

**Levi:** Hi, everybody, welcome to the hands-on healthcare machine learning broadcast. I'm Levi Thatcher, here with Mike Mastanduno.

**Mike:** Whoa, Levi, what are you doing here?

**Levi:** I'm back. I was on paternity leave. I have a new baby.

**Mike:** Well, congratulations.

**Levi:** Thanks, thanks. Nice to be back.

**Mike:** Yes, it's great to have you.

**Levi:** Thanks, guys.

We're excited to bring Risa Myers on the show today, from Houston Methodist. A data scientist there. A new group. And so, we'll be chatting with her here in just a moment.

**Mike:** Yeah. We've got a great show. It's really awesome that Risa could join us.

Before we get started, as normal, we want to make this interactive so feel free to engage with us in the chat. You can log in to YouTube and get your name next to your comments. You can up your resolution, although we don't have any code on the screen this week, but if you need to, it's in the gear in lower right corner. We'd love to have you join our community and participate on Slack outside of the broadcast.

This week was pretty quiet in the news. There wasn't really any news to discuss. And it was also pretty quiet in the user-submitted questions. I guess, we did such a great job last week that—

**Levi:** Answering all the questions, weirdly. Yeah.

**Mike:** Yeah. Nobody had any other questions.

But we did ask the users to download the latest version of the package straight from GitHub and check out the new XGBoost multiclass classification functionality.

**Levi:** I'm excited for that.

**Mike:** So we'll have to tell Levi about it. But if you had a chance to use it, please let us know in the chat how it went, how we can improve it. And maybe I'll do that when I get some time.

But with that, I think we should just kind of move into the main event here.

**Levi:** Yeah.

We're excited to have Risa Myers join us from Houston Methodist. Thank you so much for joining us, Risa. Thrilled to have you on.

**Risa:** Thank you. Thanks for the invitation.

**Levi:** Super exciting stuff you guys are working on down there. And just to start with the broad level here. So Houston Methodist is a health system based in Houston, kind of scattered throughout the part of Texas there.

**Risa:** Yes. We have eight hospitals throughout the greater metro area here. The last one is opening up this summer in the Woodlands.

**Levi:** Beautiful.

And Houston Methodist Hospital, which is in Houston, been ranked #1 hospital in America by U.S. News & World Report.

**Risa:** Yes.

**Levi:** I'm not sure if that was last year.

**Risa:** We're in the top list. I don't—

**Levi:** Top list—

**Risa:** We're #1 in Texas.

**Levi:** In Texas?

**Risa:** I think we're #19 by US News.

**Levi:** Super—

**Risa:** Yeah, #1 in Texas, #1 in Houston. And I think it's #19 in the US on that list.

**Levi:** Superimpressive.

**Risa:** Yeah.

**Mike:** Wow! That is impressive.

**Risa:** Thank you.

**Levi:** That is amazing.

**Mike:** Houston is a real hotspot for medicine.

**Levi:** It is.

**Risa:** It is. We actually have the largest medical center in the world here.

**Levi:** Oh, wow.

**Risa:** I don't know how many institutions off hand but as I was saying earlier, looking out the window, I can count Baylor College of Medicine, Texas Children's, MD Anderson, St. Luke's. They're all right here in one place.

**Levi:** A hotbed for healthcare.

**Mike:** Yeah, that's super exciting.

**Risa:** It is.

**Levi:** That is exciting. We're thrilled that you could join us. We'd love to just hear a little bit about your background, Risa. So, in general, how did you get to where you are? Tell us a little bit about your path.

**Risa:** Okay. Sure. Thanks.

So I'm a computer scientist, pretty hardcore. I actually have four degrees in computer science which I'll tell you is probably one too many. I got involved in healthcare. Really, my first job out of my undergraduate degree was— like I went to work for Hewlett Packard Medical, back in the medical products group. We were building cardiac ultrasound imaging systems. And that just really got me engaged in a field that helps people out. It was not invasive which was great. So there was no blood. But I just felt like I was building something that made a difference.

From there, I actually started working a Master's in computer science part time and remote. This was back in the dark ages. They actually videotaped classes and take them across the country, a week delayed. And part of that deal was actually I went out to California, to Stanford, to finish my Master's degree. So I did that.

