

Nick Marko on Leveraging Health Data

A Conversation with David Harlow at HealthBlawg

March 2016

David Harlow: This is David Harlow at HealthBlawg, and today I am speaking with Nicholas Marko, who is Chief Data Officer for Geisinger Health System, and is also a practicing physician. He is the director of neurosurgical oncology at Geisinger as well. Welcome to HealthBlawg Nick, and thank you for joining us.

Nick Marko: It's my pleasure to be here, thanks for having me.

David Harlow: So Geisinger has been on the bleeding edge of a number of innovative healthcare improvement initiatives and I wonder if you could start this off by giving us some insight into the Geisinger experience with electronic health records in the context of the shift from volume to value. Have there been roadblocks to progress in that arena or any particular Aha! moments you could share with us?

Nick Marko: Sure, absolutely. So just by way of introduction – as you said – in addition to my clinical practice, I also serve as Chief Data Officer here at Geisinger. So the pieces of the puzzle that I oversee include our data management system, our data warehouse and our big data stack, our data science department that does a lot of our predictive modeling and other such analytics, and then our data strategy and data governance arms as well. So the perspective that I'm coming at these questions from is essentially as part of the team that helps to make sure that we bring all of the pieces of information together, that we've got about all of our patients, and then serves those up to the folks, whether they be the clinicians in the population health department, or the practicing clinicians who are in hospital seeing patients every day. Also folks in our health plan and other similar parts of the system that really help deliver care. So that said, when you talk about how the EMR fits into all of this, Geisinger has had a long tradition, long history with EMRs. So we were one of the earlier Epic clients, I think back in '96 was when the system was installed here and that certainly predates me – I've been here for about three years or so – but the point is we've had a big repository of data for a long time. So data has always kind of been part of our culture here at Geisinger. For a long time we were collecting it and treating it largely as part of the transactional record in helping us do business every day, and then a number of years ago we said – Well, to do things like more accurate patient predictive analytics, catalyzing a switch to a value-based system where we really have to know more than just open up a chart and see what Mr. Smith has going on today or last week or something or we really had to look at the corpus of data in a different way, one of those things that we had to do was change the structure of our system. So that's when data warehousing was sort of very popular and like everyone else, we built a relational database and it was working pretty well and serving a variety of specific projects and specific needs. Lately, as the type of data that we've been dealing with has changed as the volume, but more importantly, the variety of data has changed, and as those folks who

are on the front lines – the clinicians, the population health folks, the folks in our health plan, the push has become more and more to learn as much as you possibly can about the patients as people so that we can focus on the value of their care and the overall management of the patient. The desire has come to know more and more and more about people so that we can make sure we are really getting to the heart of the problem. And so our backend systems have evolved to have, to incorporate a variety of big data type technologies and other such things. I don't want to dwell on that, but the point is throughout all of that evolution, what I think we have been progressively coming to realize is that as the data and information space changes very rapidly, and as the demand, the nature of the demand changes very rapidly, the EMR is really not a place that you can rely on to be the one-stop shop for all things medical and healthcare information related. So these EMR systems were built some many years ago as transactional systems designed to replace paper charts and automate workflow and so we've had an opportunity to work with many of the major vendors and they are all producing a very high quality product that does exactly that. The challenge is when it comes to secondary use of data, the systems were never really designed with that in mind, some are better than others, but regardless what a lot of people have been trying to do and ourselves included for some time, is coerce a transactional system into also being a system that serves secondary use needs, comprehensive querying and the ability to feed predictive analytics and to create interactive data visualization and manipulation spaces. And that's just not what they were designed to do, so our big, I think Aha! moment, has really been an evolving thing. It wasn't like a light switch coming on, but I think it was us saying we've got a great relationship with our EHR and helps us transact our business every day. It's a critical and invaluable part of what we do but at the same time, we need some more depth and we need to treat the data like data and to find a home for the data that it can live in and that we can do all of the things that we want to do with it and not necessarily say, "Boy I'd really like to do this except the EMR can't do that", or "I'd really like to do this, how do we engineer the EMR to do that?" We are seeing the information as one thing, the EMR as a transactional system for that and our more comprehensive data management system as a place where we do all the other things with data that we want to do. So I think that's been sort of the real transformation over the last few years for us.

