AI in Healthcare: Navigating Risks and Safeguarding Patients

Presenter: Josiah Dykstra, PhD

It is my pleasure to welcome back Dr. Josiah Dykstra. If you'd like to learn more about Dr. Dijkstra, you can read about his bio on the course registration page. But at this time, I'll hand the mic over to you.

Thank you. Thanks for that great introduction. Welcome, everyone. I am pleased to be here. My name is Josiah Dykstra and I'm really excited to be talking with you about AI today.

Whether you came for curiosity or you've been begun to explore, or maybe you have some healthy skepticism already about what it can or cannot do, I hope that this is some new information for you and useful and practical. This is the state of the art as of today, although like all technology, it is rapidly changing and evolving. But I want you to have practical things that you can do in your in your own situations, in your clinics and in your practices. Here are my disclosures, here are the limitations and risks and our learning outcomes for today.

So, as I said, I'm excited to have created this course for you and to discuss some of the most important factors and practical tips about what is AI, how does it show up in healthcare today, and how can you be prepared? How can you think carefully about what are the risks, whether it's the appropriate solution or not for your situation and how to be ready for the future. I want this to be as future proof as possible. And so I'll talk about a lot of products that exist today, but these are just examples. These will.

Commercial companies are building new things all of the time, and I am always excited to read more about them and to try them out. Myself, as somebody who works in technology and cybersecurity, that's the part that interests me a lot. I have a couple of day jobs. My role and my expertise is in technology and in cybersecurity. Artificial intelligence is another kind of software.

And so my perspective on this is how can it be used for effective uses and what are the threats and vulnerabilities to its use? Cybersecurity, technology, AI, these are all tools. They are all helping you do your primary goals. Your goal is healthcare, and I'm glad that you have that. And I want to make sure

that you know how to achieve your goal with the technology available in a way that is safe and effective.

I love doing research in the area of technology. I love discussing how to mitigate cyber risks. And that is where my passion lies. I do a lot of this.

So we're going to start with the general landscape of what does AI look like today, this year, this moment in time, I'll ask a question to you hypothetically, which is how do you feel about using AI for managing your own personal health? You, as a patient, when you go to the doctor or to another health provider, would you be excited or not excited if that provider offered AI solutions or if you were purchasing a health solution that had AI technologies in it? Well, one recent survey said that 80% of Americans would be willing to use at least some kind of AI powered tools in the management of their own health. That's a pretty high number and I suspect the number ebbs and flows. But it will soon be so ubiquitous, the use of AI in our tech, in our, in our lives, that it may be hard to escape.

It may be hard to even know that it's happening behind the scenes. One of the things I want to do is help peel back the curtain a little bit so that you know, what is this, how does it work and what does it mean? I also want to take this moment just to on a brief aside about hipaa, the HIPAA laws do apply to AI. That being said, the HIPAA rules are also written in a very tech agnostic manner. There are no specific HIPAA rules about iPads, There are no HIPAA specific rules about AI.

And the rules that apply apply to every kind of technology. And so things like making sure that privacy is ensured, making sure that patients are treated equitably and equally, that the security mechanisms that HIPAA has in place are, are applicable no matter what technology you use, including AI. So as we go through the presentation, I'm going to point out a few different places where the HIPAA rules intersect. But some of this is just general education as well.

The place where we really need to start is with terminology. And this is helpful for us today in our conversation, but it also will help you understand what you see in the news when you talk to your colleagues and perhaps even when you talk to your patients, making sure that we're all on the same page. Artificial Intelligence AI is actually a field of computer science. Computer science is a big umbrella that deals with hardware and software, with security and privacy. And AI has been around in this field for quite some time, for quite a number of decades.

And it is that umbrella of things that our computers trying to mimic, emulate tasks that traditionally required human intelligence. I think that's what generally we're all on the same page about, which is the computer trying to do things like driving cars or creating images. If you think about those in terms of traditionally human tasks that now computers are trying to do, we're on the same page. That's what I mean here.

The way that AI works at a very high level is sort of similar to how humans learn. And humans learn based on examples. Humans are very, very efficient at this. You can show a child sometimes one example of this is a fire truck, and the child can very easily apply that one example to many other situations and say, oh, yeah, this is a different color in a different setting, but I still know that it's a fire truck. When we train computers, it often requires many, many more examples.

Computers don't work precisely the way the human brain does you. Even though we've started to use similar words and terminology to describe this. The output of that training is something called a model. And you might hear about models. A model is just a piece of software.

It's a piece of software that has been encoded with the examples and has a computer representation to say, this is pictures of cats. That, in essence, in reality, in the physical world, is just a piece of software. And those pieces of software can be downloaded. You can have models that are shared between organizations. Models can be loaded onto a device like your smartphone or a hearing aid, and the

models can be retrained.

We'll talk a lot more about this as we go along. The field of AI also has changed quite substantially in the last couple of years, and that's the next bit of definitions that I'm going to talk to you about. But in the beginning, most of AI was about classification and it was teaching a computer. What category does this represent as an example? So if I said to the computer, here's a picture of a cat, I want the computer to say, this is a cat and not a dog.

