what does this mean for public health, and how do we move forward?

So I start any discussion about AI and public health with this slide because I think it can be very tempting when we're talking about a transformational technology like AI to go into the technology, but actually this is about the impact and making sure that we're focusing our attention on the impact for public health. In the simplest form, if we know– how will we know if AI advances are having a positive impact for public health. And if it is positive, it means fewer people are developing diseases or poor health outcomes that are preventable. We're promoting health in our environments and in our communities, and we're reducing health inequities. And if it's not, we have to ask ourselves, "Why?" And in particular, we

have to focus our attention. There's a lot of distraction and noise in this space. And I think one of the most important jobs of anyone interested in this area is actually, to reduce the noise and to focus our intention on what's going to bring us towards these outcomes.

[This slide has a diagram showing essential public health functions, including health promotion, disease prevention, protection, surveillance, emergency preparedness, policy, workforce, and knowledge functions.] And just to step back, of course, if we think about this from the broader perspective, where is AI fitting into the essential public health functions?

And some of those functions are around our particular areas that we work in, from health promotion all the way to emergency preparedness, but it's also around our policy and program work, our workforce, our evidence and knowledge synthesis. So there are many opportunities for AI to contribute. There's also some where we should maybe not be jumping to AI, and we'll talk about telling the difference, and we have to understand the full impact of that. So hopefully after today, you'll have a very good sense of that.

[This slide has a timeline illustrating major developments in artificial intelligence from the 1950s to the 1990s, including the Turing Test, ELIZA, expert systems, AI winters, and Deep Blue.] Before I get into some of the foundational pieces, I think the history is important, and some of you may know the history. I saw there were some sort of data scientists experts. Some may be new to it, but I think a couple key historical points are really important as we enter in sort of our, you know, some of the most transformative technology we'll see in our lifetime. So, first of all, AI is not new. It actually started in terms of its methodological foundations in the 50s. We have Alan Turing's seminal paper, which is "Can Machines Think?" That's kind of what he wrote about. And there was a very famous meeting around this, and there were some early developments. There was a lot of optimism early on.

So again, this may feel familiar. This is gonna solve a lot of problems, this can do all these things we couldn't do before. And there were lots of different versions. There were some health applications as early as the seventies. Not the best ones. There was a rule-based MYCIN technology made to identify bacterial infections. It wasn't very good. And then there were a few things that actually weren't great. And so there were actually two what we call AI winters.

And again, important to understand what a winter is, because what this– what the winters effectively were, were a lot of enthusiasm around something, and then a lot of critique and maybe some failures on how good can it actually be. And it kind of quieted the field down for a long time. And its rebirth came in the 90s, and it was really around games. So we started seeing computers winning games and winning chess competitions, and then it started to revive again.

[This slide has a timeline showing AI developments from 2000 to today, including ImageNet, deep learning breakthroughs, AlphaGo, ChatGPT, AlphaFold, and multimodal AI.] The AI that we're talking about today was largely defined after 2010, and I'll talk about why, but we started to see this resurgent. We had increased compute capacity, increased.

Digital data. And in 2010, we had deep learning breakthroughs. This is, of course, why Jeff Hinton won the Nobel Prize for deep learning advancements. And the big one being around images. There was a lot of competitions around can you identify correctly classifying an image using a computer? And that happened. And that really gave rise to a lot of technical pieces. And of course, 2020, we had Alpha Fold, which is determining the protein structure, which really transformed basic science and AI. Of course, we had chat GPT, which transformed the way people interact with AI, and now we're talking about agents, which I'll talk to you in an autonomous– generalized intelligence, which I'll talk about in a moment.

But again, the point of me showing you this history is to say it's actually not new. It's been around for a long time, and there have been waves of activity. The technical breakthroughs past 2010 have really enabled technology that we see today.

[This slide shows a graph from the Stanford AI Index showing rapid improvement of AI systems across multiple benchmark tasks, with many systems now reaching or exceeding human performance.] And every year I really look forward to this report. If you're interested in this space, it's a Stanford AI puts out the benchmarking report. So AI progress up to now has really been focused on beating benchmarks, a human benchmark. So you can think of something like handwriting recognition. How well can a human recognize handwriting compared to a computer? The thing I want you to focus on this slide comes out every year.

So you see that human baseline at the front is first of all the shape of those lines. You see them going like this. That they the progress of that benchmarking is extremely fast. This has changed. A lot since the early days. I've looked at graphs since the nineties in this space, where you, you know, it takes 10 years for computers to be as good as humans on handwriting recognition. So that pace is a dizzying, actually. So just important for us to keep in mind. Secondly, the tasks are a lot more complex.

And I'll talk about this later on. They're not, we're not just recognizing patterns and doing predictions anymore. It's actually doing PhD level mathematics and reasoning, which is something that people thought AI could never do around reasoning. Reasoning being one of the most, if not the most complex human task. And we're seeing a lot of PHD level science and reasoning capability as of 2026. So the capability is there, and it's getting faster.

[Nested diagram showing the relationship between artificial intelligence, machine learning, and deep learning, with deep learning as a subset of machine learning and machine learning as a subset of AI.] Another piece around just the round foundations, I've already thrown these terms around. And I do not like giving a definition for AI because in some ways it doesn't really have one definition. It is really an umbrella term. When I say AI, I'm talking about a broad category of methods that really is about using computational approaches to something that we would typically think of requiring human intelligence. Now, again, this could be very simple to very complex.

Machine learning very specifically, which I'm gonna talk about, is really about learning patterns from data and deep learning, which you saw in the historical timelines, accelerated it so that we could use complex data. And I won't give you a technical lecture on this, but just to say deep learning, the way it works, is about breaking up complex data into

smaller components so we can handle it computationally and solve problems that were not solvable. Before that's why we had the breakthrough. So image being one of them, it was very hard for a computer to analyze an image when you turn it into a convolutional neural network and break it up. It's possible. I'm simplifying, of course. And it's not just images, it's text, it's sound, it's language, which is the underlying technology behind a lot of our GPT, Claude, Large Language Models.