And then life circumstances brought me to Houston. When I first moved here, I went to work for a research group at Baylor College of Medicine. We were doing some great early EHR work. We ran out of funding which often happens to research groups. So I did something completely different for a while and really hit a point in my life where I realized I wanted to come back to healthcare. It's just a field I feel very passionate about. I like the fact that it can help people. You understand disease better. You understand what's going on better. And you make a difference in the world and in people's lives.

So it took a number of steps to make that happen. A few years ago, probably about eight or nine at this point, I joined MD Anderson Cancer Center. And I thought, "This is great. I'm set. I can bike to work. It's the top cancer center in the world— certainly, the country, probably the world." And there, I had a chance to work on a data warehouse. So I was part of the IT team and we were taking the transactional-based data from the EHR and transforming it into a format where people would do longitudinal research on it.

And that was great. And it was useful and neat. But what I realized was what I really wanted to do was I wanted to use the data warehouse, not build it. So in order to do that, to really sit at the table with the docs, you need a PhD. So I wanted me, a good time in my life, to really step out across the street - Rice University is right across the street from the Texas Medical Center, and I took some time and got a PhD in computer science with a focus on machine learning.

While I did that, I had the chance to continue to collaborate, both with people at MD Anderson and then I met some other people, some faculty at Baylor College of Medicine, and I did some research with them as well. So all of my PhD work, again, was focused on predicting outcomes or— some of my work was predicting outcomes from [inaudible 00:05:15] clinical data. In particular, I tended to look at intraoperative or ICU data and predict the outcomes from there. So I collaborated with MD Anderson. People there have continued to work with me. [inaudible 00:05:25] collaborated with the Baylor College of Medicine as well.

And now, I'm at Houston Methodist and hoping to do some similar work here. I've been here just over a year now so.

**Levi:** Wow. What a fantastic place that you're a PhD.

**Risa:** Thanks.

**Levi:** And such a hotbed of healthcare innovation. That's beautiful. It's always great to see more folks from computer science coming into healthcare as well. It's nice that they have that tie-in.

**Risa:** Right.

**Levi:** So new group there at Houston Methodist. It started up about a year ago. Do you mind just talking a little bit about the origins of the group and how you're structured within the hospital system itself?

**Risa:** Sure.

So the group is not as new. The addition of data scientist is new. There are two of us here now. We've been here just over a year.

We're part of an IT informatics group. And the focus here, we really sit on that boundary between research and operations. It's one of the things that really appealed to me about this job, is we have a chance to—

So Houston Methodist is, it's an academic hospital but the main priority is clinical care for the patients. So we do research but we really are looking after our patients, a little more than some of the other more traditional academic hospitals. And we have that balance. So we have the [inaudible 00:06:26].

The IT informatics group was formed— hang on, I had notes on this. I don't remember. So originally it started from a request back in 2012 from the CEO, Dr. Boom. He asked us to put together a data warehouse. And the group kind of grew and evolved from there.

So we have people on the team who do web development and app development – mobile app development. We have some database experts. We usually also always have a resident medical doctor on hand who also helps us understand the terminology better and helps us interact when we need to get some advice from the clinicians or get some more information. So the team has really evolved. And our—

So the way we work, we get a request from a number of different places. We'll get a request from the operational side. I'll get a VP who e-mails me and says, "Hey, can you look into what are the differences, we want to look at this particular intervention that we've done with patients. Are they patients that it helps more than others? So, is there some population where it matters more?" Or we'll get a doc who's really excited about predicting re-admissions in a

particular subset of patients. So, they'll champion and come to us and say, "Hey, we want to build an app to predict readmissions on this patient body." And we'll spend time and work with them and figure out what's involved? Do we have the patients? Do we have the data?

We have an oversight committee that meets about monthly. I think we're going to switch to bi-monthly shortly. And sometimes in the summer, we don't meet as often. But on that Board, we have the Chief Medical Officer. We have the Senior VP [inaudible 00:07:52] the Executive VP, so basically the CEO of this hospital which is Methodist Hospital, which is our flagship hospital and our high-acuity hospital here in the medical center. The CIO sits on that Board. The CFO, so the Chief Financial Officer as well. We have the Emeritus Director of the Research Institute there. The Chair of the Surgery Department just recently started attending. We have a new Center for Outcomes Research and that Chair started attending with the group. So we have really the right people in the room. We have someone in charge at the physician's organization and quality. We have also people in the room who are able to balance the hospital's interest overall. So the project's lead come from them, or from other executives, or from clinicians and we basically have a round table

[inaudible 00:08:41] here's what somebody wants to do. There's a form they fill out that says, "Here's what's novel and here's what's different. And here's what the patient cohort looks like and what we think we might be able to answer, both in terms of operations and in terms of research questions." And then we talk about it. Does it make sense to do it? Does it not? We have limited resources. Do we want to go in this direction or not?

