

Aaron N.: Thanks Amy and thanks everyone for joining today. We are very excited to announce the public introduction of Touchstone. Let's go ahead and get started and I just want to begin with an opening question and that question is, just something I'd like you all to think about as we have our discussion today. The question is what does your organization need in order to get better at opportunity analysis and strategic planning? There's generally two types of things I like to think about that organizations need to be good at, when they're doing opportunity analysis and strategic planning. When you're trying to figure out where you need to focus, you need quantitative information that shows you where your opportunities are and you need ... also needed to bring to bear qualitative information.

The qualitative information could be things like, where do I have teams in place to execute on improvement projects once I've identified an opportunity. As we go through today and we talk about Touchstone and how maybe Touchstone can help your organization with the quantitative side, I just want you to be thinking about how ... what your organization needs to get better at. Just a little bit of background, Touchstone is an idea that has been around Health Catalyst for a long time. Really it started several years ago when Dale Sanders, our president of technology started thinking about what value we could provide to the Health Catalyst customer base if we were able to aggregate data from across customers into a central data repository?

What products could we build on top of that data and we launched a product line two years ago to start tackling this question and initially that product line was called CAFE. We recently decided to change the name to Touchstone because CAFE, it was an acronym that stood for Collective Analytics For Excellence but we ended up ... we kind of dropped the full name and started calling it CAFE. What Touchstone is, is it's a next generation benchmarking product or what CAFE was and now Touchstone is, is a next generation benchmarking product and the word CAFE didn't really mean anything to do with benchmarking versus Touchstone, literally means a standard or a criteria that you can measure yourself against.

We like the name and we ended up changing it to Touchstone. Two years ago and we started the product line based on Dale ... Dale kind of pushing us into this space. We identified four customer development partners to start figuring out what product or products we would initially build on top of ... if we were able to aggregate a very large, very detailed healthcare data set from across our customers and from external data sets. Those customer development partners are Allina Healthcare in the Minneapolis area, University of Kansas Health System, MultiCare in the Seattle Tacoma area and Partners HealthCare in Boston. I just ... I really, just wanted to take the opportunity to sincerely thank these customer development partners for their help.

It's been endless conversations and interviews and pestering for me, trying to really make sure we build something that's useful to them and our initial interviews with them resulted in a number of ideas. We could build a benchmarking product if we had this data set. We could build a machine learning platform, a clinical decision support tool based on pattern recognition, looking across the very large data set to develop concepts like patients like this. There were really a lot of different opportunities in terms of what

products we could develop but what everybody said was the low-hanging fruit, what all these customers said was the low-hanging fruit is benchmarking.

That's really where our initial focus has been so as we started to hear from our customers, "Hey we think you should focus on benchmarking," we tended to agree because Health Catalyst's historical approach for conducting opportunity analysis is very internal looking. Health Catalyst had ... our approach kind of comes out of the inter-mountain doctrine, which is you want to identify ... if I could sum it up very succinctly, you're basically trying to identify care processes where there's a lot of clinical variation. The idea is that if you find care processes that have a lot of variation and you're able to develop a shared baseline that says, "Here's what the standard is."

Then measure how that variation compares from that standard, then you're able to understand which variation is good and results in better outcomes but where variation is also bad and results in worse outcomes and then try and eliminate the bad variation and improve outcomes. That is very internal looking though and doesn't take into account how you compare externally. Benchmarking was really a gap in how we looked at opportunity analysis in the past so that's another reason we were ... we thought that, "Hey, maybe our customers are onto something. We need to listen to them and incorporate this into our opportunity analysis tool set."

Before we delve in, we wanted to do some additional market research to really validate the importance of benchmarking in terms of opportunity analysis and if you look at the benchmarking market, it's kind of been around for a long time. There are a number of established players that have been doing it for probably close to three decades and so it does play an important role in the healthcare analytics industry. We did a survey of roughly 125 healthcare executives and we found that almost 90% of respondents report benchmarks as either very or extremely important to outcomes improvement prioritization.

Overall when we ... we look across a number of factors to see which are the most important in prioritizing which types of areas you're going to focus on in terms of improvement, benchmarks were the second most important factor and the only factor that was more important was regulatory factors. For better, for worse, if CMS says they're going to penalize your health system for readmissions, you're going to focus on reducing readmissions regardless of whether or not you're already performing well in that area. This sort of validated the need to go and build a product that was focused on benchmarking. That really generated the start of Touchstone. I'll pause for a second, we'll ask the first poll question. I'll turn it over to you, Amy.