I do need to talk about generative AI just for one moment, just to give everyone a grounding, because it's probably, if I did a poll, the most popular AI use, not necessarily in work, although it could be in work, but most people interacting with some form of AI every day. The difference with generative AI is that this is about taking patterns from existing data that we know, but actually creating new content, new language, which was how most people use. These models. It could be also music, if you've ever heard AI music, it could also be sound and images, if you've ever worked with AI images. So this is again a bit of a shift, a different category. And to bring it to our public health discussion, people are using this to create reports, maybe guidance, adapting guidance for different audiences. It does require checking. There's lots of issues with generative AI, and again, we'll get to some of this in some of our more detailed sessions that are coming up. There's hallucinations, there's confabulations.

Now the models are always improving, but lots of things to consider here.

And then just a couple more on the technical bits before we get into the examples. I do want to emphasize that when you work in this space, that the models don't just appear, right? There's some really, really important steps. And the reason so the models have to be trained, they have to be validated. There's really rigorous approaches around this. There are standards now. When I started in this space, I've been in the space for about nine years, almost 10 years, believe it or not, been working with.

Machine learning, and there weren't a lot of standard. Now there's a lot more standards how we think of good validation and testing. What I want to emphasize on this slide, though, is where I think there needs to be a strong focus on public health, and by the way, where public health has a lot of expertise. So the idea of how are we evaluating these in a real world setting so that we know they're improving outcomes. Those three outcomes I said at the beginning, are we improving health outcomes? How are we monitoring the models change and evolve over time really tricky compared to saying something that's just a static intervention that doesn't change and we see what happens. You actually need ongoing monitoring. So I would– I'm gonna come to this at the end, but one of the gaps is actually thinking about how we in our public health institutions and our settings embed real world evaluation and monitoring, which is the hardest thing we need to do right now.

And then to give you a couple little more details for those of you that are like, this is still too high level for me, just to give you some examples of how it's being used. And I'm going to go into some very specific examples in just a couple of minutes, but we have things like clustering, which is about finding groupings in data, and that's used a lot in the surveillance realm. We have, I already mentioned, deep learning in complex data. One of the places I want to call out are environmental images. So, for example, we AI for Public health actually

ran a summer institute on AI and imaging. We weren't talking about imaging of chest X rays. We were actually talking about imaging of the built environment, for example, or environmental imaging. And I'll give you one of these examples later on. And then, of course, we have the large language models, which is the fundamental basis around generative AI and increasingly something called multimodal models.

And essentially what that means is not just taking data from one data source, but taking data from multiple. data streams at once. Technically, that's a bit more challenging, but we're now, if you think about that capability slide, at the point where this is becoming less of a technical barrier and is going to be the gold standards of all modern AI. It's all going to be released on multimodal data.

And if we just think about the future a little bit, just to get you thinking, we could talk about these quite a bit. Where are we going? And again, that's why I like showing those benchmark trends right up at the beginning. We're going to something called foundation models, which are moving away from very specific models to specific tasks to general purpose models. Again, these models underlie a lot of the GenAI technology. They're called pre-trained transformer models. That's what. GPT stands for, and that they can do lots of different things real time, we're moving away from just one model one time, but actually, how does the model keep feeding information and then keep retraining itself, right? Because the model should take the experience of it being used in practice to improve.

Multimodal, I already mentioned agentic, which is a big change, a big focus. Right now, which is moving away from I'm going to use one model for one thing to actually a system that plans and carries out tools. So it predicts something, it does an action, and it completes the whole piece, the whole pipeline. And we'll talk about that. And then the last one, which lots of philosophical debates on if and when this is going to happen, but the concept of our generalized intent intelligence across many tasks.

So before I go into the examples, I've given you sort of that foundational piece. That's the most technical bit of the talk. So if you felt like that was a lot, it probably was, but we've established our technical foundation, not just for today, but actually for all the speaker series to come. So just so you have a bit more of a grasp that AI isn't just this concept. These are actually tangible things, and you know what they are. But that the very simple conceptual level, I want you to think of AI as doing one of these things. It's– and if you instead less think about what the model is underneath, but more what the model is being used for, it becomes a lot more practical. So descriptive, predictive, causal, or prescriptive, what does that mean? It's often being used just to tell you what's happening.

Or it can be used to tell you what is going to happen.

It in very specific cases can be– can tell you what would change the outcome, but that's a lot harder. And then in the very prescriptive cases, actually filling out the decision workflow, which again, very rare. So we have lots in the first two buckets, very little in the bottom two, which I would argue is appropriate, given the fact that we want whole human oversight over most of the applications that we would think about in public health. So lots of surveillance in the predictive, you might be using this for prioritization and planning. In

the causal, this is about specific interventions or policies, and then the recommended and automated systems again being the lowest frequency right now.

[This slide shows a diagram showing where AI may support public health organizations: public-facing services, analysis and decision support, and internal operations.] So let's get into some examples. because this will give you a good sense of how AI is actually being used and what we can learn from some of these examples. And these examples are all public examples where you can read more about them, which is why I chose them, and I'll draw some pieces around this.

So, the one thing before I give you the examples I want to acknowledge, again, I said at the beginning, there's a lot of noise in this space. AI can be used in lots of different ways. And there's a big gap between AI that's being used internally for operations. So I'm sure your organizations or your institutions and yourself are thinking about applying AI for operational workflows and document processing and administrative. Then there's the piece where in public health, the idea that we're doing surveillance, maybe it's automating some evidence synthesis, it's helping with risk assessments, using more complex methods. And then there's that public facing we see in some organizations, chatbots, giving information to community. Obviously, these have completely different risk profiles, but I want to acknowledge that it sits in all of these settings, and examples can fit in all these settings. And if you're thinking about AI strategically in the public health organization, you're probably thinking of examples actually in all of them.

So let me talk about the first example. So this example is really about looking at foodborne illness and turning free text from food histories into a standardized food outbreak analysis. So very, very practical example. A language model classifies those food categories, which are unstructured, into categories which can theoretically more quickly identify exposure information at scale. And there's probably lots of people here that Do this work and time is of the essence. When there is an outbreak, you want to know as soon as possible the source of that outbreak and or narrow it down to enable an investigation. The investigation will always be required. And this allows that task to happen in a much quicker way, and using the free text observations. This is a useful example.