David Harlow: So the data management system is something that goes beyond the EMR and is collating other information about patients' conditions as well?

Nick Marko: Yeah that's exactly right. At a large system like ours – and of course there are others that are even larger – you've got multiple hospitals, you've got multiple clinics; many of these have been added over time through partnerships and mergers and acquisitions and the general scope of business. So we've got lots of different EMR vendors, systems, in various parts of our organization, we've got data stores in lots of different formats and structures, and so we really have to have some capability of being EMR agnostic when it comes to answering many of the questions that we want to answer to drive things like value-based care. And so we focus a lot on getting data out of its parent transactional systems, be that an Epic system, a Cerner system or something else, and putting it into a much more open data management stack that includes relational databasing, Hadoop, Apache-type databasing, and a variety of in-memory analytic

pathways and other such things that are all sort of designed to meet the various needs that are emerging. So we've really tried to decouple the strategy to meet the need of secondary use of data from the primary EMR systems, and we've done that by creating sort of a much broader space where data from lots of different systems, from lots of different EMRs comes together. And then also meets data from some other places, non-EMR sources, be those our transactional systems in-house, be those external data sources where we go to add additional information, be that more basic science kind of stuff or we add molecular data of our patients; all of these things are in play for us. So we need a place to bring together data from lots of different EMRs, but also lots of different other sources, and then use that as before, I mean, which we sort of ask our questions much in the same way as not just about treating a condition, but it's about treating as a patient as a whole, same thing with this, it's not just about the data in the context of the transaction.

David Harlow: Looking beyond individual patient encounters, the question is: What are the categories of things that you are looking at to build these tools to do? We can speak generically about population health management, but is there anything more specific in that you could identify?

Nick Marko: Yeah, sure --

David Harlow: You've built the system to do x, y, and z --

Nick Marko: Yeah, so population health management is a funny term because it means a lot of different things to a lot different people. And so to some extent I try to shy away from using the term because it is not specific. But essentially I think there's a couple of levels of things where we are trying to do something, you know, where we are trying to use as much information as we can to help either a patient or a group of patients. So on a large scale, or on a large level, it's looking at a group of patients who are fundamentally similar in some way and trying to make sure that they are getting uniform quality of care delivery, that every one of them is getting the highest possible standard and that the care that's being sort of doled out consistently to everybody. I think traditionally that's what a lot of people think of as the population health space, and we certainly do that. So we've built systems, for instance, to look at patients with certain chronic conditions - heart disease, diabetes - these kind of things. We have care pathways that we've put together and general management principles that we've put together that says every day I should get the following, right? So then we leverage our data systems to make sure that every patient on that list, who's got that diagnosis, is getting all of the following things on our best practices list, and when somebody misses something, we get an alert or their primary care doctor gets sent a message, or whatever may be the system of choice. And so that's really managing a large population with a consistent type of care that is what we think is the best possible standard of care. The sort of extreme opposite to that is looking at very personalized approaches to medicine and we use our systems to do that too. So more recently, the idea of whether you call it personalized medicine, precision medicine, whatever it is, for me as a data scientist it's making individual predictions. Predictions or assessments for a data set where the end is one, where you are looking at one specific patient in front of you. And so to do that what we've got to do is not necessarily think

about the population as an aggregate, but really focus on each person as an individual and learn as much about them as we can from their individual sources of data that we can integrate. And then also learn as much about them as we can as them being a representative member of one or more populations. And then from there that's where say our data science department kicks in and tries to build individual predictive models that will help us figure out for instance, how long may someone expect to survive with a certain form of cancer? What type of treatment strategy may be best among these two or three alternatives for somebody who not just looks generally like you, but for someone who has the exact characteristics of you. That's the very individualized part of things, but at the same time it's an effort that parallels the way the population health is done in many of the same technical details, it's just sort of how you ask the question it differs. So we've worked on projects across all parts of that spectrum and we've really tried to engineer a system, both the backend data system and the front end support systems for our clinicians and for our end-users, that really meets those goals at whatever level they may be, all the way from one patient to many patients.