Amy Barone: Great. Aaron, the first thing we'd really like to know today is how satisfied are you to our prospective audience, with our existing ... with your existing benchmarking or opportunity analysis tools? Please answer one of the following. Number one, unsatisfied. We have seen zero actionable improvements. Number two, somewhat satisfied. We have problems but some improvements. Three, satisfied. We see steady improvements and an ROI. Number four, extremely satisfied. We have significant improvements and five, not applicable. For example, you don't have currently a benchmarking solution. Please choose one of those answers. In just a moment, we will

display those answers on the screen so everybody can see kind of where our audience falls.

Okay, great. Thanks for answering so 40% said not applicable so they currently do not have a benchmarking solution, 39% said satisfied, they do have some problems but some improvements, 12% satisfied, we see steady improvements, 6% unsatisfied and 3% extremely satisfied, so that's interesting.

Aaron N.: Yeah, the thing that I find interesting about those results is when we interviewed the four customer development partners and some of our others ... from the Health Catalyst customer base to really try and dig into the problems with the existing benchmarking tools, we found a number of problems, which I'll get into in a second. Most customers, we're getting some value out of the benchmarking tools that they had so this somewhat validates that. The other thing is all of our customers that we're working with had some benchmarking tool so the majority here that actually are reporting that they don't have an existing tool is pretty interesting to me.

Let's get into some results from our analysis. We did a lot of research going back to 2016 when we are first looking to build Touchstone. We tried to figure out, where are the gaps today with these existing tools and the thinking is that, if we can improve upon those things, that might provide value for our customers. The first gap is really, when you look at the incumbent benchmarking tools, how it typically works is, somebody at your health system is responsible for packaging up the data, your data and submitting it to the benchmarking vendor. Essentially what happens is you get a file spec, a specification that says, "Here's the file format your health system must submit their data in."

Then you basically package all your data together and then electronically send that over and if it doesn't meet some requirements, they run a bunch of automated checks in your data. If it doesn't meet those checks, they kick it back to you and it's basically a very manual error-prone process, that you have to manage on a monthly or quarterly basis just to get your data over to the benchmarking company. Then, they scrub your data and they send your own data back to you, along with benchmarks and we thought, "Boy that just seems like something that could be automated with software." That stuck out to us. Another opportunity that we saw was the user interfaces are a little bit awkward.

I like them into self-service SQL so you're kind of ... it's a lot of clicks to try and get at any particular answer so we thought that, that could maybe be improved upon. Third was, and this relates a little bit to the awkward user interfaces. Users, when they're in these tools, they have to really manually search through a lot of benchmarks to find out where they have opportunities. The tools aren't that great at recommending opportunities to the user. Four is once you have an opportunity, once you know what your opportunities are, when you're trying to dive into a specific opportunity to understand what factors are driving that opportunity, I call that the why, why do I have an opportunity here?

That seems to be very difficult with these tools as well and usually, what happens is when we taught ... when we interviewed users that use these tools to find

opportunities, we then ask, "Okay now, what happens once you know you have an opportunity?" This is where the really manual process came out. They typically form a work group to try and figure out what they should do to improve. Maybe you have an opportunity to improve readmission rates for your heart ... for a heart failure cohort. You form a work group, the work group meets. In that first meeting, you come up with the hypotheses about what you think is driving the poor performance on a particular metric.

Usually what happens is you end up sending a data analyst or a couple analysts off to get data to either prove or disprove your hypothesis. Maybe your readmission hypothesis is, "We are discharging people to, we need to improve care transitions. We're discharging people to rehab facilities or centers that have high readmission rates and we actually have access to other facilities that have better readmission and care transitions so maybe we should focus on those." For example, this type of analysis where you're trying to figure out the factors that are driving your opportunity, tends to take weeks or even months where you're sending an analyst, your analyst off to gather data.

You bring the data back to the work group next week, you look at the data. You tweak your hypothesis based on that information. You send the analyst off again to get more data and so, we just saw this ... we thought, "Wow, if we can shorten the the time that it takes to do this hypothesis testing, we could really improve this whole improvement process," because you can't really start improving until you understand what factors you have to change. It really stood out for us. The fifth thing is, a lot of these tools tend to focus on narrow data sets and they tend to be fairly inpatient focused. We looked at that and we said, "You know with the transition to population health and and value-based purchasing, having data sets across the full care continuum could be super valuable." With that, I'll turn it back to you Amy, for the second poll question.

Amy Barone: Great. Let's take a look, so what is the main problem with the benchmarking opportunity analysis tools you have used? Again, to our audience, if you'd answer one of these questions please. One, non-intuitive interfaces that are difficult to navigate. Two, difficulty finding improvement opportunities in my data. Three, inability to drill down into the data behind recommendations. Four, incomplete data, limited to narrow data sets for the inpatient audience. Five, mistrusted risk adjustment algorithm. I'll go ahead and leave those questions up for just a moment and then we'll share the results. While we're waiting please just take a moment and ask any questions you may have, open your questions pane and just type in your question.