**Levi:** Wow, that's fantastic. So I think this is kind of common with data science groups, in general and specifically in healthcare, that you have sort of these grassroots efforts of, okay, well, let's figure out who's working on what and who's [inaudible 00:09:08] machine learning and doing these sorts of things. And get a project going with them. But then, also, you have an Executive Committee. Are they specifically for data science type prioritization? Is that why—

**Risa:** For this team, that team meets mainly to prioritize what this team works on. So the team focuses on the corporate data warehouse, any mobile apps that we do that are tied to research and the data science initiatives.

**Levi:** That's fantastic.

One thing that we've noticed a lot is, it's good to make sure that you're solving the questions most important for the hospital.

**Risa:** Right.

**Levi:** A lot of times, with machine learning, it's so cool to just want to go make predictions for everything but, obviously, if they're not helpful then you aren't really accomplishing much.

**Mike:** Yeah. I've got to say, it felt like a failure when we decided that machine learning wasn't appropriate for a specific product. Like, but we need them in all of the products.

**Levi:** Yeah, yeah. Exactly.

**Mike:** It's like, "No, not that product."

**Levi:** You have to be practical about it.

**Mike:** "We don't need it in that one."

**Levi:** Exactly.

**Risa:** Well, I think another thing I found in my time here is we also need to do some foundational work. One of the groups I'm working with is focused population health. And I'm helping them build a database and work with a vendor right now. And they just needed some help understanding their data better and processing it and doing a sanity check on it. So it's not really data science per se, it's maybe more data engineering, more database design but it's fulfilling a need and it's getting to know their domain, and it's getting to know the people. And that's a lot of what I think I've spent this first year doing is getting to know people and getting to know the data.

Houston Methodist changed EHRs two weeks before I joined so, okay the data warehouse is built on the old data and then we've got the new system coming in and trying to figure out. You know, my PhD work and my background has been mainly in surgery and anesthesia, so now I've got a broader set of data that I'm dealing with. And now, I've got to learn what does that mean, and how does that work, and how do I understand that data, interpret it? And really, start building these friendships and collaborations and understanding where we're going and what's important to the organization.

**Levi:** That's fantastic. A lot of common themes, I think, we'll hear throughout healthcare is, "Okay, well how do we prioritize? And how do we find folks to work with?" So with this population health initiative that you mentioned-- like with Mike and I, like working with Health Catalyst folks, you're not always helping with data science or machine learning really that matters. Sometimes, it's like, "Hey, how can I get P values out of this? Or how can I run some T tests?" So with the population initiative that you mentioned, how did that collaboration happen? Like you kind of just strike up conversations in seminars or— you know, what's your best advice?

**Risa:** So I'm going to talk a little about this one [inaudible 00:11:29] project that I'm really excited about. So this one, I actually interviewed-- part of my interview process for the job was I interviewed with the physician in charge of that operation. So she was someone that I just felt like, "Oh, we kind of click and we have that dynamic." So when I came on board, I made an effort to get to talk to her and say, "Hey, can we get involved in this?" It turns out, at one of these oversight meetings, the Chief Medical Officer said, "You know, you guys should get involved in population health." And that also triggered the, "Okay, now, we've kind of got that executive direction to go get involved."

And sometimes you go in and you say, "How can I help? How can I-- what can I do?" And it's not always that glamorous, "Okay I'm going to build this fancy model." Sometimes it's, "Okay, let me combine your data for you so we can get the answers we need to figure out. Is the program working?" And these are just simple metrics. How many patients started? How many patients finished? But again, it's the ground work you need to get going.

Another example that I want to give is—so I've got a lab coat on, right? I'm a scientist. I get to wear a lab coat.

**Levi:** Looks good.

**Risa:** Thank you.

It's back to that. If you want to sit at the table with the MDs, how do you get them to pay attention to you? How do you get them to notice you? And not just, "Oh, you're just another administrator or [inaudible 00:12:30] in the hospital."

