

Webinar: Machine Learning Informed Artificial Intelligence in Healthcare

Empiric Health's mission is to reduce unwarranted clinical variation, thereby improving healthcare outcomes and affordability. Primarily focused on surgical services, Empiric Health integrates Electronic Medical Records (EMR) and supply utilization data, which is then processed by a proprietary AI engine that sorts surgical encounters into clinical cohorts, or groups of comparable surgeries. This engine was created hand-in-hand with clinicians in order to facilitate direct clinician engagement. While AI engines provide traceability and clear explanations, they struggle to scale and expand quickly when exposed to novel data or changing practices.

Empiric Health partnered with DataRobot in March of 2020, implementing a machine learning (ML) model factory to monitor performance and increase AI engine expansion. Integration of a machine learning-informed AI engine provides benefits from both methodologies — the ability to scale and expand quickly (ML) as well as traceability and clear, human understandable logic explanations (AI).

TRANSCRIPT

Andy:

Hello everyone, and welcome to today's webinar. Machine Learning Informed Artificial Intelligence in Healthcare. My name is Andy and I will be your moderator for today's webinar. Before we begin with our presentation, there are just a couple of housekeeping items to note. First of all, today's webinar will be recorded and everyone attending today's webinar will receive it. We will also have time allocated for Q and A at the end, so if you have any questions throughout the presentation, please enter them in that question box within your go to webinar panel.

Now, I'm pleased to introduce today's speaker Megan Bultema, Chief Data Scientist at Empiric Health. Megan obtained her doctorate in biochemistry and molecular biology in 2012 from Colorado State University and continued her research as a postdoctoral fellow at Stockholm University in Sweden.

During her academic endeavors, Megan applied statistical methodology towards inter-atomic force field parameterization and novel cancer therapy discovery, publishing a total of 11 publications in peer reviewed journals and serving as a visiting assistant professor at the University of Colorado.

Following this academic career, she leveraged her experience as a data scientist and focused directly on machine learning methodologies. Megan is currently the chief data scientist at Empiric Health, leading a team of highly skilled individuals focused on developing innovative solutions to support the mission and reducing unwarranted clinical variation. With that, I'm pleased to now turn it over to Megan to begin the presentation.

Megan Bultema:

Thanks Andy, and I want to thank DataRobot for having me here to present this webinar and all of you in the audience for attending. I'm very excited to share the work that we're doing at Empiric Health, as well as the lift

that we've gained by partnering with DataRobot.

Empiric health has an AI-powered clinical analytics company. We're focused on improving outcomes and reducing costs in surgery across the nation. The underlying methodology that we use was developed over time at Intermountain Healthcare to further their initiatives in reducing unwarranted variation, and thereby improving outcomes and costs. As a critical step on this journey, Intermountain developed a proprietary AI rules-based engine, that sorts surgical encounters into groups or cohorts to enable comparative analytics and opportunity identification.

In 2017 Empiric Health span out of Intermountain as a partner startup company with not only the mission to carry this out into the broader healthcare market, but also to enhance and innovate upon the original methodology. What does success look like for us? What does it look like to reduce unwarranted clinical variation? One of our focus areas is on reducing variation in surgical supply cost.

In these figures, we're looking at a histogram of surgical supply costs for a single surgery group or cohort, the total knee arthroplasty or total knee replacement cohort. On the left-hand side, we're looking at a distribution of surgical supply costs for the surgery at a hospital system that's just beginning their engagement with us. They are just starting to work on this. On the right hand side, we're looking at a similar distribution at a health system for this surgery, that's been working on reducing supply cost variation for years.

We can quickly notice two things. First, the variation and the distribution has decreased through this work. The curve has tightened and the standard deviation has decreased. The second thing we notice is that the overall average supply cost per encounter has also decreased. Now, to be successful in this work, there's one crucial audience that you have to engage with. Engaging with supply chain is important. Engaging with an executive leadership team is important.