We'll be answering those at the end of the webinar. Okay, so we have the poll results number one is 28% difficulty finding improvement opportunities in my data, which is kind of one of the things you mentioned as a concern, 24% said incomplete data so they had limited data sets, 19% said non-intuitive interfaces, 15% mistrusted risk adjustment algorithms and 14% said inability to drill down into the data.

Aaron N.: The thing that jumps out of me here is, when you use them to figure out where you have opportunities, it was just like kind of looking for a needle in a haystack. That 28% difficulty finding improvement opportunities in my data, speaks to that a little bit so I'm

glad to see we're on the right track there. The one thing that also jumps out at me, that I'll comment on is 15% said that the biggest issue was not trusting the risk adjustment algorithms and this is feedback I didn't listed in the major problems that we identified. There were some other major problems too that we didn't identify or that I didn't go through. This was kind of one in the second tier problems.

I'll walk through this a little ... I'll approach to solving this problem, a little bit later so I'm excited to get into that. Just real quickly, here's Touchstone's approach to these ... the five problems that we just went through. One is we created a software to automate the data feeds so there isn't any human involvement when it comes to sending the data from your data warehouse to Touchstone. Two, we've tried to build a very intuitive user interface, something that enables quick iterative analysis. Three is, and this speaks to the poll question, the difficulty finding improvement opportunities. When we get into ... I'll show some screenshots of Touchstone. We've built a recommendation engine that automatically surfaces opportunities to the user and I'll talk more about that in a second.

Four, we have algorithms running and in conjunction with our user interface, that we're aiming to automate quick iterative hypothesis testing. When I think about really the problem that I hope Touchstone solves, it's two things. It's identifying what are my opportunities, I call that the what. That's what we hope the recommendation engine does and the second is improving the cycle time it takes to understand the why. We feel like when that work group gets together, to try and understand what they need to do to improve heart failure, if instead of sending that analyst off to do, to gather data and do analysis, to bring back to the work group next week, if they can pull up a tool that already has their data in it to drill in, to get to, "Hey, it's this subcohort of patients and they're not getting this medication."

"This other subcohort of patients isn't getting a lab that contains some important information and everybody else is." Whatever that information is we feel like if we can speed up that cycle time, we've just ... we've gotten to the change that needs to happen faster and improvement work can then get started as opposed to just waiting for data. The last one is, we've included data sets across the care continuum in Touchstone. I'll show you in a second which data sets we're starting with but some examples are clinical data from your EMR, claims data, cost accounting data and patient satisfaction data.

In this slide, what I'm trying to convey is Touchstone is really a framework application. With other benchmarking tools, you might have to have, one tool for inpatient, one tool for population health, one tool for supply chain. Touchstone brings all that together and we essentially have different modules in Touchstone. We're launching with an inpatient module that runs on your clinical EMR data plus your cost accounting data and a population health module that runs on your claims data. Over time, as we listen to customers to understand which use cases are most important for them, we'll add additional modules. Some examples of those are a regulatory module that runs on a combination of claims and clinical data and cost insights which we're really excited about.

It runs on clinical data and it also runs on CORUS, our activity-based cost accounting system. You can actually get down to an apples-to-apples comparison of true activity-based costing and compare yourself to others that actually have that software as well. Now, I'd just like to dive into a few screenshots of Touchstone, to kind of give you a sense of the recommendation engine and the user interface. This is a screenshot of the home page and this is where users, when they first log into Touchstone, will land. As a user, I notice a few things here. I see a bunch of these cards and what these cards are, each card is a different recommendation.

The recommendation is a cohort of patients, in this case, coronary artery disease at one of my particular hospitals, Promontory Hospital and I see it's an inpatient recommendation. Touchstone on the home page is looking across all your different types of data sets to make recommendations. Sometimes I have an inpatient recommendation, some are population health recommendations and as we add additional modules, we'll surface recommendations for those modules as well. The red text is the metric that is leading our algorithms to recommend this cohort. You can think of it as the metric that is primarily driving the opportunity or the metric where I have the most area to improve in. In this coronary artery disease cohort, my mortality rate, compared to risk adjusted benchmarks is a 1.9.

My mortality rate, you can interpret that as a ... it's 90% higher than expected on a risk-adjusted basis. Over time as a user uses Touchstone and we see that, "Oh you're focusing on certain types of patients," maybe you're a clinical analyst that's only focused on cardiovascular, we will learn, just like Netflix learns which type of movies you're interested in. We learn what you click on. We capture that information and we filter the recommendations so we're showing you things that are relevant for your clinical area. As a user, I can drill into any of these recommendation cards, to learn more about how Touchstone recommended it.