David Harlow: How do you see the adoption of these tools among the clinicians, is it driven by different sorts of payment systems or payment incentives, are the clinicians incentivized, do you see any relationship between the two?

Nick Marko: That certainly is the million-dollar question, isn't it? I think I can offer an interesting perspective in this because I've got a foot in both worlds, right? So I see patients in practice and at the same time I help design these systems. I can tell you that it's pretty much always been my experience - whether it was in my residency training or fellowship training or as an attending physician I'm learning my own practice - that many physicians are insulated from the day-to-day pressure of the finances as it associates with this patient or that patient, and I think here it's a very similar thing. It's not like you got somebody constantly breathing down your neck saying, "How many RVUs are you generating today, how many dollars are you generating today?" And in fact that's not where the demand for these data tools comes from. So we tend to not see physicians, for example, asking us for - "Oh give me a system that can tell me in real time how many dollars I've made today" - or - "How many RVUs I've made today versus yesterday?" Right? That's not where our demand generally comes from. I think our demand generally comes from our docs wanting to do the right thing for people, for patients, and so much of, most of what we get, if not almost all of what we get, are people saying, Boy I've got this kind of patient, or this group of patients, or a complicated subset, a small number of individual patients who I really wish that I could do better in terms of providing the absolute best care possible, and I know we've got the information in our system that could help us figure out what's the right thing to do for this one patient or for these 30 patients or whatever that may be. And I feel like I just need some help accessing and putting in the right place. So I think our demand is largely driven by docs wanting to do the right thing, and then the payment and reimbursement questions are often dealt with on sort of a much larger, higher level business scale rather than the level of the individual. But I think most of our demand comes for tools that are some kind of clinical provider really wanting to do the right thing for the patients that their tasks are taken care of.

David Harlow: So you see demands coming from the other direction, dealing with payers and innovative payment arrangements and trying to drive behaviors in a certain way?

Nick Marko: Yeah absolutely, I mean the hospital healthcare system is still a business after all and there are whole pieces of the organization that exist to make sure that we are running the business in a way that keeps it solvent and keeps us able to have the lights on and the doors open. And so there are definitely huge demands for different kinds of products coming from those folks and I think that we see it in a couple of different areas. So one is we see a combination of people on the provider and payer side who say, "How can we make sure that as we switch to this value-based care model, we really get a good comprehensive sense of who the patients are so that we don't engineer a plan that we think solves one problem and then we find out six months later it creates another one?" And so the only way to really try and make sure that you are doing something like that is going to be a comprehensive solution with a strong net positive is to know as much as you can about people, so there is a lot of demand for integration of patient-level data that includes claims data, clinical data, and other things so that the people tasked with providing that value-based care can feel like they've got a good comprehensive view of the patient. There is also a lot of strategic demand, so that the folks who do business strategy in the organization always want to know where is the next place we can go; where are the areas where we can strengthen our patient base; where do we have an un-served or under served sub-set of people and how can I make sure that as soon as we identify them, we are doing something that makes good sense to bring better care to those people. So I think it's a combination of operational and strategic and both of them I think are largely driven by evolving cost utility type models right, that say we've only got x number of dollars to do the right things for patients. You better make darn sure you are not wasting any of those dollars and that you are using every piece of information at your disposal to get it right.

David Harlow: Are there any particular clinical areas that you've had some interesting projects, or successes, or anything you could speak to specifically?