But direct engagement with physicians is critical, and the single most crucial step in that dialogue is providing clinically meaningful data. The key to that success for us lies in our cohort methodology. We're going to use that [inaudible 00:04:46] true analogy of looking for the apples-to-apples comparisons here. There is no shortage of discrete data elements that are available to describe a surgical encounter. However, if you use something like procedure codes to group all patients together, you may end up with a random basket of fruit.

You're going to end up with patients that received a specific procedure for a multitude of different reasons. Including cancer or a traumatic events or an elective surgery. This understandable variation is going to affect the surgical supplies that are used and the length of those surgeries. This is warranted variation and it really obscures our ability to find the opportunities. Next level in refinement is looking at the diagnosis codes that are available on a surgical encounter. These would be your DRGs, your CPTs and your ICD10s.

This gives you information about why a patient would have received a specific procedure, but you still end up with some variation. You might be down to a basket of apples, but you have green apples and red apples all mixed together. The next level of refinement that Intermountain discovered that they needed, and that Empiric Health has continued to refine, is the ability to extract information directly out of the operative note text.

That operative note text is dictated by a physician during the surgery, and it provides a host of information about the context of the surgery, why the surgery is being performed, any complications that arise during the surgery, and this can inform our grouping or our cohorting process. One example of this is for a total knee replacement or arthroplasty surgery, the context of whether that surgery is a revision of a previous knee replacements or whether it's a primary surgery.

This information is consistently available in the operative note text, but is inconsistently documented within the discreet data elements. Knowing this is crucial to our conversation and to our opportunity identification methodology, we have taken over the AI rules-based engine that allows us to get to these red apples to red apples comparisons. This cohort engine has more than 1500 different logic rules that are applied to over 70,000 different data elements. This includes over 150 different concepts that we've extracted directly from that operative note text. We use this AI rules-based engine to create 310 different distinct surgical cohorts.

This is a snapshot of what that logic engine looks like for a single cohort family. We're going to stick with this example of total knee arthroplasty or total knee replacement. I'll use that term interchangeably here. We're going to look at the cohort family that we have developed over time. In the flow diagram, going from left to right, we're getting into increasingly more stringent groupings and more apples-to-apples comparisons.

We're starting off by gathering all encounters where the patient received a knee arthroplasty surgery, and then step-by-step through this decision tree. We are further refining our groups into five distinct cohorts at the very end. We are evaluating procedure codes, diagnosis codes and also those concepts that we pulled out of the operative note text. All of this gets us to a clinically meaningful population that we can both discover opportunities from, but also engage in physician dialogue.

At the very end here, we have five different individual cohorts and we also have a group of encounters that were excluded from all of these cohorts. We call that population affectionately, the export. These cohorts that we'll be working with today are the total knee arthroplasty or total knee replacement, a partial knee replacement, a bilateral total knee replacement, bilateral partial knee replacement, and then a knee replacement revision surgery.

One thing to note here is that if we were trying to use some of the discrete data elements that are commonly used like the MS-DRG code, this population would be very mixed together. These cohorts were derived with physicians, every one of our cohort families, every one of our cohorts were developed with a physician leader involved in the process and reviewing the final encounter placement to really drive that clinically meaningful conversation.

That gives you a bit of a picture about Empiric Health, some of our foundational methodology, but what I want to talk about today is how we're leveraging machine learning to enhance and grow our methodology. To level stuff with that, there's a few terms that I'll give a basic working definition for. AI is a term that we often hear for artificial intelligence. It's been recognized even in the 1980s, that AI or artificial intelligence means many different things to many people.

This is one working definition. Artificial intelligence is any technique that enables computers to mimic human intelligence. This includes if then rules, logic, decision trees, expert systems. It also includes more structured techniques such as machine learning and deep learning. Machine learning is a subset of artificial intelligence. It includes structural statistic techniques that enable computers, algorithms to improve at tasks with experience. By this, most oftentimes we mean training data. Deep learning is a further subset of machine learning focused around multi-layered neural networks. We are not going to go that deep into this, so I'll leave that definition there for you to read. How is Empiric Health leveraging these technologies to further our mission?