One more thing before I go on, we have different types of recommendations. We call these recommendation categories. This all metric opportunities is one category, outlier opportunities is another category. What all metric opportunities is, is we're essentially looking across all of your metrics simultaneously and slicing and dicing your data in order to find cohorts where when we look at all your performance on all metrics, where do you ... on a risk-adjusted basis compared to benchmarks, where do you have the greatest opportunity? Outlier opportunities tries to look for cases where within a cohort, you have a few providers that are performing well and a few that are not performing well or you have a few departments that are performing well but not performing well.

There's other recommendation categories that I'm not showing right now. An example is individual metrics, where we only show one metric at a time. We're going to continue to try and listen to users to figure out which are some good recommendation categories to roll out in the future. Some examples are, that we might roll out in the near future are try and find cohorts where the length stays low but the readmission rate is high. People are leaving the hospital a lot faster than expected but they're also coming back to the hospital more frequently than expected. Another example is high readmission rates from skilled nursing facilities or from rehab centers because there might be an

opportunity there to target care transitions to ... or target patients to certain facilities and thereby reduce readmissions.

If a user drills into any particular recommendation, they go to the prioritize page and you can actually just click on prioritize here at the top and that'll take you at the prioritize page as well. What the prioritize page is, is it's kind of like an Excel pivot table. I have metrics across the top, you can see I'm in the inpatient module right now and I'm looking at ... I have a few different metric categories, financial-variable direct cost metrics. I have outcome metrics and along the left I have ... we call this the group buy and it's all the dimensions in your data. Right now, we're grouping your data by care process but you could group your data by DRG or gender or any of the dimensions in your data.

You can look at your metrics a few different ways. In the outcomes metrics, you can look at them as observed over expected, meaning compared to risk-adjusted benchmark or you can just look at the observed value or you could look at them as opportunity, which you can think of as like the number of days that you could avoid for length of stay, if you were performing as expected. The number of readmissions you could avoid, if you were performing as expected, relative to benchmarks, et cetera. We also have this thing called the opportunity index and the opportunity index is just a formula that we came up with, that essentially looks at all the metrics that you currently have selected.

You can add or remove any of these metrics, if you only want to look at length of stay for example, you could close out all the other metrics. The opportunity index, based on whatever metrics you're showing, you can use to sort the table. You can use it as a global sort. I could sort the table by length of stay, just to find the care processes that have the highest risk adjusted length of stay but if I want to take all my performance on all metrics into account at the same time, I can use the opportunity index to sort the table. The explore page has a few different graphs that we give the user to help drill into their data. Essentially, it's trends over time and stratify. With trends over time, I can look at ... I'm able to look at my financial metrics or my outcomes metrics.

Right now, we're looking at variable direct cost of supplies and mortality. I can see my performance for the current year and the past year. For the stratify, I can actually pick the dimension I want to stratify this metric by. Again, we're looking at variable direct cost of supplies here but I'm stratifying by provider. I can see that some providers have higher cost and more variation than others. Let me go back. In the outcomes section, we have risk-adjusted benchmarks and so you can see your performance compared to those benchmarks. The gray ... The bars in each of the charts in the outcomes section are the observed metrics. The blue dots are the risk-adjusted benchmarks and the lines are the observed over the expected.

These two things are screenshots of patient level detail and actually, how we think we're going to solve the problem of not trusting the risk adjustment models or at least one of our solutions. What a user can do is they can ... Once they've identified a cohort, they can drill down to the lowest grain of the data and they can see the individual ... Let's see, if I can move my mouse without advancing the slides. They can see the individual expected values for each patient. Right now, for each discharge, I see these

patients all died and here is their expected probability of dying. We've sorted this in ascending order so the patients at the top here had a very ... are risk models flagged as having a very low probability of death and they actually died in the hospital.

Those are kind of interesting results but it's a little bit black box, I don't have the ability to drill in and understand how the risk model is arriving at these probabilities so that's where I can click on any of these expected values and drill in ... and right here, I'm surfacing exactly how the risk model arrived at that 1.7% so here are all the features the risk model is taking into account. I see that this person is less than 60 years old. They never had a heart attack or cancer, they never had diabetes. The only thing that they had, that the risk model was able to use to predict their probability of death with was uncomplicated hypertension.

There's kind of two things that could be going on here. It's either a documentation problem or the risk model is not doing ... picking up something that it should and maybe it's an opportunity to code differently or maybe it's an opportunity to actually improve the risk adjustment model so we can risk adjust better. The second type of opportunity is, it's actually an opportunity to improve care in some way. In either of those cases, we think those are interesting opportunities that we would want to act on. Just a little bit more about risk-adjusted benchmarks, I don't want to get too technical here because we can really spend a lot of time on risk adjustment.