It's very complicated in terms of the model, but it's actually descriptive and surveillance. So again, we can think about the complexity of the technology, but what this example is doing is actually describing something that's really used for surveillance, and there's probably action that will follow. This is a useful one because it's easy to locate, right? There's lots of text to organize. There's a very clear defined classification task. There are people, public health officials that you will use this data and validate it, and we can clearly evaluate whether it's working or not.

So a nice example of an application of AI in a public health context.

Next one again, very simple example, but one that again, very practical, right? So here there's something called med coder, which is used by the CDC to convert causes of death recorded on death certificates into standardized codes. Anyone that works with death data, I know I have for many years. It's very delayed. There's very manual processes around this. And using machine learning, language processing, they were able to code nearly 90% of

records automatically. Some couldn't be coded and they had to would have to go in and be done. But this is, again, descriptive. It's often used for surveillance, priority setting.

And this can happen quicker; so all the other next steps can happen quicker. That one of the examples that's used on this one is around monitoring opioid overdose deaths and where that can be localized and identified quicker to inform public health action.

I'll give you a couple more very quick examples. [This slide has charts comparing a machine learning model and traditional approaches for predicting COVID-19 hospitalization risk, showing improved identification of high-risk individuals.] This is a COVID example, which is around [how] can we estimate the risk of hospitalization for those that have tested positive with COVID for the purpose of vaccine priority? So we know that we did vaccine prioritization based on the things we know that are important, age and comorbidities and some other factors as well were used in that, but they were sort of these simple rule based systems– Where in this study, we were able to show that using existing data, no requiring no data collection, we can do it a lot more accurately and faster. So the idea that we can identify these individuals quickly could then again feed into processes in real time.

[This slide shows an example of computer vision identifying potential cooling towers from aerial imagery to support Legionella investigations.] This is one that's really interesting. It's gotten a lot of attention, very popular. Again, deep learning, computer vision methods. So very sophisticated methods. Again, many of you in the call, I'm sure, have been dealing with the Legionella disease outbreaks, which have been happening across Canada the last couple of years. Typically, what's required, and it was really helpful, someone that works on this in a local public health unit was telling me about this that they have to a lot of these is cooling towers and that someone would be driving around or looking up, "Where are the cooling towers?" and then doing subsequent investigation. And so here, computer vision can search aerial imagery in moments and identify the towers, and then narrow the search. So the actual action of public health is not really changing much here. This investigation still needs to happen.

They are just narrowing and making sure that that action can happen quickly and more precisely, using advances in machine learning. And they made some really nice open available methods to do this.

[This slide has a series of graphs of AI classifying social media posts according to positive, negative, or neutral attitudes toward vaccination.] And last example, I wanted to just highlight [an] example of using language models, looking at public discourse. This is from Laura Espinosa's team, and what they were doing is using publicly available language model to classify whether social media points to a positive or negative or neutral stance towards vaccination. This type of discourse analysis is very common in public health, and there's been lots of advances, again, to be used to prioritize efforts. And this is also highlighting a limitation, of course, because there are actually, with all of these, I'm sure– We could have a very robust discussion on the strengths and limitations, but wanting to give you a sense that these pre-trained models offer a really scalable way to do this. Before this would happen… This work's been happening for public health for many years, but it

would involve a lot of manual labeling of posts. And the way that we do this in the more traditional sentiment analysis.

Okay, so what do these examples teach you? I gave you a range of examples, identifying in these examples that for the most part, we're seeing AI structure information, identify patterns, and support prioritization, but each of these needed a connection to a public health action. In some way, it was more clear than others. I would say that starting with that point is really important. So what public health action can we enable?

Or what workflow can we enable? And then going for what are the possibilities in these tools. I think one of the pieces, I didn't talk about the evaluation of all these models, but the way we should be thinking about evaluations, like, are these good? Are these valuable? Are they not valuable? It's above and beyond performance. So I think that performance has dominated the literature for many, many years in this space. And what we're really looking for now is, is it actually changing outcomes?

Is it impacting workload? Is the goal, ideally avoiding a preventable illness or whatever, the goal is for all those tasks that we saw happening quicker, more effectively? Is it reaching more people? Is it reaching the right people? That's the way we need to be thinking about these evaluations. And of course, there's benefits and harms across all of these that need to be considered, and I'll talk about that in just a moment. And that needs to be worked into our analysis.

So before I get into some of those challenges, I think again, I've said this a few times, for us in public health, the importance of us sort of changing the conversation about we should use AI or interested in using more AI to having some version of this that you could fill out. So as a public health decision maker, if I knew X, which the AI is helping me get faster, easier, et cetera, I would do Y to improve the measurable outcome. And this is really important because in AI, we have a lot of capability. That's why I started with that. Amazing capability that I never thought we'd achieve by 2026. We are achieving it. But what can we actually do with that information? And is it improving an outcome? And to be clear, I do think there are very good use cases where that would be the case.

But right now, we're not as clear as possible as we can be with on that. And I think the quicker we get there, the more we'll start actually focusing our attention on the value versus any AI application, which we could add. And by the way, I do, I do include some of the administrative tasks in there, which I know are super important to day to day work and workload. And we have to understand which one of those is actually also related to the goals of public health that we're trying to achieve.

So let's talk about some of the impacts. Okay, so right now I've given you a sort of foundation of what AI is, how fast it's moving, what's coming. I've showed you some examples of how it's actually being used in the public health context. I want to talk about some of the impacts, which again, we're going to get into in lots of ways in the Speaker Series. And specifically, I want to talk to you about data.

I want to talk to you about the impacts in terms of an equity perspective, and I want to talk to you about environmental impacts. And all of these could have their own session, but

we're going to talk about those, just briefly introduce them, and then we'll talk about how we think about translating that into some sort of responsible implementation.

Public health has been using data for at least six hundred years. So this is not a new challenge for us. The idea that we are constantly struggling with data quality, representation, and whether or not it's fit for our task. And in AI, this is amplified quite a bit.

But one of the benefits of AI is that we can work with new data we've never had before. So if we take, for example, the imaging data that's identifying the cooling towers, we've never been able to do it that scale and that speed before. On the other hand, we have to be critical about a lot of things. And again, this is where I'm gonna call on public health because we have really, really good expertise on this. So if we think about who sits around the table in advancing AI applications, public health must be there. Who better to ask about data representation and quality and what you can do with data than public health? I truly believe beyond any sector in the world, this is the sector that's most poised to answer that question.