Well, in April of this year, we partnered with DataRobot, and before our partnering with DataRobot, we had worked on a few models. We had toyed around with putting some of these models in production. But when we started our engagement with DataRobot, a few things become really clear to us. One, we could iterate much more quickly using an AutoML platform. We could test out ideas and concepts. We could experiment with what

structure was going to work best. We were also able to leverage the deployment structure within DataRobot to integrate our models into our database pipeline.

Now that we have access to both our AI rules-based engine and to machine learning models, we wanted to look at that cost benefit analysis of using one technology versus the other. Our AI rules-based engine provides a lot of things that are very helpful, very beneficial to our company. A couple of those are the traceability and logic explanations that you can get out of an AI rules-based engine. We just spent probably less than a minute on looking at that flow diagram for the total knee cohort family, and it is easy to grasp that conceptually.

It's easy to describe that to our user base and they have a high degree of user comfort with how an AI rules-based engine works in this context. But rule engines are going to struggle in a few areas. They struggle most notably when they're exposed to novel data. For us, this occurs when we're onboarding a new client. If we don't have an existing rule for novel data, our AI rules based engine is not going to perform the way we want it to or expect it to.

Another place that we run into this is with changing practices at existing clients. The medical field is always innovating and changing and if there are new codes or new procedures that are described in the text, this rules engine is going to struggle with that. Alternatively, using machine learning models in place of an AI rules-based engine, these machine learning models can scale and expand quickly. They are flexible to integrate novel data if handled correctly, but there are some drawbacks to this.

A couple that are relevant to us, there's clear logic explanations that we need to engage with our user base. And also traceability, although DataRobot does have some functionality that helps with these. But again, I've just highlighted that user comfort. This is highly important to us. It's part of our critical business approach that we're able to engage in a dialogue that's meaningful and explain our cohorting process to our users. Within the broader context of healthcare in general, I think we're seeing right now where there's a degree of comfort that needs to be gained for machine learning models to be integrated into the health system. Especially when patient treatment is on the line.

What we did was we decided that we wanted to integrate both of these available methodologies. Instead of choosing one or the other, we wanted to leverage the benefits from both. We have developed a machine learning informed AI engine. This allows us to scale and expand quickly, but we can still maintain that logic explanation that's important to us and that user comfort that is also critical to our business approach.

I'm going to share a bit, this project is how did we do that? What was the structure of our models? How are we integrating these two processes? A very high-level view, what we've done is we've taken the results from our clinically driven foundational AI, that's our AI rules-based engine. We've taken the output of that engine and we've used it to train machine learning models. We then turn around and use the predictions from these machine learning models to inform expansion of the AI engine.

One of the important steps is that in between these two, we have clinical review process happening of the results from our machine learning models. Then also reviewing the results before we are expanding the AI engine. As we work through this process, we're able to leverage the benefits of both AI and machine learning, and this allows us to get past those pain points that we have previously had with our rules-based engine.

I'm going to take a few slides to talk about our model strategy and some of the modeling decisions that we made during this process. At a very high level, if you think about our problem here, we have the entire surgical population from a health system. Our goal is to be able to classify which of those surgical encounters was a

partial knee replacement versus a total knee replacement, versus any other surgery that's also occurring within the health system, and these extra bubbles here are just to represent that we have many different cohorts within this population.

One of the decisions that we have in front of us is whether we're going to use a multi-class model or many 1vRest or 1vOld models. These are two different classification approaches that are available to us. For a multi-class model, you have a single model, and when you're training that model, you're training it to identify for a single encounter, should it be placed in class one, class two or class three.

When you are going to multiple 1vRest models, you're training multiple different models. Each of those models is specialized to recognize, is this encounter in class one or not in class one? For the second model, is this encounter in class two or not in class two? So on and so forth. Now, when we are approaching this decision, there's a couple of known challenges with multi-class models. Multi-class models struggle when they have many classes to be able to classify. In our case, we have 310 different cohorts.