So I go to grant rounds we're an academic institution and there are numbers of groups that have grant rounds. Which for those of you not familiar with, they're basically teaching opportunities. They can be someone within the hospital giving

a talk. Or a lot of times, they bring in outside speakers. And it's great to go to these (1) because you learn something and (2) you learn who's there, and who the important people, and what are the topics that are being discussed, because these are also probably important to the organization.

So I'm terrible at raising my hand at the end of a lecture and asking a question. My mind goes blank. But what I am good at is going up to the speaker afterwards and saying, "Hey, you're doing X, I'm doing Y. Can we talk about—" And I did exactly that at one case. There was a physician who had given a talk. And she had big bold red letters on one of her slides, "I don't know how to predict intracranial response in patients." So I wanted to jump up and down because that was my thesis. So I went up to her afterwards, I said, "Okay, I know how to do this." And she was a really busy woman. And she said, "Oh, yeah, we'll talk."

And I think the next two months, every time I saw her in the hallway, "Hey, can we talk? Can we talk?" I had friends put in a word for me and we finally sat down. And what you find then is that it turns out there were only three patients they could've looked at. So we didn't go there. But you have this negotiation of, "What questions do you have? What skills do I have? What data do we have? How do we find the right sweet spot where we can start doing something that matters?"

This is a fantastic project that I'm very excited to be working on. So in this case, we're looking at liver transplant patients. And there's a metric that the national organization uses to evaluate centers. And we feel that there is a measure in that evaluation process that we think is very subjective, so we want to replace it with a more objective measure. So here's a chance where we're going to say, "Okay, we don't have all the national data but we have our data. So let's look at our data with current score. Let's take out that objective measure. See how it does as a baseline. Let's put in something else that we think— I'm sorry, take out the subjective measure and put in an objective measure and look at those different models and see if one of them can more accurately predict what's going on.

**Levi:** That's a fantastic point.

**Mike:** Yeah. That's so cool.

**Levi:** So a lot of what data scientists do is not just straight up machine learning. We play that up along the broadcast because we're excited about it and getting into it more and more and that's what our packages do.

**Mike:** It's titled—

**Levi:** It's titled [inaudible 00:14:38] we've got to, you know, blow up our title.

**Mike:** [inaudible 00:14:36] you know?

**Levi:** But really, a lot of the work of a data scientist is, "Okay, how do we gather this data appropriately? Are we measuring the right things? Are we using T-tests appropriately?" It's basic statistics a lot of the time.

**Mike:** Yeah.

**Risa:** Right.

**Levi:** It's fantastic that you brought up the social aspect of it. I feel like—I think we're all pretty nerdy. But there is, I think, a social component of the data scientist's role that's maybe not there - the software engineer's role or the statistician's role. Has that been, I guess, pretty integral to your work then, being willing to step out of your comfort zone?

**Risa:** I think it's about communications.

Absolutely.

**Levi:** Yeah.

**Risa:** So it's standing in the elevator with a bunch of docs after grant rounds thinking, "I've got to introduce myself." "Oh. Hi, I'm Risa." Just working of it because maybe I think some of the docs have the same problem we do that we're a little introverted and antisocial but you have to be willing to communicate.

Why am I talking to you guys, right? My adviser drilled into me, "Any chance you have to talk about your work, take it. Get a chance. Meet people. Have random conversations."

Somebody approached me on LinkedIn and said, "Hey, I'm working with one of the docs here, do you want to--?" "Okay." So now we've got another potential project that we're working on.

Be willing to talk to people. Not everything is going to turn into a great research project but you don't know where it's going to go. And in the meantime you're going to learn, what else is going on at your institution or what else is going on in the field. And that's been really valuable to me.

**Levi:** That's fantastic advice. Kind of the always say yes mentality of "Sure, I'll help you out for 20 minutes." Like, we'll see where it goes. Why not?

**Risa:** Yeah.

Right, that's exactly right.

**Levi:** And I feel like across machine learning and data science in general, data science's roles a lot of times are there to help the company learn more about data science in general so like at Health Catalyst were helping others become more familiar with R and Python and machine learning.

So how has that played into your role specifically, so the education component? Is that—

**Risa:** That's a great question.

**Levi:** Not really a mandate but is that part of what you guys were supposed to be doing?

**Risa:** Great question.

So that's not something we're doing a lot of today, outside of our group. Within our group, we have a monthly journal club where everybody takes a new technology or a new topic and once a month we'll present about new SQL databases or TensorFlow.