Some of the questions we want to ask ourselves, will using this data allow for new information? Okay, really important because some data is really cool and it's really impressive that we can use AI to process it, but is it giving me something new? Is it giving me an insight I didn't have before? And again, I'm gonna come back to this point because it's really important for the environmental impacts. Who? Does this data represent? So the selection biases in this data are something that are significant. Anyone that works in public health, epidemiologists doing this work, it's almost the first question you're asking because you know, even if it's a prevalence estimate, not a complicated model, you know it doesn't represent everyone. Who does it represent?

Who does it not, and how does that impact the public health decision?

Sometimes it's not new information, but it's faster, which is a good thing if faster translates into better outcome. If it's just faster, but then you have to plug that into a workflow that's still moving at a snail's pace, it may not even help. And it may actually stall and jam up the system even more. But if it's faster, and mean faster means we can do something more quickly and reach more people or reach the right people in the right way, or use a resource that was being used on that data before to make the workflow work better, then it's valuable.

And then the last one's really important because I've seen this time and time again in public health. We've seen this in our AI for public health platform projects and our initiatives. We have amazing scholars and interns working on really innovative applications that are really well done, but then how is that maintained beyond that one project? And we saw this, and I'm gonna talk about wastewater surveillance early on, some of those initial innovations were done in in an academic setting. So then how do we actually make sure it gets out of the university project and into standard operations for public health? Really important for us to think about that earlier rather than later. Otherwise, we're going to develop a solution which we can't maintain or scale, which is a problem for public health.

So if we think about some emerging data sources, I put two here as an example. Wastewater data, I mentioned, came up in COVID. Really positive pieces about it. It was timely. We were getting more population coverage, not 100%, but populations that are connected to sewer systems. We had fewer privacy or legal risk, not none, but much fewer, because it was a population estimate of infection. And it was very actionable. So you saw those signals during COVID, we've also seen them now since for broader respiratory surveillance, and it's being incorporated into broader surveillance efforts, which of course are used. But is the pipeline reproducible?

Where does it exist? Who's constantly validating these models to make sure they exist? Depending how you sample which– wastewater, you have to consider that in the model. The different sewer systems have different pieces. So it can actually be a bit more complex, but we can see that we're actually getting new information, which could make a difference. If we look at social media data, it's a bit trickier, right? We know it's timely, which is good. It's more cost effective than trying to get those sentiments one by one. We're definitely getting some new information, new perspectives that we probably weren't seeing before, but there's a lot of selection bias in this data.

There are constantly questions about how appropriate it is to use that data, and it can be difficult to action. So we can know that certain trends are coming up– But what do we actually do with it? It's not impossible, but I just present this as an idea every time we're thinking about these starting these new examples, we need to think critically about what it's adding and our ability to maintain it. And I believe the more we do this, and the more we focus on those efforts and those data sources that are really bringing the value, we'll start seeing that value very quickly, which might be a bit diluted right now.

[This slide provides an illustration showing how biases can be introduced throughout the AI lifecycle, from problem selection and data collection through deployment and use.] I could spend hours, many hours talking about one of the biggest challenges with AI, which is related to algorithmic bias and specifically how those biases can impact equity. And again, this is gonna, as was said at the beginning, this is going to come up in every section. So what I'll say here for the interest of time it is a few things. So obviously the biases in the data can propagate through the algorithms and then those can bias decisions. And that's a huge problem. And it happens with and without AI just in general and how we use data, and we're very aware of that in public health. Now, how do you mitigate that? What are the solutions around that? And that's where the conversation has really shifted.

Again, in the early 2020, 2019, we were talking a lot about just raising the alarm bells that this bias is important and we need to consider it. And we need to continue to do that. Now it's much more around solutioning around it. What are some of the methods, de-biasing methods? How do we make sure through entire pipeline that we're considering this. And it starts with, and I really like this framework, anytime I talk about this, it starts in the whole pipeline. So what problems are we even choosing to work on? Why are we choosing certain problems over the other? It's as fundamental as that.

When we collect the data or we use existing data, where are we looking at biases in that data? How are we defining outcomes? What outcomes are important? How are we

developing the model? There are practices, actually, that can increase the risk of bias if you don't address that in. And there's ways of looking at subgroup performance, making sure that you have always multi-site validation. There are a number of practices that I don't have time to get into that actually are the gold standards. And even if all that works very well, once the model is used, how it's being used can also cause bias. So let's just say we have a model that's actually really positive.

It's very effective. It works well across all subgroups at identifying people that have a certain need that would benefit from an intervention that we have. So it's sort of the best case scenario. Well, if that intervention is, if that model is only being used on a subset, or it's only being used by people that know about it or can afford it… Then we're going to start seeing bias. Even though the bias is not necessarily in the model, it's how that model is propagating bias. So a very complex topic, but what I'll say is that, the role of public health in solutioning this topic, including when we shouldn't be using models, is critical. And we have a lot of expertise to bring to this.

And so that brings me to the third kind of impact area I wanted to mention, which is around the environmental footprint. And so in AI, we have a very complex tension to navigate. Actually, the world has it, but I would say in AI, we in public health, we think about it even more. So we know that AI has an environmental footprint, and the footprint is not necessarily just one thing. It's around the hardware production and all of the material extraction, rare earth metals, mining that has to be done to create all the materials that go into all this computing. How we train and use model requires electricity. It requires GPUs. We have data centers, which I know is a very live discussion now in terms of the infrastructure. Data centers need to be cooled.

And the more we use them, the more environmental impact on electricity, water use, and there's also local impacts. There's some great work just recently published out of an environmental group in Harvard where they were looking, actually, you can actually detect the CO2 emissions. And in fact, there are carbon estimates now for lots of our AI technology. So we're getting much more sophisticated. And then even after it's all said and done, all of this technology has to be replaced.

And there's e-waste and all of that. So this is a big challenge because in public health, on the one hand, some of these tools can actually help the work that we do from a climate perspective. We can have early warning signals for wildfires. We can identify who's most at risk for climate displacement. There could be lots of efficiencies that we can use that actually have a negative carbon footprint with some of this technology. So there's potential benefits from using this, these technologies in the right way, and then there's harms.