We could identify quickly that a multi-class model would struggle with this many classes and would even struggle with a much smaller set of classes. The other thing that a multi-class model is going to struggle with is imbalanced class sizes. We have a very small cohort population for the bilateral partial knee replacements. This does not occur that frequently. This imbalanced class size presents a problem to a multi-class model.

We wanted to test the effectiveness of the multi-class model versus multiple different binary models. We did this for our single cohort family looking to get it as total knee replacement cohorts. What you're seeing here is the training volume that we had for each of the cohorts within our training data, and then the hold out F1 score for the single multi-class model or where we moved to multiple different binary models.

As we would expect, when the training volume is low, the multi-class model struggles significantly. But when we transfer that over to a binary model, we can see a large lift in our accuracy, in our F1 score. This allowed us to identify that we wanted to move to binary models. The other thing that we get out of moving into multiple binary models is access to the prediction, explanation insights within DataRobot.

What we're looking at here is every row is a individual encounter. You can see the predicted probability that this encounter was placed into that total knee replacement cohort. Then the additional information that we receive here is the explanation for why that encounter was predicted with that score. The strongest indication for the encounter on the top was the operative note text.

Then if we dropped down to the bottom, some of these encounters that were predicted not to be within the cohort, we can see that some of the strongest indicators for those encounters to be predicted not in the cohort was the primary procedure on the encounter. This information is very relevant to our business problem. We want to inform our AI rules-based engine. If there is a contextual situation that the models can bring up to us, they can indicate cohort placement, we can incorporate that into our AI rules-based engine.

Another model strategy that was in front of us was whether we wanted to use one layer of models versus two layer of models. Taking this back out to the high level again, if we are looking at our entire surgical population and we have one layer of models, we would have an individual model for each cohort, and those models would be trained to identify, is an individual encounter a partial knee replacement or is it any other surgery?

That population of any other surgeries would include lap appendectomies, Lap Choles, hysterectomies. That's the setup for a one layer system. But in a two-layer system, what we do is we introduced another layer of modeling. These models are built and trained to predict whether an encounter is in a cohort family or in the

general population. For this example, we built a model to predict whether an encounter should be in the total knee cohort family or not.

Then in our second layer of models, we're predicting within that population of the total knee cohort family, which one of these individual encounters should that, or individual cohorts should that encounter be placed in. This drives model specificity. You can imagine if a model is trained to recognize a partial knee versus a lap appendectomy, versus having a model that is trained to recognize a partial knee replacement versus a total knee replacement versus a revision knee replacement, those two sets of models have different specificity.

All of these decisions that we made in our model strategy was to drive model specificity and accuracy. To take it back to our business problem, we rely on these cohorts to be able to identify opportunities, but also engage in dialogue with our physicians. Our threshold for an accurate predictions is very low. We really need to strive to have extremely accurate models, and the decisions that we made in these two model strategies were both in support of that.

With those two decisions, we have now created a situation where we have a lot of models. We have a single first layer of models that just 86 projects, it's 86 models large, and those models are predicting cohort family placement. Then in our second layer of models, we have 310 different distinct cohort models within that layer. We're looking at about 400 different models within this project. We have built a model factory in effect.

Again, just want to reiterate that this is really in support of getting the most accurate model for our business case. Before I talk about how we're managing all these models, I just wanted to briefly go over what features we're using to train these models. This is the first round of features, we're already testing out an additional round of features, but what we did here is we grabbed those fields that were important to our AI engine.

That includes some patient information and includes the bills, procedures and diagnoses that are available on a surgical encounter. As well as some free texts from those operative notes. We're extracting out a portion of operative notes, this is called the procedure perform section, and we're submitting this text to model training. I'm going to refrain from talking too much about the feature engineering, but one thing I did want to point out is that we are highly dependent on-text processing.

We have found that not only do we need the texts from the operative notes, but we had a huge lift in our model accuracy if we were able to use the procedure descriptions and the diagnosis descriptions instead of the numerical codes. That was one challenge that we were facing when we started to engage with DataRobot. We weren't quite sure how the text processing was going to work, and I am very pleased that the text processing has been what we needed. It has enabled our models to be successful and accurate.