This is a cat and not a clown. And so the computer was trained to put things into categories, and it was quite good at that. And we could get closer and closer. And actually, today, things like computer vision, the computer can perform sometimes better than humans. You might have heard this in the detection, for example, of cancer.

Radiologists now sometimes in some cases use artificial intelligence to help diagnose cancer in an X ray, for example. That's an example of this classification problem. But that's probably not the kind of AI that you hear the most about in the news. What we are starting to hear more and more about is something called generative AI. And that's not necessarily putting an example into a category, but it's the generation, the creation of things that didn't exist before, generating new content like text, audio, images, and that kind of thing.

Generative AI is a subset of the bigger field of AI and it's based on probabilities. What is the probability of the next thing in a list? I'm going to give you a whole bunch of examples of that, particularly in something that sort of looks like autocomplete.

But let's dive in even more. This is a picture of models that are available for download. Software developers, people who are developing tools, can share and play with and contribute to these kinds of models. This is one that was created by Meta. It's a very, very popular kind of model that runs, that

doesn't have to be connected to the Internet.

Lots of the AI that I use every day I go to a website to interact with it. But there are also models that developers can download to their own computers. These are available to you as well if you want to play around without having to reveal sensitive information online. So we talked about AI, generative AI. Let's keep going down levels of granularity.

If you've heard of LLMs, that's an acronym for large language model. LLMs are a more narrow subset of generative AI. So they are also about creating new things that didn't exist before. They are trained on large quantities of information, usually text. So companies like Google, for example, are trying to train their model on as much information that is Internet accessible as possible.

Books, Wikipedia, news information software that might be on the Internet. And the LLM then has a very broad exposure to how to answer questions. So we might see the opposite of large language models, which are small language models that are more narrow and focused. Instead of being trained on everything on the Internet, you can imagine a lawyer, an attorney, might want to have a model that's trained on contracts. It doesn't need to understand every book that's ever been written.

It doesn't need to know every web page on Wikipedia. But that model, that piece of software has been trained on something very narrow and specific. This certainly exists in healthcare as well. There are health specific LLMs that exist.

You may also have heard of GPT. GPT stands for Generative Pre Trained Transformer. This is a specific family of AI models created by the company called OpenAI. OpenAI also has a commercial product called Chat GPT. ChatGPT uses this GPT family of AI and it makes essentially a webpage where you can ask questions and get answers back.

This is what I often hear people talking about. If they haven't thought very, or they haven't been exposed to very much AI, they might have been exposed to things like ChatGPT. ChatGPT is one of many competitors. Google has a version called Gemini. It exists on phones.

It exists in the Google search engine now. And so what I wanted to help you see here are there are many layers. Terminology gets to be confusing when we talk about AI. Some people might assume that they, that you mean or that the conversation is about ChatGPT. I just want you to know that that is a very specific instance made by one particular company that does one particular thing.

So with that introduction, when you search in Bing or in Google, you might see predictions. And when you query for things in those search engines, it might give you suggestions. This kind of autocomplete is meant to help users, to help you and me, to be, for us to be efficient. And it's based on predictions about what might the next word be that I am typing as I type in this query into Google what is the first? And Google is offering me suggestions in this case.

Google has trained their computer to make these predictions. And these are predictions perhaps based on what other people have looked for in similar situations. It might also be that they're using AI to say what is the most likely next word? Maybe the next word is likely to be day. And so that is the number one recommendation in the search engine.

We don't know always how these companies are are doing what they do behind the scenes, but it's very likely because they have AI capabilities. Both Microsoft for Bing and Google for their Google search are probably using this technology to do this kind of text prediction.

Google also has a search engine that is built into the Google search and that is called Gemini. And I want you to show you example before I even explain what's sort of going on here, Google Gemini is an AI very explicitly and it's also sort of a fancy autocomplete. And when I type in a question like am I more

likely to get a cold or the flu, it produces output. It doesn't look like Google search. It's not just links.

It looks like text, like what a book might say or what a webpage might produce. And what's going on here is that Google doesn't know the answer, as in it's not trying to understand my question. It's not trying to produce what a human brain might produce. It is in fact doing more like what its search engine does, which is to predict the next word. And so when the answer says it is more likely that you'll get a cold than the flu, it isn't that this system understands the answer.

It's literally trying to predict the likelihood of the next word. And it does that again based on lots of training of examples. And in this case probably they have information from medical websites, from research journals, from news transcripts. And the more that Google has trained, the more it knows the probability of the next word to make a prediction. This also gets better over time.

They continually refine these models and some of them now can go in real time on the Internet and try and find as up to date information as possible. So instead of this being old and outdated information, the AIs can give very current, relative, recent and relevant information for people asking these questions. These are fun things to play around with. And I often find myself people or people asking me, well, well, how does it know? How does it know that it's right or wrong?

And this is the real danger we're going to come back to a couple of times. The AI does not distinguish truth from fiction. That is the, the domain of humans. We need to look at this output and say, does this seem reasonable? Can I validate what it says?