[This slide has a diagram showing strategies to reduce AI environmental impacts, including transfer learning, pruning and quantization, and smaller task-specific models.] So what do we do about it? It's quite the conundrum for it, for us.

I'm going to say that we need to think about it in three ways. So, first of all, just knowing that there are things we can do. One of them is not using AI for certain things. And that's why I keep reinforcing the fact that we have to drive home what are the most valuable applications for a public health outcome and focus on those, not think about it as a

generalized technology that we apply everywhere. But there are things, and I don't want to get into too much detail on this, but just to say there's things like transfer learning, which is this beautiful concept where you can use a pre-trained model that works for one task and just transfer it to another. Why is that important? All that, we don't, if we don't start every model from scratch, you have a much reduced environmental impact of that tech. And that's been shown again and again empirically. So it's sort of like this needs to be the best practice.

There's things called pruning or quantization, which is about maybe we don't need every single variable, which is maybe hard to believe, but actually doing that will have a reduced carbon footprint. And then the idea that maybe we don't use these generalized models that have much more of a carbon footprint, and we do more smaller models for targeted tasks. So not saying that all of these can or should be done, but there are ways to actually reduce this. Same with our workflow. We don't have to retrain every model.

We should have a robust evaluation and method to be able to understand what models work in what context. There's lots of work around data centers, and then most importantly, avoiding unnecessary use, which I mentioned. And then lastly, I'll say in our governance. So AI governance, which we talk a lot about, we often think about data and privacy, of course, first. We think about ethics and bias and equity, and we have to now start thinking about environmental impacts.

Like we do for a lot of the other work. And so we really get into have we thought about is this the most environmentally sustainable solution compared to others? And there are ways to actually do this. And in our course, which I talk about at the end, we go into some of those calculations. Just to give you a sense that there's more tangible ways to do this.

Okay, so I'm turning towards the last part of the presentation, which is about what this is actually, what do we need in public health? So I've told you there's all this capability, I've given you some examples, I've given you some things that we need to really think about from a critical lens on the impacts of AI. What does successful deployment actually look like? And it really is about…

Which I'm reiterating several times on purpose. First and foremost, aligning with public health goals. But we also need to make sure it's effective and safe and ethical. [This slide has a framework describing requirements for successful AI deployment: alignment with public health goals, effectiveness, safety, ethics, equity, privacy, and sustainability.]

And I want to talk a little bit about how we make sure that happens with the caveat that this is not every individual's responsibility. And there's some structures in place that we need to establish that. So back in 2022, we wrote this paper, which is talking about what priority areas do public health organizations need to do [for] this advancement into practice. And you'll see some things here that I've mentioned one way or another, except for workforce and training and capacity building, as well as partnerships, which I think are really key. And if we have time, I'll talk about those at the end. And there's been some thinking since then.

I would add to that, it's not just infrastructure to do modeling. I think a lot of people think like, oh, I need the infrastructure to run the models. There's actually infrastructure to do the testing and monitoring as well. There's a life after these models get deployed. And the good AI practices, there's, as I mentioned, a lot of standards now, but we also need time for the humans to actually critically assess those, that time needs to be built in and maybe structured in a formal way. So this still remains relevant with some newer editions.

[This slide has a diagram showing responsible AI implementation supported by engagement, oversight, regulation, guidelines, and standards.] At the crux of it all, I see these two bookends as being crucial. On the one hand, of course, there's going to be regulation and guidelines and standards. Those are evolving every day.

We cannot rely only on those, although they're important and we're a bit behind in Canada on some of these, but lots of people are working on it so I'm confident we're catching up soon. On the other hand, it's engagement. Engagement with people that are affected by this technology, the practitioners that use it, the operators of the health system in an authentic way. And of course, this is how we operate in public health. We have very strong foundations in community based participatory methods, for example. This should also be what we apply as we think about AI.

And so there's individuals like everyone on this call educating, learning more, understanding how this works, organizations that have practices and policies, and higher-level regulation. [Swiss cheese model illustrating multiple layers of protection for responsible AI use, including regulation, standards, monitoring, and oversight.] And I, you know, public health loves a Swiss cheese model. This is how I think about it, because individuals have responsibility. We have some policy and regulation and standards, and we also need monitoring and oversight, as well as individual responsibility.

No one layer, for example regulation, will take away all the risks that we talked about.

I think in the interest of time, I'll pass some of the frameworks that we have just to say there are a lot of wonderful frameworks around ethics, equity, and AI, especially in the last couple of years. So we have a great starting point. [Example of a participatory AI project for diabetes prevention showing collaboration among researchers, practitioners, and community partners.]

And just to end on a very practical example of a project that we've been working on to do deployment of AI related to diabetes risk prediction. And what I want to emphasize in this project is that the core of this project is based on engagement. So the technology is there. It's done. We keep improving the model, sort of a side part. But what we're doing is really purposeful engagements on what how the model is being used, how it's being presented, how it influences policy. And it's been a tremendous experience.

And there's some papers there you can read on how that actually changes what we do. And all of the responsibility pieces of AI come along with that, because all those issues are surfaced.

[This slide shows a lifecycle diagram showing stages of responsible AI deployment: define, develop, prepare, evaluate, monitor, and adapt.] And I'll leave you just with the lifecycle of deployment, which is, again, I don't want you to leave here thinking, oh, it just ends. We just figure out what AI we're gonna use, we're gonna put it in, and then we're gonna use it. There is a cycle here because once the AI goes into practice, after all that really careful co-design and engagement, you need to continue that engagement all the way through. We need to be much more rigorous about evaluating the outcomes, ideally in a prospective way, which is starting to happen, but it's very slow. And then that monitoring has to feed back into either this problem or other problems.

This is just my last slide before I'll just close to say that evaluation piece is something that I think is our next priority for public health.

Not just thinking, are we adopting AI, but how are we showing AI is adding value? And when you're asking that question, you really are needing to be able to report on, yes, the performance does it work, but how is it changing our practice? What's the public health benefit and equity and sustainability, which I've talked about. Quite a bit about.