**Levi:** Wow, that's a good idea.

**Risa:** And so, that helps the whole group stay up to speed at least a little bit on some of that and we try and spend some time playing with those tools as well.

We haven't done anything here but I'll tell you, I've given a grant rounds talk at MD Anderson in anesthesia about my work.

**Mike:** Awesome.

**Risa:** So there are the opportunities to say, "Hey, here's what we are doing. Here's what is going on in this field. Here's the domain but here's the machine learning part to it." So there's certainly the opportunity of doing that here as well.

And some of it is, it really is, "Hey, here's who I am, and here's what I'm doing. Here's the kind of things I do." I mean, I tend to do that more in a one-on-one basis, go up to talk to somebody, "Hey, I'm doing XYZ. Can I help? Can we talk?"

So nothing formal in terms of learning to the hospital yet but I think that might come down the road.

**Levi:** Yeah. It sounds like a lot of informal opportunities. With all the research going on, there has to be a lot of folks interested in machine learning and stats in general and, of course, data. So that's a fantastic opportunity.

**Mike:** So, on that note, Risa, we had a couple of questions from the chat. People are interested in what your day-to-day toolset, your dataset, what that kind of stuff looks like. I think you mentioned you're using R for your analysis. Do you have like a sandbox of test data you can use? Can you talk about that a little bit?

**Risa:** Okay. Sure. So [inaudible 00:17:52] what do I say?

Most of my time, I spend writing a SQL code.

**Mike:** Okay.

**Risa:** Honestly. You know, I'm pulling a lot of data from— we have a data warehouse so I pull a lot of data from the data warehouse and then I'm starting to pull data from the new EHR and we're trying to integrate that as well. I do that. I do a lot of the statistical analysis, I'll do in R.

Because I'm a computer scientist, I'd like to program. I haven't had the right opportunity to do that here. So one of the challenges I have is my PhD work was on predicting outcomes from real-time data, so I have this really cool model that I want to use but we don't have the data yet. We actually don't have the infrastructure.

Like a lot of hospitals, real-time data takes up a lot of volume, and they don't preserve it unless there's a reason to. Or they preserve only some subsets of it. So back to that groundwork, I'm working with the medical device integration team to start thinking about "How do we start saving that data? And what do we save? And how do we save it? And how do we access it later?"

So then, hopefully, I can go back to either like C++ code, or R, or something else to start analyzing some more of that data. But I would say, mainly here, we use R kind of a lot of database programming. Some people use [inaudible 00:18:56] as well. I've used that a little bit for something else.

**Mike:** Yeah. That's interesting to hear.

Yeah, I guess, no matter how much you want to— even with four computer science degrees, you can't get away from SQL, right?

**Risa:** No.

**Levi:** It is the common currency for data.

**Risa:** [inaudible 00:19:10]

**Mike:** It's a necessary thing in data.

**Risa:** Yeah.

**Levi:** Yeah, great advice for you folks out there. If you want to get in the business, learn SQL.

**Risa:** Learn SQL.

Oh, and not just for manipulative but it's— when you get a new dataset, how do you store it?

One of things that I think is so important is make sure whatever you do is repeatable. I'll get a project from one of the VPs and I'll do it. [inaudible 00:19:29] it's done. It's not done. That was just phase 1.

If somebody asked you [inaudible 00:19:33] with a number, make sure you can always defend where did that number come from. I literally keep a text file. "Okay, here's the query I read. Here's the file it went into." Because you always end up going back to that information in doing the next level. And you get—

Even though you think at the time, when you're in the middle of it, it makes sense, so I'll never forget this. You forget it. So I really try and document. What did I do? How did I generate this graph? Where did this data come from? So that later on—

Sometimes, right? Sometimes, you find a mistake. You misunderstood something. You left a condition out. Or oh, they want to do this with a different subset of the population. It's much easier when you have that query available to go back, pull the data a little differently, manipulate it a little differently but you know exactly what you did.

**Levi:** Yeah. Maybe, four months down the road, the executive slightly changes the question or some other executive wants it for a different division. That's fantastic advice.

**Risa:** Exactly.

**Levi:** So in terms of the prioritization process, you mentioned it's kind of grassroots and you have this committee as well. So have executives been excited about clinical decision support using data science and machine learning?

**Risa:** Pretty much so.

**Levi:** Or what are the kind of the flavors of projects that are getting prioritized?

**Risa:** So I think it's the same thing that you see across the country right now. Everyone is looking at re-admissions, probably more than anything else. They know that changes are coming in the policies. And how do we start looking at those numbers and doing the right things for our patients.