Is there other mechanisms that can help me feel confident that what the AI is producing is true and truthful? But that's a problem that remains in this field. We're going to come back to that.

Companies have built an enormous amount of commercial software and other services for applying AI in many aspects of healthcare delivery and wherever whichever of these pieces you are responsible for, have some interaction with. Let me give you a couple of examples about how this shows up. For one, maybe you have to do some marketing, maybe you have to attract patients and referrals from other healthcare providers. Google Ads, for example, uses a lot of artificial intelligence to try and put the right information in front of a potential patient who's searching for your services. And if you use Google Ads that is powered by a tremendous amount of artificial intelligence, many other companies offer different ways to do marketing.

Again, powered by AI. There are altogether different services for helping patients schedule with you and with other health providers. That might be the form of say, an interactive chatbot on your website, or maybe when they call your phone number, there's an automated AI powered system that is trying to predict, will this patient schedule with you? Will they cancel? Can you, can you give them information that makes it more likely that they will come to see you.

AI is good for that because it can learn from examples to help predict again the probability of some other event. There are technologies for appointment preparation and if you need to come up with a new form, if you need to come up with examples, you might use an AI powered tool to help you Prepare before the patient comes in to be as effective in the delivery of healthcare for them as possible. That might also be reviewing old notes from the last time they came in. For example, to summarize whether what you tried last time were all very busy people. The AI can try to review previous records, previous notes for that kind of thing.

In the appointment when you are in front of a patient, there are plenty and growing number of technologies available for AI powered for the time when you are with the patient. And that might be AI powered software for programming a device or for example, for transcribing the appointment that you're

having right then. We will talk a little bit more about this in a second. But if you want a technology that can listen to the appointment, can take notes, can generate reports for you, that technology exists today after the patient leaves, there are yet more technologies that can help with the next steps. If you want, for example, an AI that can generate a medical record, that technology exists.

If you want to do research about how the likelihood that the patient, based on the clinical observations and the data that you've collected, needs different treatment, there are AIs that can help do that as well. I've seen a couple recently that can help with billing and coding. Certainly payment is a big part of the healthcare ecosystem and this is a niche skill that not every healthcare provider has. And now there is technology available to help you do that. One of the most interesting and quite frankly mature areas is in continuing education.

I will very often ask an LLM to summarize recent research, even in my own field. You can do this in yours. I can say what are the latest advancements in a particular area of healthcare? Or here are two recent research documents that I don't really understand. Can you summarize them for me in a few sentences as if you're explaining it to a fifth grader?

AI and LLMs are very, very good at this. And that is part of all of our jobs, is to keep on top of what's going on. I think this is a pretty effective and safe way for you to think about doing that. So again, AI in the whole ecosystem of healthcare delivery.

So where is AI? In addition to those professional services, one that I didn't mention on the last slide are AI powered devices themselves. This exists in smartphones and glucose sensors and hearing aids, many of which have AI features. Now as part of the health related devices themselves. You might also see it in clinic specific software.

So it might be built into the ehr, it might be built into other medical equipment, the vendors now are seeing the advantages and the opportunities here to help you do your job, to help the delivery of healthcare be as effective as possible, not trying to replace you, trying to make you an effective healthcare provider. And then I would classify everything else as third party applications. Chat GPT is a commercial software that you can purchase as a third party and you can use it in some aspects of healthcare delivery with caveats that we'll discuss. Heidi Health, for example, is a third party application that does transcription. It records audio during an appointment and it can transcribe that.

It can create notes that might be useful for you. And then it has started to permeate AI has started to permeate all of our individual devices. Google Gemini is built into the Google Chrome browser. Apple Intelligence is built into the newest version of iPhone. Windows Copilot is now built into Windows.

And so all of those are available to you and we are going to talk about how to use and use them safely. All of them exist to help you save time and burden. That really should be the primary reason for thinking about whether you should adopt AI is how can it help? How can it help me, my business, my clinic, my patients? And it does a lot of that through automation.

We've had a lot of automation for a long time. That is not AI. But really the biggest benefit here is I can do things faster, I can save time, I can save money. All of those are reasons to think about adopting AI.

Some of the most mature and commercial solutions exist now in things like chatbots, in scribes and in digital assistants. So I want to just touch on those three things because you are likely to see and maybe want to experiment with them very soon. Chatbots tend to be external facing, as in the patient might interact with the chatbot to ask questions, to schedule an appointment, to get information about your hours or your email address. The chatbot is intended to help again, help you be efficient because it can do triage before they have to call and talk to a human. Chatbots have pros and cons.

There are lots of commercial software now that you can integrate into your website. And so that is one thing that some patients prefer, being able to do that efficiently at their own schedule and at their own pace. For you as a healthcare provider, it's likely that you have or may soon be thinking about something like a scribe, which is an AI that's there to, as I said with Heidi Health, to listen to an appointment, to generate notes, and it's as if you had another person in the room with you who could do those things. While you focus on the patient, it frees you from having to type as you are trying to listen, to try and not have to multitask, to give more care to the patient. While this other technology does what you would have had to do as a human, there are lots of commercial softwares that can do that now.