So I'll just close here. I'm really looking forward to your questions to say that we covered foundations, we talked about applications, we talked about impacts, and a little bit, albeit I know we were quick on what responsible implementation actually looks like. The good news is everything I talked about is going to be touched in one way or another in the rest of the series. I've seen what's coming, and in a lot of the ways we'll focus on specific areas in a lot more detail. [Photo of the Introduction to AI for Public Health class] I thought I would just end on, this is my class. And last year at the Dalla Lana School of Public Health, we had our first introduction to AI for public health class, where we dove into all these topics in a lot more detail. Amazing, thoughtful group from all different disciplines coming together to think about this. And we did a little exercise at the end of the class where I said, What do you, what do you, where do you think the possibilities. Are what are you worried about?

And what do we need to do to make sure this goes well for us in public health? And we came up with a class, generated these themes, and this is a little bookmark that we made after to always remind ourselves. And so just wanting to end that there is a lot going on with AI, but the future is not yet written, and it really is ours to shape, especially those of you that have been working in public health for a long time. The sort of invitation to come up to the table and make sure that we define this technology for how we want it to be used, with possibilities and a commitment to the goals of public health, and not that someone is going to define it for us, because I do believe, as I've mentioned several times, that public health has a very unique and important position to make sure that this technology is used for good. So I will end it there. I'm sure there's tons of questions and thoughts, which I would love to hear, thank you so much for having me, and I'll take the slides down now.

Marc-André Bélair:

Wonderful. Thank you so much, Dr. Rosella, for your fascinating and balanced presentation. It was really great. I'm going to quickly give you a chance to catch a breath and we'll share, I'll just share the link to the Slido again for folks to post your questions. There have been

some questions that have been filtering in throughout your presentation, which I'm happy to see. So just a reminder, if you'd like to join us at slido.com and use the code AI4PH1. You can pose some of your questions and upvote each other's questions. So we have about twenty or so minutes where we can tackle some questions. So I'm just gonna jump right into it.

I will stop sharing my slide in like just a minute just to give folks to get connected. But one question that I want to know is that you emphasized starting with a public health problem. How can an organization recognize when AI is worth pursuing and when a simpler approach would serve it better?

Dr. Laura Rosella:

Yeah, really, really good question. And by the way, I want to say that this is where we should be spending the most time. It's the most important decision. And often it, I mean, various things happen. Some health units and organizations are spending a lot of time here. Sometimes someone comes with a technology and says, I think this could really help us. And we're trying to like retrofit it into our system. So I would begin by naming the decisions that you want to improve or the limitations of the current approach. So the way we do X isn't working as good as we could.

It's slow, we're missing groups, and surface those first, and then compare it to the alternative. So what could improve it? And sometimes when you do that, you find that there's lots of other things that could improve it, aside from AI. And then there's even sometimes simpler methods. But if it's very clear that AI can actually make a difference here, then that's the place to start. And the cooling tower example was a nice illustration of that, right? This is how we do it now. It's slow. We're definitely missing places.

When I look on the map, it was there last year and now it's not there anymore. So it was very, the use case was very clear. And then the decision that we change would be very clear. And I think it's also really important to write out who's going to act on it. Like what's the entire workflow? So there's the part where AI plays a role, which generally is very small, and then the entire workflow and really understanding that. And I'll just say this reminds me, from one of our AI for public health mentors, Professor Dan Lizotte, who's going to be moderating, I believe, the next session. He ran a workshop in 2022 And he's a computer scientist by background, but he brought together local public health and then, you know, the best of the best in terms of the AI experts at the time, and said, Okay, local public health, what are your problems? We're listening, as opposed to, and then a brainstorming activity of like, oh, did you know that there's this way to do this?

And here's a method I've been using that might be useful here. So that's another, I think, great place to start that sort of brainstorming work at the local level.

Marc-André Bélair:

Great, thank you so much. So there are a ton of questions pouring in on Slido, which I'm really happy to see. I did note that one question was asking whether this presentation will be shared after the fact. I just wanted to note for everybody that yes, we will be providing

the full video as well. Dr. Rosella has agreed to share their slides. So we'll be able to package all that content, and we're translating it as well en français so it'll be accessible to everyone. Just keep an eye on your emails and on the AI4PH site for more details about that. So just give us a little time to process it all, and we'll get it. So one question coming in here from Slido is what AI models are you currently using in your work and where are they hosted?

Is it locally hosted or with a private organization?

Dr. Laura Rosella:

Yeah, great question. So in my work personally, I do a lot of prediction work and surveillance work. Generally, we're using typical machine learning models, ensemble models, boosted trees, et cetera. They are all hosted locally, for the most part, because I'm building these models based on data that's in a protected environment. And that's been a big challenge. And I know that's a challenge for a lot of organizations, but usually these models are held within a secure local data infrastructure. Some of them have secure cloud-based infrastructure. So not all of them are local, but that cloud infrastructure is owned by an organization that has authority and privacy to personal health information or health information. But it's a good question because this is rapidly changing.

And I'll just say the broader AI conversation in Canada and focus on sovereignty is really focused on how we make sure that infrastructure is sovereign to Canada and secure and protected. So I know a lot of organizations are thinking about this.

Marc-André Bélair:

Great. Thank you so much. Now, I see a question here that many teams across the government of Canada are seeing the move towards agentic AI as the next big shift. What do you think for public servants and maybe for public health professionals who are also public servants, what do you think they need to do to prepare for a world where agentic AI is increasingly integrated into workflows?

Dr. Laura Rosella:

Yes, I love this question. Thank you for asking because I didn't have a ton of time to talk about agentic, but again, just to recap, if people are hearing this term for the first time, agentic is a technology that's not just doing one thing, but it's actually coordinating a series of actions. So it's finding who's at risk, it's booking the appointment, it's finishing the report after it's doing the full end to end. And there's a couple really important points here. So one is obviously there's a workforce implication, and as we get more and more towards AI agents that can effectively do tasks which humans are doing now. There's some questions about what that means. And I think if, you know, I'm thinking about this a lot too. It affects lots of– there's scientist agents out there as well now too that can ask the research question, do the research and write the paper. So all of this is happening all around us.

So I think it's important to be aware of these trends. And this is why I look at the trends reports and what's happening.