Nick Marko: Sure, so I think one of the areas that makes a good, sort of a good example of how these things can come together is a project that kind of overlapped between our radiology and our heart and vascular folks. So every time somebody gets a CT scan, the radiologist would frequently read or comment on the thing that was specifically indicated to be ordered for, and then will also comment on incidental findings that they see. Many of the ordering docs don't tend to focus on the incidental findings because they ordered a scan for a particular reason. They want to know does this guy have a big liver or not, does this guy have an inflamed gall bladder or not, right. So they look at that, they zero in on that, and then they go about their business. What we found was we would frequently find patients who had incidentally-identified abdominal aortic aneurysms and because it wasn't the reason people were ordering the scan, it generally was a high risk for not being followed up on. And so we integrated a couple of pieces together, we integrated text analytics that we applied on top of our radiology reports with some of those data management tools to bring pieces of the patients' clinical and counter history

together to see who and what they had followed up with and about. And then for patients who had an incidental diagnosis of abdominal aortic aneurysm, but who'd never had appropriate follow-up for that, it keyed right into our care managers who could reach out and contact the person and say, "Hey you know you had this incidental finding, and that we've got a whole pathway of taking care of folks who have these issues, and for monitoring and surveillance, would you like to be connected with that pathway?" And so we've been doing that for quite some time now and we have definitely identified people who have gone through that pathway, have needed that surgery; who've done very well as a result of that, we've got a variety of people who are on active surveillance now, whose disease is stable, but got to keep an eye to make sure it doesn't get worse. So I think that's a real example of informatics and data management getting right to the level of the patient, and it's a combination of text analytics and serving up information to care managers and combing through diagnoses in an automated fashion, but I think it's a good example of how you can put those pieces together to actually reach the level of one individual patient to do the right thing. And in fact it's worked so well that I've adopted that technology into my own area of clinical practice – I'm a neuro-surgical oncologist, I do primarily malignant tumors of the brain and spine, and we've started doing this for metastatic disease to the brain because particularly here where the system has been centralized, folks may get imaging, folks with cancer may get imaging and there may be suspicious findings for brain metastases on a head CT or PET scan and that's not what the oncologist was necessarily looking for, it gets easily overlooked. And so again, we've got a comprehensive brain tumor clinic, multi-disciplinary brain tumor program and anybody who's had one of these incidental diagnosis on their imaging, but who hasn't specifically had a surgical follow-up or comprehensive or oncology follow-up, we can reach out to them and see if they want to be engaged in that system. So I think we spot patients like this on a regular basis and I think hopefully, we are catching people who might have fallen through the cracks and helping them live better and longer lives.

David Harlow: That's terrific. Do you see any aspects of that of effort being influenced by the way Geisinger is organized versus the ability of other kinds of health systems to deliver that kind of care, that kind of information, getting people into those care pathways?

Nick Marko: Yeah, that's a good question. I think for us, one of the biggest things that we have come to realize is that rigid separation of your technology team from your information or informatics team and those from your clinical teams stifle creativity and also hamper implementation in these kind of areas. So it was not an uncommon thing a couple of years ago here to have a great set of technology developed and designed for doing something, but 80% of the clinical folks didn't even know that that thing existed, et cetera et cetera. So what we've been really trying to do and it's come actually from the IT side of things or the technology side of things, is to first of all bring technology and informatics and data together, sort of a comprehensive team so the – we are challenging the traditional idea of an IT department that's separate from the informatics group, that's separate from a data warehouse – and we are saying, really these should all be part of the same system. So from the leadership level to the operational level, we are trying to break down those verticals and have people from traditionally each of those teams working

together in unison on a variety of projects. And that's meant shifting reporting structures from the top down and I think it's made a lot of good sense. At the same time, there's got to be a lot of clinical integration with those backend technology teams, so we try really hard to get clinician informaticists involved in what's happening on the information and technology side of things. We try very hard to get folks who are in the IT space to go and be with clinicians who are delivering care on a daily basis. I shouldn't be sort of the rare exception where I spend some time in the hospital and other time with the tech folks and the tech stack, we should see a lot of this. And so we've been trying to catalyze a lot of that crossover. And so I think, when you ask is there a way that Geisinger has arranged that makes it easier to do here that hasn't been done or there might be challenges in other places, I don't think that there is necessarily a secret organizational structure that we know that nobody else does, but I think one of the things that we've come to realize is the natural tendency of IT and administrative type things is to silo themselves off from everybody else, and it actually takes a lot of active effort to push against that and to really try and bring end-users and clinical users into the fold of the tech space and vice versa so that people know what cool things we are doing and so that everybody who has an idea about how to tap into a new technology or new resource can say, "Hey I want that because that's great;" demand should always be driving what we do and there should always be more people who want to do cool stuff than we can possibly even meet the demand for. So that's really -- what we are trying to do is break down a bunch of those silos and bring IT and end-users together in a much more collaborative fashion.