Methodist has started a continued care organization with CMS. So, again, looking at how do we care for these patients best? How do we reduce our costs? How do we care for everyone? Reduce readmissions - all those classic things that everyone is interested in finding out. How can we do that?

**Levi:** And when you say coordinated care, is that making sure that once they leave the hospital they're not sort of abandoned?

**Risa:** Right.

**Levi:** How do we keep them out of the hospital, in general?

**Risa:** And help them be healthier. Right.

**Levi:** Yeah. It is really interesting. So much of healthcare is focused on what's going on inside of the hospital but if something gets there, of course, it's failure of the system in general.

**Risa:** It's changing. I mean, as everyone knows we're moving from a pay-for-service model to a— I don't know what the right term is, but improved health and preventative model. How do we improve care and this capitated model is coming where you get a certain amount per patient for amount of time. How do you care for them in the right way rather than just bring bodies in and do imaging or do whatever to raise money. That's not the goal anymore. It's how do

we— and not that the goal was ever just focused on making money but caring for patients but there's a different focus now.

You want to make sure, "Has somebody already done this test? Let's not do it again." And you see that throughout the organization. We have a health information exchange with other organizations in the medical center. So now we can get more information about the patients who've been seen somewhere else. And [inaudible 00:22:00] less effort and not spend the time re-doing a test. And so "You've already got the answer, let's go from here."

**Levi:** That is fantastic.

**Mike:** Yeah.

**Levi:** You hear about that issue so much where, "Okay, well, we went to a different hospital system. And so, the doc can't see your past info." It's kind of crazy outside of healthcare, people don't— we understand that that's how it works but mind boggling.

**Mike:** And kind of on that note, we had a great question from the chat about de-identified data and privacy. Now, I imagine since you're pulling most of your work from the EHR, there's going to be names and dates attached to all of that but do you do any work with de-identified data?

**Risa:** Sometimes I de-identify the data. So, for example, this population health project is a great example. We've pulled some of their data into a separate database because not only is our population health project about patients but it's about our employees. We're focused on employee risk. So we motivate all of our employees. We incentivize them to give us their blood once a year and we look to see if they're hypertensive or of high cholesterol. And if they are, then we work with them to try and lower that. So not only are these patients but they're co-workers.

So we've pulled some of that data and we have a vendor handling a lot of this for us. But to do some of the integration, we've got stuff in a separate database. And I've got a co-worker I work with, if you need to do some stuff [inaudible 00:23:14] set views that hides— she knows the names but let's hide their Social Security Number. We don't need that. We need it to communicate with the vendor to get everything done right but let's not expose any more that information than we need to.

**Mike:** Yeah.

**Risa:** So it's something we're very careful about and very sensitive to here.

**Mike:** Great. Thank you.

**Levi:** That's a fantastic question across healthcare.

So we mentioned clinical decision support is of interest. How do the clinicians feel about that from what you've seen? So the idea, right, is that okay, well, instead of a clinician trying to keep everything in their head in terms of journal articles and Joe's past history, the idea is to serve up something to them. "Okay, well here's Joe's risk of readmission, or high length of stay or an infection." How do clinicians respond when you talk about those sorts of things?

**Risa:** Well, I'm unsure on the other way. We don't normally drive— we drive some of the projects— most of the projects are clinician driven or executive driven.

So most of those projects, they are clinician coming to us to say, "I think we can predict readmission using some small number of features. Let's do it." So they are very happy to work with us. Or if I go talk to [inaudible 00:24:19] they don't even realize we're here, so they're excited, "Okay. Well, we can do— how can you help me? What can we do?"

So maybe I'm only talking to the people who are really excited about this but that's still a good place to start and it gets other people motivated. And they start seeing success stories.

**Levi:** That's amazing.

**Risa:** So we haven't hit barriers at all there.

**Levi:** Beautiful. It's best when they pull from their end because they're engaged and then they appreciate the tool. They will use the tool. That's beautiful.

And you mentioned readmissions being a top priority which, as it is, a lot of health systems.

**Mike:** Yup.

**Levi:** Is that specific to a certain disease category, most often? Or do folks just want to know, "Okay, broadly, who's likely to be readmitted within a certain time period?"