To understand where you as a human in the workforce add the most value. And I think it's also important to be clear where the public wants us humans to be involved in the workforce. The one thing I will say about agentic because I'm looking at a– I'm participating in an agentic evaluation now, and I'm looking at a lot of the literature on the evaluation, it's very thin at the moment. And so while I know that this is something on the horizon people have to be aware of, we don't fully understand how these are working in a way that we can just plop them into all our workflows. We need a lot more evaluation and research on the full pipeline because we can say a model is performing well, but the impacts when you have these multi-step processes are quite a bit different. So my advice is, to get involved, to pay attention to how AI agents are evaluated really rigorously from a scientific point of view. And there's some papers on this. A lot of it's in healthcare, not as much in public health, because I think we're quite far from this. And then secondly, really getting focused on where the human aspect of the work needs to be preserved and making sure that's not lost.

Marc-André Bélair:

Yeah, no, that's a really important consideration for sure. I want to talk a little– I want to shift gears a little bit, and I think you alluded to it in your early discussions, I talked a little bit about it in my land acknowledgement but there's this concept that the training data for a lot of AI models is, you know, can be considered culturally homogeneous and the question here in the chat, and I think this is something that I'm particularly interested in is– Are how are the data that's built into training models is representative of the communities that we serve as an organization, not just, you know, Canadians, you know, Canadians at large, but when looking at, for example, Indigenous communities or, you know, non-English speaking communities. How can we examine the biases that exist in, you know, the training data and the data sets? And somebody specifically referenced here ensuring that AI is culturally intelligent, similar to what Tulong Tech are trying to do. I don't know if you're familiar with that, but what are your– some of your thoughts on that?

Dr. Laura Rosella:

I think this is a really important question.

I'll point to a few things. So, first of all, there are lots of really good thought leadership coming from Indigenous scholars and Indigenous leaders. So there is, I'm aware of, there's an Indigenous protocol and artificial intelligence position paper that I know was released in the 2000s. And so there's lots of really good thinking about this and lots of good work from the Indigenous perspective that needs to be recognized, first of all, everyone needs to read them and we need to informally put it in our governance. One thing I'll say, just take a step back. We have OCAP principles. We have care principles. We have– now, and it not to say we have them and we're done, but we know what the problems are with the way we do things now with data. Period. With and without AI.

And so I think the first thing I will say what's most important is making sure we're foregrounding that work and we're making sure how we're thinking about anything that we're doing adheres to those principles. The next thing I'll say, I mentioned it early on, is we need to be thinking about this way earlier, like to the problems that we're choosing. Never mind the specific technology and the data and the bias which come.

But I think backing this right up to saying what are we using AI for? What impacts will it have for these particular problems. And I think a lot more of those earlier conversations is what we need to do. And I think that Indigenous where I think we need to do a much better job at understanding how Indigenous ways of knowing are shaping the discussion. Again I think that there's very good ideas already out there. It being implemented in practice, I don't see as much. So I think that's a really important area for growth. It's part of governance, but it's more fundamental than that. Actually, it's just about like how we're actually approaching this transformation, this change that's happening from the outset.

And you did ask a question about data bias specifically. So let me mention that. How do we know that AI is representative? That has to be part of the standards of developing these AI, which it's slowly becoming, but again, that has I would integrate that into governance. So before deciding if an AI tool or an AI based technology can be used, understanding has it been validated across population groups that we know are important to our community, Who's in this data, who's not in that data? And then depending on the answers to those questions, you may say, well, we can't proceed with this, or we can only proceed with it in a very specific way with all the caveats around this. And we do this already for non-AI based data to some extent. And we need to just make sure that that's translating in. I don't want all this to be about governance, though, because it can't just be governance.

It actually has to be with like how we develop the problems to begin with. How we develop the models. Yes, governance is part of it, but we can't wait for governance to catch all these issues if actually our fundamental way of framing the problem hasn't considered care principles or OCAP. It has to be throughout the lifecycle. Right.

Marc-André Bélair:

No, absolutely. Thank you. And I we we've got way more questions than we'll be able to answer today, unfortunately, which I mean I love to see that, you know, folks are engaged. But if we don't get around to your question today, I apologize. But there's a question here I think is really interesting, about AI and this, you know– this talk about replacing entry level– entry level staff. And this one specifically is about, you know, is AI expected to replace entry level MPH or or you know, MSC public epidemiology type training roles, and how is it gonna shift our responsibilities towards oversight, community engagement, and algorithmic equity? What skills ensure that we remain indispensable?

Dr. Laura Rosella:

Yeah, great question. Again, such a good question. I love this question so much. And this is where, I mean, I'm, of course, biased because I am a public health person, but this is where I think in public health we wrestle with these tensions more than any other sector, right? So, first of all, I will say, that we have done some testing in some of our research projects. We've actually said, okay, we've compared directly in a structured way, if you used AI based, like full analysis for a typical public health problem to human, to some hybrid in between. And I feel very confident that we aren't necessarily just talking about a replacement. We may be shifting our focus, which I think the question is.

getting at where we're spending maybe less time coding, but we're spending more time checking and doing oversight and directing and doing the interpretation. So that's where I– how I think people should see this. The part, the reason why I was saying we wrestle with this tension a lot more is that in public health. Occupation is a determinant of health, right? So we are thinking about not just, oh, is it going to replace a bunch of people, but what happens when people don't have entry level positions? Like, what does that do for population health? So I think we need to engage in this question in a much more meaningful way. For those of you that are training right now, I want to say: You know, getting as much exposure and training as possible so you can be in the conversations. That's why I wanted to end on that.

We can shape what this looks like and how this looks, but you do need a sort of basis of what even is it? And can I get to the table to make sure I'm shaping those conversations? And am I shaping the solutions? And then we need to start seriously thinking about all of our education systems, all of our workforce. Work, what it means, what it means to have experiential learning.

You know, many of you would have seen that in New York they banned AI in elementary school effectively with this idea that we want to make sure kids are doing things first before using AI. And it's kind of a similar thinking, like where are we actually intentionally building in redundancies? Cause if we just let it go, it will take a lot of this out. And we, and people do need to evolve. They need to have those first experiences, struggling with data and understanding what it means.

to be able to evolve. So I think I'm blowing up the question a little bit. I think it's much bigger than what the question is being asked. There's definitely room for entry level. And I think that, frankly, we don't have enough entry level public health people working. Our workforce is so minimal that I don't see that as the immediate problem, but I think the way we think about what it means to be entry level and what it means for the future of work is something that, again, we have to engage in as public health– [as] people working in public health and knowing that it's actually determinant of population health.