Just a few comments and if any of you are interested, I'd love to have a follow up call to really dive into this in detail and get your feedback. We ended up building our own machine learning models to create risk-adjusted benchmarks and the techniques that we're using are some of the more modern day machine learning techniques so random forests and deep neural networks are some examples. We're actually using a lot of the same features, that when we ... Risk adjustment has been around for roughly 30 years and the features ... and it really started with claims data and the features for risk adjustment that are commonly used haven't changed a whole lot over time.

It's administrative data, its clinical data and it's demographic data. The thing that we're really excited to start doing in 2018 is to start to go into two other areas that we think could really improve the accuracy of our risk adjustment algorithms. One of those is socio-economic data and another is mining clinical notes because there's just a ton of information in clinical notes, that is frequently not coded and you can really change the disease picture for any patient, if you're able to effectively extract the information from those clinical notes. The last point I'll make here is, there's two ways you can get at these benchmarks if you have access to Touchstone.

One is going through and using the tool that I just walk through some screenshots of. The other way is, we actually route your expected values back to your data warehouse so you can consume in them any application that you want, whether that's a health catalyst application or any sort of reporting you want to do or a third-party tool, they're yours to do with what you wish. Just a little bit on the data that we have in Touchstone, currently we have about 125 million patients in Touchstone. It's starting to be a pretty large data set, that translates into a little over 640 million patient visits and if we look at

all the diagnosis codes, all the procedure codes, all the medications, the labs and sum up all those facts, we have over 300 billion patient facts.

It's starting to be a pretty large data set and when we project out how ... Well, I don't want to say how big we think it could get but we're excited that it has a lot of room to grow yet. Importantly, we've purchased a few third-party data sets to include as well as our customer data in the Touchstone data set. The three that I'll just mention real quickly are HCUP, which stands for the Healthcare Cost and Utilization Project. It's a fairly detailed data set that's a ... it's a 20% sample nationally of all payer hospital visits. The second is Medicare claims so Medicare has a lovely data set called, their standard analytical files limited data set.

We've purchased 100% of the Medicare Part A claims. The last is we started doing some research on some state all-payer claims databases and the first one that we've purchased is the Utah all-payer claims database, basically because it was the most detailed and it gave us the most relevant features to do our risk adjustment with. Just real quickly, there's another exciting announcement that we want to make today and that's ... that we're actually making a free version of Touchstone available to all health catalyst clients and we're calling this Touchstone Public. It's a limited edition version of the software running on some of those third-party data sets that I just mentioned.

Regardless of whether or not you buy Touchstone, you can get Touchstone Public for free and the fun thing about some of these third-party data sets that we're using at Touchstone Public is you can actually drill in and see yourselves, like in the Medicare data, you could find your health system and drill in and see your performance on the Medicare data set. Hopefully, even if you're not able to get Touchstone right away, you can start using the software and really see if it's useful to you and hopefully, get some value out of it just on the third-party data sets that we've put in there. Then Touchstone is ... the full version of the software is what we call Touchstone and that's running on your data and all the other customer's data.

I just want to highlight two ... what our two main focuses areas are for the rest of 2018. The first is really to improve our recommendation engine. You heard me talk about that a little bit. It's trying to come up with recommendation categories in algorithms that are consistently surfacing up personalized recommendations that you're interested in. This kind of goes back to the ... I forgot if it was the first or second poll question but finding opportunities in the existing tools is difficult. We want to automate that as much as possible so that, we know it's not the only factor that's important for deciding where you focus your improvement opportunities, like there are qualitative factors that are very important.

If you have an opportunity for heart failure but you don't have a team and a work group that's ready to start improving and making change, if you don't have that organizational structure in place then your improvement isn't likely to be very successful. You need both but when you're trying to think of, "Gee where do I need to improve," with the existing tools, you're kind of flying a little blind just because there's so many insights buried in your data and crunching data isn't something that humans are particularly

good at but it is something that computers are particularly good at. We just want to automate that as much as possible and the same thing for the why.

In the second half, we're going to really focus on ... after we've really nailed the what and we think we're on the path of that, we want to really drill into the why and enable that quick iterative hypothesis testing. The work group can sit in the room and use Touchstone to really validate or disprove hypotheses so they can decide, "Do we have an actionable change that we can make?" Just a couple closing comments. Hopefully, this has been helpful in thinking about just generating some ideas for you, as you think about what your organization needs to do to improve from an opportunity analysis capability set or from a strategic planning standpoint.