**Risa:** So it's a little bit of both. I mean, one is— and I'm not an expert on policy but I know there are certain areas where you'll get hit and you'll get penalties. I think COPD is one of those areas. There are a couple of specific cases or diagnoses where if those patients are re-admitted, you get financial penalties or they're coming.

But in general, everyone is looking at "How do we reduce readmissions?" Those 30-day readmissions are key. And again, making the person healthier so they don't need it. And what's involved in that? What are the different predictors? What are the sub-populations that tend to come back? How do we help them?

**Levi:** Beautiful.

And you mentioned the clinician engagement being great. And so, when you find these features to drive the model or to understand what's causing these readmissions, how do you work with the clinician in a literature? What's kind of, I guess, maybe the workflow, you could say, of how do you find out exactly what is driving readmissions for heart failure patients, for example?

**Risa:** So I think data science is a really, especially in this capacity, is a very iterative process. So that project I mentioned about the liver transplant patient. So we had a big powwow last Friday with the Head of Transplant and the Head of Liver Transplant and some other physicians to really sit down and say, "Okay. Where are we? Where are we going? What are we doing?" And they had some really fantastic insight about, "Well, here's why we think this might work. Here's why we think this might not work. Maybe you want to consider this."

Really, if you can get those people in the room, which I think we're really lucky to have had that opportunity, it really helps. That' when—

Right, I know enough clinical data to be a little dangerous. I know enough about statistics to know what I don't know and where I need to be really careful. So if you can get the right people to bounce those ideas off of you end up with something better and you don't waste your time.

That meeting really helped us focus, "Okay, let's look at this in more detail." So now we're off to do more of a lit review. We're off to figure out, "Okay, do we have all these measurements that we want to use?" We've gotten some great insight there.

**Levi:** Yeah. The clinical inputs—

**Risa:** And now we're going to go back and then we'll follow up and iterate, "Okay, here's what we learned."

When I started [inaudible 00:26:53] project, we did the same thing. We would meet, really. And I'm a big fan of 10- or 15-minute meetings with clinicians. They're busy. They've got a lot of things to do. Let's meet for 10 minutes. Let's talk about where we are. I've digested the next piece. Okay, now I'm going to go off and work for a week or two. Let's meet again. Here's what I've learned. Now, I'm ready for the next level. Both, I'm ready for the next level and they're ready for the next level because they're learning what we can do with machine learning as well. So that type of rapid turnaround, you know, short meeting, short focus, I found to be very productive.

**Levi:** Yeah. And just— I love that. Nobody likes long meetings.

**Risa:** No.

**Levi:** Or at least, most people I talk to don't love long meetings. So be persistently polite. Get 10 minutes on the calendar. Be brief. Be to the point.

It seems like people are really excited about machine learning and healthcare. What are the docs saying that you've been talking to or the clinicians? Have they heard of any of these buzz words or were they--?

**Risa:** Well, they have. I mean, because we just went live with a new EHR, I think the biggest focus right now is how do we get to the new EHR data?

And then it's been a concern for everyone. It's been a concern for our team. It's been a concern for the EHR team. So that's really been the focus as people have kind of been hung up for a year like, "Okay, we're using all the old data. Where's the new stuff?" And so that's a concern. I think we're a little stuck on that right now. We're in the process of fixing that. And once we fix that we'll be back to, "Okay, what are some more interesting questions we can ask? And how do we partner more?"

**Levi:** Yeah. That's, I think, the problem in a lot of places. How do you actually access the data?

So a couple of points, I guess, if we're talking to budding data scientists out there, folks looking to do this, is get clinicians engaged. It sounds like you've done a fantastic job with that. And then actually have a data warehouse or a

functioning data warehouse that allows you to pull data from disparate sources. You mentioned some silos or—

**Risa:** Actually, I was going to back up and say, one of the things that I think this team has done really, really well and it hinges back to that or goes back to that de-identification question. We actually have a cohort discovery tool. We have a tool that people can search – the research community can go in and say, “Okay, well, how many male patients between the age of 17 and 34 are there who have diagnosed, you know, congestive heart failure?”

Probably a bad example but you can pick those ICD-9 codes or you’d look for certain procedure codes and you can look for a patient demographics. And this is all de-identified but it gives them a sense of “how big is the cohort?” So if you're going to start writing a grant or you want to start looking at a project and you want to see, “Do we seen any of these kind of patients or not?” They can do that without even coming to us.