And virtual assistants are sort of one step even farther than a scribe. A scribe is just sort of listening and turning that audio to text. A virtual assistant can do many more kinds of things and it tries to be more of a partner in the delivery of healthcare. There are some commercial services like this. It's the same kind of idea as the sort of virtual assistants you have in your home, except very, very tailored to the delivery of healthcare.

Now, I showed you an example of ChatGPT, that is the or and or Gemini. So the version. The two things I want to show you here are side by side comparisons. I typed my video that I showed you a moment ago into Gemini, which is the screenshot on the right. I typed the exact same query into ChatGPT, which is the window on the left.

Functionally, they are doing similar services. It's two different commercial companies and sometimes the answers are different. It takes a while to learn how to ask questions of this kind of technology. And this is what people refer to when they talk about prompt engineering. Maybe you remember when you first started using Google that it took a while to figure out what words do I type in that give me good results.

We sort of have to retrain ourselves how to do these queries again with large language models. Because just typing in words like restaurants near me, that isn't how these systems expect to get questions. It's much more like a human. It's like having a conversation with a human. That's also a change, which is when I search for something in Bing, in Google, in any search engine, it's usually a one time question and then I find some options and I click on one and I go try and answer my question.

What is different about ChatGPT and other LLM chatbots is that it can be a conversation, sort of like learning to talk to a human. If you don't quite get what you want, you can refine the question. And so if I ask am I more likely to get a cold than the flu? I could ask follow on questions like how is this different in the United States versus Europe? Or I am starting to feel these symptoms.

Which one of them is more like a cold versus the flu? It's much more conversational, but it does take Some amount of experimentation to figure out what works the best. So that's what prompt engineering is.

Let me come back to HIPAA for a second. Most of these as user facing commercial services are not considered HIPAA compliant and they will openly tell you this on their website. If you go to the OpenAI ChatGPT page it will very explicitly say we are not HIPAA compliant. For people just typing in questions like I was showing you, we are only HIPAA compliant for people who develop software. So if you find a solution and they say we are powered by ChatGPT or powered by OpenAI, they can have a different kind of agreement with that vendor that could make them HIPAA compliant.

But out of the box, just buying the service or using the free versions online, you very much should not enter protected health information. If you were to use any AI service like you would any third party having a business associate agreement would be a necessary first step. This is a legal agreement that says that the vendor understands that you are going to be sharing PHI like patient names and that that information must be appropriately protected. OpenAI will not sign a BAA with you as an end user and

neither will Anthropic. The makers of Claude Gemini is the one exception here and it is the exception only if you use the business version of Google Workspaces.

If you pay for your email or Google Drive. There is a way to sign a BAA with Google so that you could do HIPAA phi related queries. This doesn't mean you can't use it at all. And that's the caveat I want to mention here. It is possible to use these services and not reveal phi.

For example, if there was an audiologist in Maryland who wanted to know how to reply to Google reviews, that audiologist could ask a very generic question which doesn't reveal any protected health information. But it still is useful for that clinic. It still might be useful for that healthcare provider. This question isn't phi. It's not found by HIPAA because it doesn't say the patient patient's name, any identifying information, anything that would must be protected under hipaa.

So you can still use these technologies even if they're not HIPAA compliant as long as you're not using them for HIPAA related HIPAA covered services. Now with all of this being said, I hope I've given you some hope that these tools can be useful, but I want to give you the other side as well. And I want to talk first about what AI gets wrong.

One thing that is very on the top of people's minds right now is bias. And bias, as in gender bias, ethnic, racial bias, and producing outcomes that are not what we would want from an unbiased tool. And this is a little bit unfortunate. And it's something that the vendors are working on, and it's a byproduct of the fact that it's trained on all the information that exists on the planet, much of which is already biased. And so it propagates that bias by saying, well, I learned that this unfortunately biased result is very common, and it will produce the biased result back to you because of its training data.

Many researchers are trying to figure out how to improve this so that ChatGPT and other AI doesn't promulgate so much bias. But it's something that exists in healthcare. It's something we have to be very

careful about and always on the top of our mind. Even us as humans, we can't get rid of bias. The goal is to understand that it happens and to produce mitigation so that we catch ourselves when it happens.

The same is true in AI. AI companies want to know that, oh, this seems like a biased result. Maybe we need to tweak that answer a little bit. Another thing that AI gets wrong is that it doesn't understand truth and fiction. And this is too bad because it might generate output that's hard to distinguish.

And one of the biggest concerns from policymakers is, well, people don't have any help in understanding whether the output of an AI is true or untrue. What I want to point out is that AI is not built for this. And if we want to solve this problem, it will be some other mechanism that accompanies the AI to accommodate for it. Maybe that's labeling images generated by AI to say this was generated by a computer to help humans who are viewing that image make better choices about whether to believe that this is a true image or one that could have been manipulated.