Beyond that, we train these models using two million encounters from three of our validated clients. You can imagine, even thinking about 400 different bottles. I like to think of myself as a capable data scientist, and I work with a team of phenomenal people. But creating the code to generate 400 different models, train 400 different models, do the hyper parameter optimization, do metric evaluation, receive predictions and integrate those predictions into our systems. That scope of modeling approach would be extremely daunting for us as a team.

The idea of doing that to create the best model for each one of these that we need would be prohibited. DataRobot has allowed us to do this in an iterative fashion. We're leveraging the DataRobot Python API. This allows us to initiate, upload data, kickoff projects, review the metrics, pull down predictions all within an iterative fashion. It is the reason that we're able to work with model factories in such an effective manner.

The other thing that I want to highlight here is just the diversity of models that are available to us within this

process. This is again, looking at that total knee cohort family. As we mentioned, our first layer of model, that is the model that's predicting placement within the cohort family or not. Then we have a model is trained specifically to recognize placement within each one of those individual cohorts. By using the AutoML process at DataRobot, we have access to a large repo of different data pipeline processes and models.

Not only do we have access to that repo, it will also generate a blender models where it's incorporating multiple different model results into a blended model that's available. And we're able to diagnose and look at and evaluate the model performance on a leaderboard of sorts. This brings that recommended model to the top of the leaderboard and we can interact with that through the Python API. I just wanted to give a snapshot of the different models that came out of this process for the total new cohort family, and really recognize that this allows us to find, identify, train and work with a model that is going to be the best model for each of these individual cohorts.

What's the ROI of this project? In our past state, the way in which we maintained our AI engine, and by that I mean, how did we engage with scaling our rules-based engine and maintaining integrity of our rules-based engine? We did this through a manual process. We sampled encounters over every cohort from every client. We had those sampled encounters reviewed by expert clinicians. We had a change management process, and then we stored those reviewed encounters in a database to monitor accuracy over time.

This is a very intensive resource process. Even as much effort as we were able to put into it, this is still only sampling a fraction of our cohort populations. Now, in our future state, the way we're going to bring this into our process is that once we have the trained machine learning model factory, we're able to predict cohort placement of every encounter from every client that we have received. We then compare the machine learning predictions back to the AI rules-based placement. And we provide a report to our clinicians as aggregate report.

Our clinicians then can quickly review that fallout report and identify places for rule changes that can improve the accuracy of our AI rules-based engine. This again goes through clinical review and change management, but at the end of this, and we're constantly retraining our machine learning models, we are able to get a continuous and comprehensive accuracy monitoring out of our machine learning models. This is a lift above and beyond just reducing the number of work hours that we needed to maintain the integrity of our cohorts.

I want to take us back out of the weeds up to that high level view of why are cohorts important to us and how they further our mission? Grouping encounters into apples-to-apples comparisons allows us to have clinically meaningful insight, both in opportunity identification and also in dialogue with our clients and our physicians. On the left-hand side here, you see a list of the service lines that we cover and the number of cohorts that we have in each of those service lines. With a deep dive snapshot into the spine service line, where the cohort families there and the individual cohorts are oftentimes focused on spinal fusion surgeries.

But how do we take these cohorts and translate them into opportunities? One way that we do that is by looking at surgical supply variation. This figure here is showing the list of different surgical supply categories that are needed for a successful total knee replacement surgery in the different phases of surgery. We're going from procedure prep all the way to patient closure. Underneath each of those phases of surgery, we have the supply categories that are necessary at that time in the surgery.

You can just see from the snapshot here, that there is a lot of different supplies that are required for a successful surgery, and these are the categories. What we receive from our clients is a list of every single supply that manufacturer information and the cost of all of those supplies that are used in the surgeries. The first thing that we do once we receive this data is to categorize those supplies into comparable categories.

These categories have been developed within Empiric using clinician input, and they allow us to identify variation in categories where we can find opportunities. What I've highlighted here is the variation that was observed for a health system within that total knee replacement cohort of bone cement. We found encounters where they were using bone cement at \$50, we found encounters where they were using bone cement up to over \$500.