We are very interested to know if you think Touchstone can help your organization with the quantitative side of that. I'm also really interested to hear feedback. One of the things we tried to do a good job of is listening and the thinking is, if we can really listen to you and really hear a feedback, it decreases the chances that we're spending a lot of time and energy building things that customers don't find useful, which would be the worst case scenario. If you have feedback for me or you're interested in Touchstone, feel free to contact me. My email address is there on the screen and with that, I'll turn it over to Amy for our last poll question and then, we'll open it up for Q and A.

Amy Barone: Great. Again, before we start this next poll question, please submit your questions. We do have a lot of good questions coming in. We'll get to those in just a moment. The final poll question for today is, are you interested in learning more about how you can use Touchstone at your organization? Please just a quick yes or no. Yes, I'd like someone to contact me to learn more about Touchstone or no, not at this time. Okay, so we're going to go ahead and address some of the questions that have come in, lots about benchmarking and where those standards come from and how the data was created initially because remember a lot of this audience doesn't have a lot of experience with benchmarking tools.

Aaron N.: Right.

Amy Barone: We'll talk about ... so, one of the first questions, can externally acquired benchmarks such as those purchased from the benchmarking companies be incorporated?

Aaron N.: It's a really good question. It depends a little bit on your data use agreements, with the other vendors but we actually have customers that have come to us and say, "Boy, like we love Touchstone's user interface. We like the algorithms and we like the direction you're going in terms of the problems you're trying to solve with this tool but I also have this other benchmarking data set that I really like, can I just plug that into Touchstone and then take advantage of the user interface and the recommendation engine?" The answer is yes, it's a qualified yes. It really depends on your data use agreements with your other vendors. What we've seen so far is that vendors are ...

The data that they send back to you, it's kind of yours to do what you want with but you can't ... we wouldn't be able to take it and then try and do anything like figure out their

risk adjustment algorithms or anything malicious like that. All we would be doing is kind of visualizing it, in Touchstone for you and the few customers that we've explored doing this with so far, their data use agreements allow for that.

Amy Barone: Okay. This one is kind of a follow up. Can you tell us more about the cohorts used for benchmarking? For example, how many hospitals, how many healthcare systems, any option to benchmark against small rural hospitals, so how can you segment the data so that you're happier with the data that you're comparing against?

Aaron N.: Yeah. Between the third-party data sets, we've purchased and the customers that are submitting data, we have coverage for the majority of the hospitals in the United States. For each metric that we're benchmarking, we're creating separate predictive models for that benchmark. We'll create a benchmark or we'll create a predictive model for length of stay for example, for teaching hospitals. We'll create another one for non-teaching hospitals or for rural hospitals, to try and get to more apples-to-apples population.

Amy Barone: Okay. This was pretty specific. Looks like you were using a financial definition for length of stay based on midnight census. Is there an option to calculate length of stay from admit, DT and discharge DT?

Aaron N.: Maybe I'll answer the question a little bit more generally because I think this is a good question that a lot of people probably have. It's ... If I could restate the question a little bit more generally, it would be, do I have an opportunity to add custom metrics or custom dimensions based on my own specifications to Touchstone. This is another one of those secondary limitations that we've found when we did research of the existing tools. Basically, the existing tools give you ... you submit a specific data model or you submit your data in a specific format and then, that's ... you get their metrics definitions back and that's the end of the game.

One of the fun things about Touchstone is we have the ability to extend our data model a little bit. If you're familiar with Health Catalyst at all, our standard data models are what we call shared data marts. They're the content that we ... They're how we infuse your data with content so that we can actually get insights out of it. There's the ability to extend those shared data marts. As long as you want to make a different definition for length of stay, to answer the question a little bit more specifically, you can create your own definition as long as it's derivative of the data elements that are in our shared data marts, which are pretty broad detailed data marts.

We can actually benchmark that metric because we'll have every customer on the same shared data mart, data model. Even if one ... if you want to create a custom metric definition and nobody else does, we can still create that metric as long as it's within the the shared data mart data elements and then create benchmarks for that metric and and serve that up to you in Touchstone.

Amy Barone: Another question about our machine learning models so they're asking, how are you validating your machine learning models?

Aaron N.: Yeah, it's a good question. We use ... there's kind of been this ... One of the fun things about Touchstone is when we looked at the existing products, they've kind of done their benchmarking in, using techniques and methods that have been around since the early 90s so linear methods, difficult train test split methodologies to validate. In the meantime, there's kind of been this data science revolution, where statistics and computer science have come together and created this new field of machine learning and so it's opened up a lot of new models that actually have some better performance and some pretty good practices for validating whether those models are very predictive or not.

I would love to have an entirely separate call with anyone who's interested to really get into the nitty gritty on our practices, for building our risk adjustment algorithms. Just to answer this question, we're essentially using fairly standard machine learning cross-validation techniques, to make sure that there's no data leakage, to make sure that the performance of the models that we train applies to a holdout data set and then, we have an additional holdout data set that we use to check and make sure that there's no overfitting or that we're not getting performance that is not ... doesn't reflect how good the models actually are.