All of this can happen for many reasons. The AI isn't explicitly trying to trick you. It's not trying to be intentionally deceitful. It doesn't know what those things are. Those are human attributes.

So it's incumbent upon you and me and all other consumers to validate, to look at the output of a machine and say, does this seem reasonable? Can I validate what it's saying in another way to make our own human distinctions? The third thing that you're likely to have heard about are what are called hallucinations. And this is sort of a step farther than just truth and fiction. But it Is the AI generating output that doesn't exist?

And it might make up a fake person, it might make up a fake event, a fake outcome. The thing that I want you to note that's important here is the technology is actually performing exactly how it was programmed. That doesn't make it very satisfactory. It. You might not be happy about that, but hallucinations are part of AI because the computer is always trying to predict the next word and it can

be the result of not having a lot of examples.

If I ask the AI a question about something that's very rare, it probably doesn't have enough information to make good predictions, but it's always going to generate output, it's always going to think of some answer. That answer might be totally fake. It might have totally made up that person or that event. And we've seen examples of this in legal cases and news articles and lots of others where the AI created things that didn't actually exist. This is a burden today on humans.

It means we have to very carefully check and validate the outputs. What else does AI get wrong? Let me give you a query that I put into ChatGPT recently about the height of a child and I'll just read it to you. It says, when my son was 7, he was 3ft tall, when he was 8, he was 4ft tall, and so on and so forth. What do you think is going to happen?

What do you think might the result be? Maybe you guessed, but the AI said at this rate, the child at 8 years old would be 12, or at 12 years old would be 8ft tall. It's blindly following patterns because it doesn't quote, unquote, understand what I'm asking. And it doesn't have this intuition that you and I have that humans don't grow linearly forever. That is not the way that AI works.

And so it's going to suggest something that doesn't make common sense at all. Just be careful about that. This is probably not a question that is meaningful to you, but as you ask other questions or use AI for other things, be aware that it doesn't have this kind of common sense. The same is true even when it seems plausible. Which is I might ask a HIPAA compliant AI to summarize a medical record and suggest a treatment plan.

You would think that it would produce truthful, accurate results based on the information that it has. And yet the computer could still produce things that sound plausible but are totally wrong. Maybe this medical record is a hearing test or a balance test and the record doesn't talk at all about hypertension.

But the AI says, oh, the patient has a history of hypertension and prescribed lisinopril. Well, why would it do that?

Again, it's just doing average case predictions about the next word. It doesn't really understand what it's doing. And so again, it relies on us as humans, us as experts, to validate the truthfulness here.

One more thing about what AI gets wrong. As I said, the rare events are the ones that can be the most troublesome. And there have been even recent research in the last couple of years about this, about diagnosing rare and unfamiliar medical conditions. You'd think that the AI would be very good at this because it has so many bits of information to consider. And in some cases that can be true.

Sometimes it can suggest something that a human didn't even think of. But in the general sense, we shouldn't rely on AI in unfamiliar kinds of questions, and we certainly don't want to overly rely on it in rare, unforeseen situations. This isn't to say that humans don't make this mistake as well. Humans certainly do. Our memories are fallible.

And what the computer can do better than us is consider all of the information at its disposal. But really, we humans, and we humans with technology, we have to be very careful about this. Now, when AI gets wrong, these kinds of errors can be acknowledged and hopefully in the future they can be corrected. There's also things that AI just cannot do, and researchers are struggling with this. And these are the kinds of things we don't know that we can make the AI do.

For example, making what are ethical or moral judgments is very much a human construct that's very difficult to train a computer to understand and to apply. You might have to make medical judgments like this, like, should I given a, given a crisis, treat patient A before patient B, or do something that isn't just a straightforward, obvious rule based decision. These are things that humans struggle with. They are also things that today AI is just not capable of performing. Another example here is creativity.

It often seems like generative AI is being created because it produced, wow, a spectacular picture based on just text that I gave it. But in fact it isn't creative in the same sense as a human. It's again, just trying to follow patterns, to use things that already exist to make something else. One example of this, if you want to go learn a little bit more, is that the composer Beethoven, after Beethoven died, had sketched out ideas about a new symphony but never completed it. And some researchers said we're Going to use computers and AI to generate the whole symphony.

And you would think that perhaps the output was an original creative symphony. And if you go read about it and listen to it, you can tell that the computer was recycling, was using things that already knew to fill in the gaps in. It wasn't trying to create what Beethoven might have done as a human, creative human. And third here, the computer, the AI is not intended, it's not designed, and it is not good at displaying empathy. There's a pretty interesting book called Deep Medicine that might interest you.

And that book talks a lot about how AI is good for some things. But it also frees healthcare providers to be the ones who can be empathetic, who can understand what is the patient going through. And maybe they have a medical condition that's more important or more dire even than the one that they came to see you about. Humans can display true, honest empathy in a way that the computer, even if it's trying to pretend, is not very effective at. I don't think patients are anywhere near wanting to get information that seems empathetic from technology.