Once we have this view of this surgical supplies that are occurring in these surgeries, we can go about finding opportunities. This is a list of opportunities that we found. We were able to implement multiple different clients. Up at that bone cement opportunity and working with physicians, we found that if we could reduce the use of bone cement from two packages to one for standard primary total knee replacement, we could save on average \$250 per case. Looking at the volume of those surgeries at a health system, this amounts to somewhere around \$625,000 annually.

These supply opportunities really stack up and make a large impact. Again, going back to that critical conversation with the physicians, being able to surface a comparison of what supplies a physician is using relative to their peers within a health system is extremely powerful. In many cases, physicians have never had access to that data before, so this allows for physicians to have conversations with their peers about, oh, you're only using two packages, or how are you being so efficient with your bone cement? Those are hypothetical. But that we have seen many physician dialogues where it is a conversation that really highlights some opportunities that physicians are ready and willing to engage with.

Andy:

Megan, just wanted to jump in and remind you that we're about 9:34 right now.

Megan Bultema:

Perfect. Thank you. One more high-level and then we'll summarize here. Above and beyond looking at supply cost variation, another thing that we started to look at as a company is the total cost of care for a patient. We've done this by integrating claims data associated with a patient's care. By using the claims data, we're able to look at treatment patterns that occur 30 days before the surgery, during the surgical encounter, and then up to 90 days post discharge. This allows us to not only look at the treatment patterns that are effective, but also the outcomes for the patient.

If they have to come back to the hospital or have an additional procedure because of a bad outcome, we're able to gather that in this claims data. And below here I'll let the audience take a look at few examples that we've found just in logistical processes that could save a considerable bottom of money. These are very significant to hospital systems who have moved from an at-risk cost model into... Sorry, have moved from a paid-for-service cost model into an at-risk patient population. But it's also extremely important for the patients. I think probably a lot of us on this call have a high deductible health care plan.

If one of us has to receive surgery, we want to make sure that I'm getting the best value out of our care is very significant and important to the patient population. This is our final slide. This is a list of some of the machine learning use cases that we've been working on over the past seven months through our engagement with DataRobot. We have kicked off over 4,000 different DataRobot projects.

We have initiated and trained over 120,000 different machine learning models. We've done this with a phenomenal team of female data scientist and we're supported by a data analyst and a data engineer as well as

a phenomenal company. I want to thank everyone for listening to this webinar and happy to take those questions now I, Andrew.

Andy:

Great, thank you so much. Really impressive amount of work that you and your team have done to solve this problem in a creative way. We do have one question from the audience. I'm not quite sure if I completely understood it but I'm going to try to summarize it. If I get this wrong, please feel free to correct me in the Q and A window.

But I just to try to summarize this, has this approach or what you've implemented here, has it reduced workload at all to design these rules or in any part of your process? And how have you merged human inference with the ML approaches that you've employed? Have you replaced human inference with the inference from machine learning models?

Megan Bultema:

Good question. Yes, it is drastically reducing the amount of work hours that we spend going through this. I just want to flip back to this slide. Before we were doing that sampling approach, we were growing a significant number of resources at doing this for every new client. Through the new process, what we're doing is we're not replacing the clinical inference of the clinical insights, but we are providing our clinical analyst a place to respond to a suggestion.

We take those machine learning predictions and we do a fallout report between the machine learning prediction and the rules-based placement. Then we're able to suggest what might improve the accuracy of the rules engine. We provide that list of suggestions to our clinical analyst and they are able to look at that and provide the expertise that we need to make sure that we are forming our rules-based engine in the way that we want. Then on the flip side of that, we can also use that clinical analysts review to retrain our models.

We're able to constantly improve our models by pulling in that clinical review and going through processes of integrating that feedback loop into our machine learning model training. Any other follow-up questions there, Andy? Do you think that answers the question?

Andy:

Yeah, I think it did. Thank you. We've got a couple more that have come in. Can you explain the nature of the toolbox for the NLP that you mentioned? This is going to be mostly within DataRobot, but we did get that question [crosstalk 00:38:42].