Amy Barone: Okay. I'm going to add a little bit of information to that, so another person ask can you address the interpretation of machine learning approach for risk adjustment? I will mention that we have recently had a webinar that talk specifically about machine learning and AI. Do go look in our archives and see if you would like to watch that webinar as well. Okay, is Touchstone using any of ACG system, Johns Hopkins?

Aaron N.: No, we're not currently, We're open to using any sort of clinical groupers that customers feel would be super helpful to them, as they do analysis. Right now, we have DRGs, we're using CMS and HHS, HCCs and we also have the Health Catalyst clinical hierarchy, which is what you saw on the screenshot, that was the care process grouper so lots of groupers. There's pros and cons with all of them and the good news is we're happy to be flexible and listen to what you all think would be super helpful.

Amy Barone: Okay and then, earlier you mentioned, remind ... Basically the question is remind me again what I need to provide to get started with. In other words, what data do I share?

Aaron N.: Yeah, so if you are a Health Catalyst customer, you would have our DOS platform, which has our data warehouse and then our shared data marts installed on top of your data. What data you need to provide is essentially whatever ... whatever data you would like to see in Touchstone, if you have an EMR and you have clinical data, you can get the inpatient module. If you have claims data from at risk contracts with payers or you are a payer, you can get the population health module. It's kind of, we meet you at whatever data you have. If you have all of those combinations, you can get both.

Amy Barone: Okay, great and someone else asks, I am interested in how or even if the competency of the healthcare team and the competency assessment of that team is factored into the data. Her statement is ... Sorry. My experience in hospital education and clinical systems administration, I realized the incredible impact of the actual providers on the hospital and primary care setting relating to the patient outcomes.

Aaron N.: Yeah, so I think how I would answer this question ... Hopefully, I'm not missing the question. Let me know if you think I'm going down the wrong path. It really gets us back to you need two things to do improvement work. You need to know where you have the biggest opportunities and what you can do about them and that's on the quantitative side, it's a little on the qualitative side, what can I do about them but it's also on the quantitative side. You also need to know where you have readiness, where you have clinical teams and providers that are willing to engage in improvement work. That's just qualitative information that is within your health system.

The good news is, we think that both are extremely important and Touchstone is going to be the ... Our services organization has recognized this. They've recognized this since Health Catalyst was founded as a company a decade ago. We're excited that our analytic services organization is going to be using Touchstone for their ... for opportunity analysis, for the quantitative side and then they'll also be helping our customers with the qualitative side as well, figuring out where they have that readiness.

Amy Barone: Okay and then this is a great question, how do you suggest that I internally build support to spend the time on Touchstone or in benchmarking in general?

Aaron N.: That is a really good question. The value proposition for improvement work is kind of ... It's kind of insidious because you have all these insights buried in your data about where you could be improving. If you don't have access to them, you don't know what you don't know. It's kind of like a leech that's eating away your organization. You have the opportunity to improve in these different areas but you don't know about them. It can kind of be a little bit of a chicken and egg problem and that I know I need to invest the time to find these things but I actually don't know where they are so I can't like prove to people that I need to invest the time to find these things.

That's one thing about automatically identifying the what, that I am pretty excited about with Touchstone, because it's more transparency right away about where your opportunities are.

Amy Barone: Another great question is, if someone is interested in a demo of the product, can you use their data from their hospital to benchmark, to see where some opportunities may exist or is it just a generic demo?

Aaron N.: Yeah, a great question. If they just want to kind of see ... if it's pretty early stage, they just want to see the tool and learn more about it, it would just be demo data but if the conversation is getting close to, "Hey, we're like actually really serious about getting Touchstone and using it for improvement work," then it changes a little bit. We wouldn't have ... You wouldn't have the Health Catalyst data operating system, so you wouldn't have your EDW with all your data in it, so we couldn't automatically pull that data from you but we would be able to say, "Hey, give us these few files and we'll put them in Touchstone. We'll give you limited access, like a 90-day access or something and you can use it on your data, to figure out where you have opportunities."

It's a little bit of a try before you buy. The other opportunity is with Touchstone Public of course. It's not as good as your data but it is actually ... you can find your institution and you can find the Medicare patients at least, for example at the Medicare data set and look at your performance on different cohorts.

Amy Barone: Okay. Someone else asked, traditional inpatient risk adjustment only uses the admin data from that admission and you're accessing a longer-term view of the patient if available and if so how do you deal with the issue that a first-time patient will have largely in visible history?