So we talked about what it is bad at and what it's impossible at. There are unfortunately some things that are dangerous that AI is good at and one of those is being malicious. And a lot of hackers, a lot of criminals are using AI as a tool to do bad, malicious things. This is important for you to know about and so I want you to know how to defend against this. One is that attackers are going after the AI on medical devices.

They want to know can they trick them, can they? For example, there's lots of research about self driving cars and hackers trying to change the size of the stop sign or something else that makes the self driving car do something it wasn't supposed to do, like go instead of stop or speed up when it should be slowing down. The same is true for attackers. We've seen research examples of them going after the AI on medical devices themselves. That's a problem for the vendors to deal with for the most part to try and detect that that's happening, to be resilient against it.

Another that you should be aware of are deep fake attacks. As we said, generative AI means that the computer can emulate real people. Maybe that is your boss, maybe it's said administrator, maybe even a patient. If you get a telephone call that says please send me my medical records, there is a non trivial possibility that that is a deep fake. I hope it doesn't happen and I want us to be resilient against this, but probably every healthcare provider should have some validation that the phone, the person on the phone is actually the person you're talking to, because these deepfakes definitely exist and attackers are using them more than ever before.

And the third one that I'm going to talk in more depth about is AI generated phishing, because phishing is a real threat to all of us and it can cause real harm by infecting our computers by releasing sensitive patient information. Let me talk about that more in depth. There was recently a research study done at IBM, and IBM wanted to compare phishing attacks generated by humans in their research lab and phishing attacks generated by a robot, by AI. The humans used information about their victims. So pretend that you might be a victim.

They would look up information about you and your interest on the Internet. They would create an email with some sense of urgency. They would say that there's not much time, and then they're going to send this email to you and say, oh, you must take this action. Trying to trick you into doing something. The approach for the AI was a little bit different, which was, give me a list of concerns for people in

healthcare or in a particular specialty, and please create for me an email that uses social engineering to try and trick them.

Please make it sort of look like marketing. And then the computer said, well, all right, who do I send these to? And off they went. Now, which one do you think tricked more humans? In fact, the phishing that was generated by humans was more effective and less people knew that it was fake.

The humans just did a better job. But it took a long time for those attackers to create that. It took 16 hours for the researchers to generate that information and send it off the computer. The AI did this whole process in five minutes. It wasn't quite as good.

It was still quite convincing and quite damaging. But the attackers probably are willing to sacrifice the time, the efficiency. And so that's why I've started to see many more emails that are almost certainly generated by AI. What can you do? Number one, when in doubt, don't trust the sender.

If you can, call the person who sent to validate if you have any suspicion. There was an old heuristic, right? You and I were very much taught that fake emails often had typos. This is not a good rule anymore. The AI can create fake emails with perfect grammar, perfect English.

And so just basing it on the fact that, well, if it's good grammar, it must be safe, is no longer a good heuristic. Also, expect that this kind of attack isn't going to happen only in email, but also in voice, on the phone and in text messages. The same technology can be used for all of those kinds of attacks. It's very important that you manage your digital identities, which means pick good passwords if you have to use passwords, use multi factor authentication if you can, so that in the unlikely event that a phishing attack is successful, that hopefully your account is better protected because of good identity management and know that these attacks are going to change. This is the worst case today, it could change tomorrow.

Attackers are always trying to keep up in this kind of game.

As things evolve, AI will certainly produce more technology. In the near future, I expect to see more and more AI in personalized medicine, in recommendations that are specific to me as a patient that understands more about my comprehensive healthcare situation, my personal behaviors, whether that's at home or in the clinic. That's an area where I think there's going to be a lot of growth. There will also be even more powerful and effective predictive analytics about what is the likelihood that some event is going to happen, that I might have cancer, that I might be the victim of identity theft. The more information that the computers learn, the more that it can be effective in those kinds of predictions.

And I can expect to see probably companies soon who are using AI for things like training, education, simulation. Generate me a fake patient and I would like to interact with it as if this was a real person. The computer can help do that without having to put people at risk or to help scale out education. Every year, the Consumer Electronics show showcases commercial technology, including healthcare. And you may be interested to go see what was in this year's awards for digital health.

Some of this technology is already on the market, some of it will be soon. This is one sort of peering over the edge of the horizon and it's available to all of us. It's easy to digest.

One thing that you can do today, right now, is if you are an owner, if you have supervisory power, or if you want to recommend it in your place of work, is to develop AI policies and procedures. Like other parts of our workplace, policies and procedures are the first line of defense. We don't want an employee to say, well, nobody told me that I couldn't upload medical records to ChatGPT. There ought to be a policy for that. Every clinic and practice today needs an AI policy.

These kind of policies can be put into an employee handbook. They can be the, they can be special training. It can be just a topic of conversation at a staff meeting. But now is the time to do it. And the goal isn't to make anyone afraid or necessarily limit their use.

In fact, I hope people use new technology. I want them to see what's possible, but I want them to do it responsibly in a HIPAA compliant manner. That HIPAA compliance, for you as a healthcare provider, that is the most important thing to consider. Is the patient data secure? Where is it going?