Marc-André Bélair:

That's such a great point. Thank you for– I think blowing up that question was absolutely the right way to go here. Unfortunately we've come to the end of our time for answering questions. I could sit here and chat with you probably for hours, so I hope we get the chance to do this again sometime. Thank you so much for taking the time out of your busy schedule to share your knowledge with us and to give us what turned out to be a really solid primer on AI and public health. And I think it's one of those sessions that leads to more questions than answers, but I think that leads to curiosity and discovery, and I'm really excited to see, you know, I'm really looking forward to our future sessions. So I'm gonna take a moment now just to point folks back to the Slido, where I've activated the evaluation survey. I'm gonna share my screen real quick. My screen is still being shared.

Look at that. Just head back to the Slido for our evaluation survey. Here at the Training and Development Unit and at AI4PH, we really value your input and we actively use the results of these evaluations to help improve future sessions. We'll also send the link to the

evaluation as part of the post section's email, along with the link to the connections form. So hopefully folks can share the projects that they're working on or look share the expertise that they're looking for or that they have.

They're willing to share, and we can get folks to connect with each other on different projects. And also just a reminder that we will be sharing a link to the recording of the presentation with transcripts and captions and slides in French and English soon. So please keep an eye out for that. I'm also really excited to share that our next session is going to be the first part of the AI Methods in Public Health. It's scheduled for October 26, 2026 at noon eastern. Dr. Rosella will be joining us again for that, as well as Lisa Fan from the University of Regina, Lisa Lix from the University of Manitoba, and Dan Lizotte from Western University. So please watch your emails and communications for information on how to register for the upcoming session. Again, thank you to Zoe Kahn for your opening remarks and Dr. Rosella for your engaging presentation and for answering our questions here today.

I'd also like to take just a quick moment to thank all the scenes that are all the folks that are behind the scenes who were instrumental in making this happen, the team at AI for public health, as well as the team here in the Training and Development Unit and A/V Services at the Public Health Agency of Canada. And on behalf of the Public Health Agency of Canada and Artificial Intelligence for Public Health, I thank you, our participants, for taking the time to spend an hour and a half with us today. And I really hope we'll see you at future sessions. So I wish everyone a really great rest of your day. And please remember to fill out the evaluation. And thank you so much. And thank you again, Dr. Rosella. I hope you have a wonderful rest of your day, too.

Dr. Laura Rosella:

Thank you so much. Thanks to you and everybody.

Marc-André Bélair:

Bye now.

APPENDIX A – Prompt & Response (DO NOT TRANSLATE)

Prompt

Can you please take this subtitle file and write a script to strip the subtitle-specific items (like the unique index number, the timestamps), and then reformat the text into full sentences. Where possible, please try to insert line breaks and paragraph breaks. DO NOT EDIT THE CONTENT AT ALL. Simply redact SRT specific formatting, and format into a human-readable format.

Additionally, there are three speakers, Marc-André Bélair, Zoe Kahn and Dr. Laura Rosella. They initially speak in that order, but after Zoe is finished her remarks, she no longer speaks. If possible, try to identify the speakers based on the context in the text. If you need to, you can see the final video here: Session #1: Introduction to AI in Public Health

Do you have any questions before you proceed?

Response

No questions needed — the request is clear, so I'll go ahead with these working assumptions (easy to change after): the script does all the mechanical work (strip indices/timestamps, rejoin lines, reconstruct sentences, add paragraph breaks at natural pauses), speaker turns are a small table in the script that I fill in from context in the text, and the output is a Markdown file with bold speaker labels. Text is untouched, including filler words, "--" cuts, and the bracketed editorial insertions.

Fixed-width lookbehind issue — I'll split with a two-branch pattern instead.

Works, word count matches exactly. A couple of paragraphs run long because speakers barely pause between cues; I'll add a hard cap so no paragraph exceeds a readable length.

One bug: "Dr." at the end of a cue is being treated as a sentence end, which split "Dr. / Rosella" across a paragraph. Adding an abbreviation guard.

The only difference is a leading blank line from the title header — the spoken text is word-for-word identical to the source. Here are the two files.

Done. ai4ph-session1-transcript.md is the reflowed transcript: 1,734 cues → 18 speaker turns and 129 paragraphs, 13,670 words in and 13,670 out. I verified the spoken text is word-for-word identical to the SRT (only the cue numbers, timestamps, and line wraps were removed), so fillers, "--" cut-offs, and the bracketed insertions like "[how]" and "[an]" are all preserved.

How srt_to_transcript.py works, so you can reuse or tweak it:

- Parsing/reflow is fully automatic: strips SRT formatting, rejoins wrapped lines, splits into sentences on . ! ? (with a guard so "Dr." and "Mrs." don't end a sentence), and starts a new paragraph at a sentence boundary wherever the speaker paused ≥

0.9 s, with a soft cap of 5 sentences and a hard cap of 9 so nothing runs on. Those thresholds are tunables at the top.

- Speaker turns are a table of cue-index ranges (SPEAKER_TURNS) that I filled in from context — hand-offs like "Zoe.", "Thanks, Marc-André", "over to you", and each Q&A question/answer pair. Edit the table if anything is off; the script just applies it.
- It prints a word-count check on every run so you'll see immediately if a future edit drops or duplicates anything.

Two things to be aware of:

1. **Cue 1579, "Right."** — I attributed it to Dr. Rosella (as the tail of "It has to be throughout the lifecycle. Right.") because the timestamp gap before it is 1 ms, while "No, absolutely. Thank you." clearly starts your reply. If the video shows it was you, change (1510, 1579, ...) to end at 1578 and start your range at 1579. Everything else was unambiguous from context, so I didn't need the video.
2. **"Zoe Kahn" vs "Zoe Khan"** — the source spells her name both ways in the spoken text (cues 45 and 1713 say "Khan", cue 69 says "Kahn"). I left the text untouched as instructed and used "Zoe Kahn" for the label, matching your message. Happy to swap the label if the other spelling is correct.

Run it on other sessions with python3 srt_to_transcript.py input.srt output.md — you'd just need to update the speaker table for each one.