**Levi:** Oh, beautiful.

**Risa:** And that’s a huge benefit. It gets them started. It gets them going.

We're now at a point where if they enter an IRB, we can then give them a de-identified subset of data. And again, they can start playing with that. And then if they really want more data that’s identified then we get involved and set the table. But we're starting to speed up that process in giving people more access to their data in a safe fashion, in a de-identified fashion.

**Levi:** That’s fantastic.

**Mike:** [inaudible 00:29:33]

**Risa:** That’s a really neat tool. Yeah.

**Mike:** That’s s kind of a sneak preview of next week where we're going to be talking about text analytics. How much of the data that you work with have you seen being coded versus free text? Or is there a trend moving one way or the other?

**Risa:** Great question.

So text versus structured data. I mainly deal with structured data. That’s what my PhD work was. We do have some people on the team who have done some text analytics and some MNLP, some of it [inaudible 00:29:57]. But I haven’t been as involved in that effort.

**Mike:** Okay.

**Risa:** But it's a huge opportunity.

And we had that problem today. Okay, here's a report. We need to pull data off that report. So it's coming. It's harder to deal with though. The intro to next week's talk.

**Mike:** Yeah. Great. Thank you .

**Levi:** Yeah.

Fantastic stuff here, Risa. Just a couple more questions, if you have a minute.

**Risa:** Sure.

**Levi:** So you mentioned clinic decision support is impotent. Do you find kind of patterns and that you're staying mainly in the critical realm, or operations, or maybe financial use case of data science a little less common? Or how's the balance there?

**Risa:** Oh, that's a great question. So I've had people mention to me that, "Oh, we need to start looking at supply chain. We need to start looking at financials." We haven't done as much of that. Mainly, we've been focused on clinical and operations, and mainly the clinical side. But there are plenty of opportunities in the hospital and plenty of interest there, to look into those kind of problems. They're huge. Yeah.

**Levi:** Yeah, just a matter of number of hours in the day to get it all done, right?

**Risa:** Right.

**Levi:** Yeah.

That's exciting that everybody's that into data science and wanting your help. That's fantastic.

So I did have a question about the data warehouse. So this is kind of a common term in healthcare. How do you get the data silos broken down and get the data into one spot so you can work with it easily? Besides EMR data, are there other data sources that you're excited to work with?

**Mike:** Ooh, that's a great question.

**Levi:** Maybe census or-- I don't know, if you guys work with a certain payor and have claims data but talk a bit about data warehouse a little bit?

**Risa:** Yeah.

So there is all of that and we get it in bits and pieces. So often, it becomes project dependent. What project are you working on? What other data sources do you need?

So we have, basically, a peer database or a data warehouse where we can load data in. And we've also found that when we start working with a new data set, we'll sometimes pull that directly. It motivates us to pull it into the data warehouse itself.

But that stuff does exist. And we're kind of dealing with that really on a case by case basis. There's a lot of silos, as everyone knows, in hospitals. And we're just— how do you break that down? How do you encourage people to share that data? How do you safely share that data is something we deal with everyday and struggle with everyday, trying to figure it out.

**Levi:** Yeah. I think we're all struggling together with that.

That's why we're here, trying to break down silos across healthcare machine learning and data science.

We really appreciate you coming on.

Mike, anything on the chat? Everything looking good? Are they asking questions?

**Mike:** The conversation has followed the chat nicely.

**Levi:** Good, good.

**Mike:** And vice-versa so.

**Risa:** Thanks.

**Levi:** Yeah. That's fantastic.

**Mike:** I'm happy on my side.

**Risa:** Great.

**Levi:** Yeah.

We appreciate you joining the Slack Channel, Risa. We may have to ping you for a question here and there in the future when we get stuck on a particular problem.

**Mike:** Oh, that's a great plan.

**Risa:** That would be my pleasure. It's a great resource.

**Levi:** Yeah.

**Mike:** If you had any outstanding questions, you can get Risa on the healthcare.ai Slack Channel. She may or may not respond but you can send her a message.

**Levi:** That's right.

Fantastic chatting with you. We really appreciate you taking the moment. And we hope to do it again soon.

**Risa:** Great, thank you.

**Mike:** Thanks so much, Risa.

**Risa:** Bye.

**Mike:** We'll talk to you soon.

**Narrator:** Thank you for joining us today. Remember to like, share and subscribe. Comment below and click the links below to join our Slack Channel and our healthcare.ai Community.