Megan Bultema:

Yeah, happy to talk about that. There are multiple different model pipelines in DataRobot that enable text processing. A lot of them will generate your engrams and provide the matrix that you need for text analysis. Then it will leverage that information much as you would in your own model. A couple of other additional things that I've noticed that DataRobot can do, and I alluded to this when I'm talking about using the text descriptions for our codes, instead of using the discrete data element codes, there is some enhanced feature creation that is available for texts comparison.

You can do co-sign similarity scores between different texts data elements or features and create new features on top of that that describe these two codes were very similar, or their difference was greater than this, and that can inform model accuracy as well. There is a lot available in the DataRobot repos. I definitely would encourage you to look at some of their white papers on their text processing. In many situations it was similar to what I had done in manual processing of text, and we found it to be very effective.

Andy:

Perfect. Thank you. Another question that we just received was, are the broad concepts you're describing/using being used in other healthcare areas besides surgery? And if so, which ones, and are there differences due to the specific area of application?

Megan Bultema:

Yes. There are definitely differences in how you need to approach your surgical populations versus other populations. We have started to engage a bit with populations outside of the surgical space. But really just in those initial discovery projects. We do think that overall, some of the methodology we've applied here is going to be applicable outside of surgical space. Just the approach of engaging physicians as you are building your processes and pipelines from the beginning, I think is very powerful and can't be understated.

I think as data scientists, a lot of times we receive a dataset and we go about discovering and doing what we can. But I think that the alternative approach of first engaging with the physicians, the clinicians, and asking them what is relevant to their approach, and then determining how to support that from a machine learning or even an artificial intelligence perspective, is a much more effective approach. I think in terms of that, a lot of our underlying methodology for forming cohorts and then being able to facilitate comparative analytics on top of those cohorts is something that we believe can be translated outside of the surgical space.

Andy:

I see. I think that's a good segue into another question regarding working with your partners and clinicians. What has been the reception of this approach with your clinical partners?

Megan Bultema:

I think that we're just starting to get the message of our ML-informed AI process out there. Now, in terms of our rules engine and the cohorting process in general, the reception has been phenomenal. There are so many situations where we have shown what we're doing with cohorts, how we're creating them, how we're monitoring them, and the physicians, I think, really feel and express that we're approaching it in a way where it's meaningful to them, and it really is exciting for them to be able to have access to this data and this process.

I think many physicians are scientists as well, and they are intrigued and interested in how they can use this approach to answer the questions that they've always had as well, and maybe haven't had the ability to answer. Many times when we present this work, a physician will say, you know what I've always wondered is does this improve patient outcomes, or is this approach better than that approach? Could we look at that with your cohorts? Which is always exciting and always very encouraging.

Andy:

Awesome. Maybe one last quick question before we wrap up. Has this problem, or how has using DataRobot maybe changed your thought process on what ML projects you pursue and how you approach them?

Megan Bultema:

That's a very good question. The ability to use a model factory has allowed us to take on these really massive projects and still maintain the accuracy we need for our business case problems. We're using a model factory for this project that I described. We're also using a model factory in predicting those supply categories. It's actually even larger model factory. But I think also just in general with DataRobot, one of the big benefits that we're getting out of it is the ability to test and iterate and do a proof of concept extremely quickly.

All of these ideas that we built up over time of, I think we could do this, or maybe we could approach it that way, or we could improve our process, just over these seven months, we have tested out a majority of those and it's been really helpful to get that fast iteration of ideas and be able to test many different approaches and then find the right path for a project to go forward. There are many of these projects we would not have done without the DataRobot platform, and all of these projects, we would not have been able to complete as quickly as we have without the DataRobot platform.

Andy:

Sounds good. Well, let me just say thank you so much for the presentation today. It was really interesting. We got a number of comments, not questions but just comments coming in saying, just thank you, very interesting and insightful presentation, so appreciate your time in doing this. And thank you everybody for joining and listening in. I appreciate the time, and we'll wrap up the webinar now.

Megan Bultema:

Thank you.