Will the third party vendor protect it just the way that you would as a healthcare provider? Making sure that there's a business associate agreement in place? That's what HIPAA compliance can look like when you go talk to a commercial vendor. But as I said, it is good to empower people to give them freedom within constraints, within boundaries about how to do this safely and effectively. What does that policy look like?

One is to give it some scope and framing and say we as a health organization are going to take reasonable steps to make sure that our use of AI is safe, that it's compliant, that we follow all the same rules that we do for our computers and our email and our other accounts. AI again is just software. This kind of a statement can be positive, encourage people to use it for good as long as it's within these constraints, and that you want them to have reasonable rules for that kind of freedom. This policy and procedure should have something about hipaa. It should say we will protect phi and personal identifiable information in a way that is consistent with our practice and consistent with the HIPAA laws.

It should have some amount of call for verification, which is, as we've talked about humans, not just blindly trusting everything that the computer produces or says or does, because us as people, we are ultimately responsible. The HIPAA rules do have sanctions and they are not sanctions against the computer. They are the responsibility of us as people, of healthcare providers, of those with this sensitive information. So it's very important to create a policy that is appropriate for your situation that

encourages people. I hope to be positive about what it can do, but also acknowledging what it can't do.

Having sanctions is another thing. If people violate any of the rules of an organization, there should be consequences for that because again, we want to make sure the business is safe, that you as employees are safe, and ultimately the patients and their information are safe.

Now, since I can't predict the future, I do want to give you some general framework questions for how to future proof yourself. One is that when a new technology comes out or when people start talking about it on social media, and you want to know whether it's right for you. What questions should you ask yourself or the company? One is whether the service, the third party service, is HIPAA compliant, whether they have instituted data protection, whether they talk about their certifications. For example, you should not use any third party service that takes health information unless they sign a BAA first.

That business associate agreement is a one time thing between your clinic and the service provider and it extends the HIPAA protections. Because you are going to give them patient information, they must protect it just as much as you would. You might ask them questions about, well, who has access to the data that we're going to share with you? What happens to it? Can it be used to train a model that other people have access to?

These are good, important questions to know. You might look at a commercial product and say, well, I want to know how it came to its decisions. That's sometimes called explainability. Transparency is another phrase that you might see used by vendors. And this makes us as human consumers more confident.

If the AI just produces a result but it can't explain, well, how did you come up with that answer? It's, I'm more skeptical about it. I would be much more inclined to believe it if it could say, well, I pulled information from these websites and I stitched the information together and I'm pretty sure about it,

right? That degree of transparency exists in different forms depending on the technology that you use.

And I also want to know how the product deals with errors. No product is 100% safe. I have to update my laptop very frequently because the software is imperfect. AI is also imperfect. It will make mistakes.

I would very much like to know from a new vendor, how do you deal with errors? How do you deal with inconsistencies and bias? If I get to talk to a vendor at a booth, these are the kinds of questions that I like to ask them. And finally, look for general culture at the company. Do they have a history of compliance?

Do they have an incident response plan if something should go wrong? Good, secure companies that care about security and privacy, they are more likely to extend that into new technologies like AI.

So let me summarize before we open the floor for questions. I hope I've convinced you, if you weren't already, that AI is a very powerful capability. It can do things and help humans do things and accomplish the goals that we want to do at almost superhuman speed and with superhuman accuracy and efficiency. That's terrific. We should be thinking about how to use it appropriately, the kinds of situations that it is good for.

AI cannot solve every problem. It Cannot do lots of things. It gets a lot of things wrong. It's not as the proverbial sort of hammer in search of a nail. It should be on our radar for particular kinds of problems, like, well, I need more effective marketing.

I need to save time in reading research articles. It is a helpful tool depending on the kind of goals that you have as you think about using it. In those cases, it's also important to think about the risks. How could this leak patient information? How could it go wrong?

How might my patients be uncomfortable with it? You might ask yourself questions about is it appropriate for your situation? Not only is it safe and secure and compliant, but does it? Is it right? For me, that's a fair and appropriate question.

Today, we should not let the computers run loose. They're not capable of replacing us. Humans plus AI are much more powerful than humans alone or or AI alone. The combination of the two is not only necessary, but much more useful. You can feedback information that makes it more effective.

You can validate the information and decide what is likely to be true or not true based on your expertise, your human experiences. That's very important. Don't neglect that at all. And finally, today is the time to start. It's a good time to create an AI policy.

Even if you're not yet using AI in the clinic. As I said in the beginning, it is very likely to come into our lives and be unavoidable, just as it is built in now to our operating systems into our computers. Better to have a policy that you don't need than to not have the policy too late. We don't want to create policies after a data breach. Much better to be proactive about that, to be prepared and to empower people.

I would love to be able to tell lots of stories about AI being a terrific companion in a clinic, and I look forward to hearing those additionally from you as well. With that, I would love to hear your questions and for you to follow up. This email address will always reach me if you have questions that you don't want to raise today or as you think about them later. But I would be glad to take your questions now as we're waiting for these to come in. I have the Q and A here open as well.