Aaron N.: Yeah, that is definitely ... It's the biggest problem with EMR data, right, because if a patient goes to your hospital once, you see them but then you don't see them bouncing around the health system, so you don't have that full continuum of care perspective. Ideally, your organization gets better at getting this claims data, setting up data use agreements with your payers to get claims data that provides that full continuum of care. I will say though, what we're doing in terms of our feature engineering, is we're saying, when we do have more than one visit, let's look back across those visits and let's not look at how the patient was treated at each of those visits but let's just figure out what comorbid conditions they had.

If we can develop a fuller clinical picture, not of how the treatment was provided to that patient because we want our risk adjustment algorithms to be treatment agnostic. We just want to adjust for factors like, "Hey, this patient was older or hey, this patient was sicker because they had a higher disease burden or hey, this patient was female," things like that that are appropriate to take into account.

Amy Barone: Perfect. Okay. For end users that don't grasp the machine learning or statistics well, do you have a user interface that visualizes a patient versus their patients historic data set and that displays the patient versus some cohort population similar to them? Then if they do, do you have third-party data sets in the models?

Aaron N.: Yeah. It's a good question. The part of that question that I keyed in on was the patient versus others like them. We think we can ... that comes into really nailing the why, we think that's important so what's different about this patient than patients like them, that resulted in them having a worse outcome, compared to the rest of the cohort. That's the user interface that we're going to start building out in the second half of this year, to enable that quick iterative analysis of hypothesis testing.

Amy Barone: Great. Another, what were some specific machine learning prediction model performance measures, AUCs as an example.

Aaron N.: Yes, so area under the curve, we also use precision recall, F1. For continuous variables, we look at root mean square there, mean absolute error, R-squared. It depends on what we're trying to demonstrate but normally, we kind of take several into consideration and that kind of helps us make ... feel better about not overfitting.

Amy Barone: Okay and you mentioned earlier some of the earlier benchmarking tools, someone ask specifically, how does this compare to Crimson?

Aaron N.: Yeah. Crimson was one of the tools that we didn't spend as much time reviewing. We did review at some. By review, I mean talk to users that had used it in the past and Crimson's tool again kind of, in our mind, suffers a little bit from the problem of where are my opportunities. The Crimson suite has several different modules so I don't want to speak too much in generalities. I think a lot of these tools do some things that are really awesome. Crimson in particular does a really good job of, I forgot which module it is, but really focusing on physicians and saying, "Here's a physician that's not doing well on X and here's the physician that is doing really well on Y."

I think that's really a strength of the Crimson product. I think again going back to identifying the opportunities of what, you're not always looking at physician specific level to identify opportunities and the hypothesis testing isn't always drilling into like, where do I have physicians that are performing worse than others either. Again, I would go back to automating the understanding of those two areas, as maybe a weakness.

Amy Barone: The competitors that exist in the market today, someone asks so why is ours better?

Aaron N.: I would point to ... To be totally honest, it's a lot of the ... at least today, the data, the underlying data is very similar. What we try and do with software is automate things that are manual. That's what software is good at, right? Software is not good at the qualitative side or like the human intuition, right? At least not yet, so where I think Touchstone is differentiated is, it tries to automate things like slicing through all your data to figure out where the opportunities are or automating that data extraction, that use to take a person's full time job or half time job, just to send the data to the benchmarking company. I think that's the direction we'll continue to go in with Touchstone, at least until we reach general AI and probably not too soon on that front.

Amy Barone: Okay, one more question. Yes, how many medical ... or I'm sorry, mental health organizations are in the benchmarks? Do you have any idea?

Aaron N.: I don't specifically have any idea on that front. I'm guessing that there are several in our third-party claims data sets. I couldn't speak to how many are in our customer data sets. I actually don't know how many mental health organizations are a part of our health system customer base either. I would be happy to drill into that and follow up.

Amy Barone: Okay that'd be great and then someone also asked, did you consider incorporating the Part B files of Medicare and why or why not and do you have a way to tell which patients are in a specific ACO?

Aaron N.: This is a really good question. I actually misspoke on this earlier. We purchased 100% of the Part A files and a 5% sample of the Part B files because that's what CMS makes available. We would love to purchase 100% sample of the Part B as well. We're using the combined Part A, Part B to build our risk adjustment algorithms for the Medicare

population. Then, we're using the 100% sample of the Part A for the Touchstone Public, our freemium product.

Amy Barone: Okay great. Well Aaron, I think that's about all we have time for today. I thank you very much. It was a wonderful event and to our audience, thank you so much for joining us today. You will receive an email with links to the recording of this webinar, the presentation slides and the poll question summary results. Also look for the transcript notification we will send you once it's ready. On behalf of Aaron Neiderhiser, as well as the rest of us here at Health Catalyst, thank you for joining us today. This webinar is now concluded.