David Harlow: Terrific, I'm sure that's an ongoing effort. So if you get to the point where you've solved the problems that were before you today, are there are some future items on your roadmap, anything that you'd like to call out in particular that you'd like to address if you have the time and the resources?

Nick Marko: I think that if I solved all the problems that are on our roadmap today, the only logical next step would be retirement -- but I think many of the challenges that we're facing today are largely operational kind of things. So there is an organizational component, there is an institutional change management component, there is a continuing-to-grow a data center culture challenge, and there is always keeping up with what's new and what's current in the space so that we can offer more and more services. I think at a certain point if you get beyond that, then you start to really get to the space where you are asking, "Why are we doing the things that we are doing with our tech stack?" and "How do we know that we are doing the best possible or taking the best possible approach to how we answer our questions?" So that can take a lot of different forms; that can be a question like "How do we know we are building our models correctly?" "How do we know we are using the right techniques to do that?" and "Have we really looked into how we fundamentally deal with the data we have?" It could be a question like, what's really the best way to take a large set of genome-level data and put that in the same space as some much smaller scale, but still probably disproportionately important clinical data, and balance those things against each other so that when we ask a question we get good and valuable input from both sides of things. And then maybe a third example would be something like there's lots of knowledge that gets captured in information and data in healthcare space that's never really written down, but is quite

important. So this kind of latent knowledge you can see, for instance, in the pattern and the timeline across which certain tests are ordered. How frequently somebody comes into the hospital and how many kind of interactions they have with in-between the doctors and in between that time? There's a lot of things that we do, you know, you say, well, I need this lab today, I don't need that lab tomorrow. Nobody really asks or there is no opportunity for somebody to write down why it is you thought that, but the knowledge still exists in the fact that the pattern is observable. So I think there are whole new sources that we haven't tapped into, that come, that are right under our noses, and there are whole new ways that we can be asking the questions and largely I think that's the real space of data science in healthcare. It's a lot less about just writing a predictive model because that's becoming more and more transactional. There's a lot more people that can do that. But it's asking how and why we do what we do with the information that we have and how do we know that we are doing it in the best way possible. And then related to that is, are we sure we are looking under every rock that we can possibly look under? So, one of my big interests right now is in external non-EHR, non-clinical data, because as you know, patients live 99% of their lives outside of the hospital. And so if I'm trying to make a clinical assessment or a prediction about how somebody is likely to do based on the 1% of their life that I can see, it's not surprising that sometimes I get that wrong. So I'm very interested in figuring out things that we can figure out about our patients' data, that we have about the patients, that is outside the four walls of the hospital because that's where people live most of their lives, and that's where most of the determinants of their health are at play. So I think if we look at how we ask our questions and if we look at making sure that we have thought to check in all of the places that where good and valuable data could live and we start to really bring those things together, I think that's when we're, I think that sort of the next generation of this, it becomes a lot less about how data moves and how we get it in the hands of this person and that person simultaneously, and becomes a lot more about how do you ask a really smart question and how do you make sure you are asking that question of the best and most comprehensive possible set of data that you can. So I think that's generation 2.0 of this whole type of endeavor.

David Harlow: Excellent. Well, thank you very much for taking the time, I really appreciate it. This is David Harlow from HealthBlawg --

Nick Marko: Absolutely my pleasure --

David Harlow: And I've been speaking with Dr. Nicholas Marko at Geisinger Health System who is working on solving today's problems as well as tomorrow's. Thank you very much.