One question that people sometimes ask me is should I tell my patients that we use AI? And the answer most certainly is it depends. And here's what it depends upon. HIPAA does not require you to disclose to a patient all of the technology in your practice. You don't have to tell them what kind of computers you use, what EHR you use, what email provider you use.

You don't tell patients those things. And so I'm not a lawyer, but as a cybersecurity person, my interpretation of the HIPAA rules is that that isn't something you must disclose. However, patients might become a little bit sensitive to the computer listening, recording, sending their conversation off into the cloud. And that could be a reason that you go above and beyond the compliance requirements and tell them, I use this kind of technology. Maybe you even offer a written consent form that our clinic uses AI to help provide the highest quality of health care we can.

Do you consent to that? That is up to you. That is not a HIPAA requirement, but it is one that could be on your mind.

So a question came in that says that the main concern is that there are published examples of professionals already ascribing unverified accuracy to AI tools to meet unrealistic productivity goals. I completely agree with this. Supervisors see increased productivity, employees see less work. And so the right question is often not asked. How do you know the answer is correct and was it verified?

I think there's multiple levels to this question. And the first is that it's good to have a policy and it's good to establish norms. How are we as a collective group going to document, to communicate, to make this a part of our culture, turning something into a culture like, well, we're always going to lock the computer when we walk away from it, or we're always going to pick passwords that are strong and secure. Those are as much cultural as technological. We can put frameworks in place that make it more likely that you all can be, feel good that things are going in the right direction than having to assume.

Assumptions are never good. And so if the question is, well, how do I know that the answer is correct? You have to in some sense trust your employee and know that they had the policy, they had the training, that you have a culture of talking about it, that there are no negative consequences to asking questions. All of that is good security and privacy culture in a clinic.

The next question that came in says, what makes you feel, what makes you feel that AI is worth implementing in the clinic Right now? It seems to be a layer of burden. Why I have to be so skeptical about what it does and double check it? You're right. I will acknowledge first and foremost that there is some overhead.

And I think the overhead is offset by the productivity that it might, that it might produce for you. So if, for example, this is an experiment you can do in your own situation. How long does it take you to document an appointment and write a report? Try a commercial service that can do that automatically and see how much time it takes you and then measure, well, how much time do I have to do to read through it, to correct any errors, to detect mistakes? It might be that it slows you down, that it's actually worse for you.

For many people using AI in, in practice these days, it actually saves them time in total. And that's why I think people are sort of excited about it. Is that all in all? Yes. There's this additional burden, but it still helps me in the same way that for better or worse, security on your laptop is a burden.

Right. It takes time to log into your computer, but the benefits outweigh the risks and outweigh the cost of doing that. But that is a cost trade off that you have to think about for your own situation.

Very good question.

This question says incentives are strong in favor to cheat with AI. So discouragement has been equally as strong. I can acknowledge that and it is a consideration. It is true for all technology. I think we all have to be appropriately skeptical.

But AI is not brand new in the same way that I had similar conversations to this when cloud computing was new or when email was new. Do I'm sure you all have lived through the experience of new technologies. AI is a little bit different. I don't think it's going to go away. I think it will become more integrated in the systems that we use, even in ways that we don't even know that it's happening.

When I stream videos on the Internet, my streaming services are recommending things, my shopping services are recommending things behind the scenes. There's a lot of artificial intelligence in marketing, just as it might be for you trying to attract clients about recommending movies or recommending books or anything like that. Can that be used for ill. Yes. Every tool since the dawn of humankind could be used for bad, criminal, malicious things.

But knowing how to use them for good, that's what I'm excited about. I think that's what lots of people are excited about.

A question came in that said I would like to incorporate AI into my practice, but I don't know how to start. Do I have to purchase an AI program? Are there free resources that I could use? Or. Or you could recommend probably all of the above, by which I mean the software that you already use may have AI features that either are optional or not turned on or could be added.

And so it's useful to look and see what's already there. You could even think about the tools built into your computer. Again, if you use services like Siri or Google or Microsoft Copilot, just be careful about whether or not to share patient information there based on the protections in place. The There are some free versions of some of the things I talked about. Gemini, Google, Gemini, their chatbot, even ChatGPT.

There are free versions of those things if you want to just try them again. Recognize your limits about what kinds of questions that you should ask. But actually doing the hands on really was helpful for me.

It taught me a lot and it might be for you too. Like, oh, it doesn't seem to really be able to do what I want.

There are some commercial products like Heidi Health, which you can go onto their website and you can purchase as a service. There are other ones like that that are commercially available. Some of them have trial periods or test accounts where you can again give it an experiment, try it for a day or a week and see how it performs in your setting and then you can decide, yeah, I want to pay the monthly fee because it is so valuable for my situation and you can buy that as a commercial service.

Thank you so much, Dr. Dijkstra. This is just a reminder if you'd like to earn CE credit. If you can, please take the multiple choice exam that populates 10 to 15 minutes after the course has concluded. We'd like to thank Dr.

Dijkstra for his time and his expertise. We'll go ahead and close the classroom. Thank you everyone.