Clinical Research Forum
Information Technology Roundtable
Meeting Summary
November 16-17, 2016
Chicago, Illinois

Contents

Introduction

The Information Technology (IT) Roundtable held its annual meeting on November 16 & 17 in Chicago, Illinois. This report summarizes the presentations and discussions held on November 17, 2016.

Session 1: Setting the Stage

Measuring Efficiency and Cost Benefit
Daniel Ford, MD, MPH, Johns Hopkins University

Giving an overview of the current clinical research environment, Dan Ford, M.D., highlighted the decisive advances made in IT implementation in recent years, pointing to EHR adoption as nearly complete and CTMS applications as common, even if significant variations among these latter persist. Productivity has also gone up, though with extraordinary new pressures being placed on research coordinators, often leading to metric-dependent processes and the potential for being dropped should those metrics not be met. On the whole, then, the amount of complexity has increased tremendously, introducing a situation where more incremental and fine-grained assessments now motivate changes and decisions.

Despite the considerable degree of murkiness that remains, along with uncertainty and even confusion—and while keeping in mind that efficiency is more than simply an IT issue—it is nevertheless worth noting certain advantages to this added complexity. Probably foremost among these is the reflection and thought now being directed to current processes initiated by this focus on efficiency. For many formerly implicit processes, we're now lifting up the hood and asking whether this is really the best way to carry this out, or whether there are costs associated with doing a process in this way that we can no longer take on. In many cases, the most basic questions are being asked for the first time: How are we doing this? What's the current state of this process? Focus has thereby been placed on goal-based considerations on the operations side, leading us to think carefully about whether the service provided meets not only our own objectives and expectations, but more and more thought is now also being given the customer.

Indeed, an important place to look for ways to improve productivity and efficiency is toward our stakeholders. Of course, this introduces a great deal of complexity in its own right, since there are many stakeholders and their needs might not always converge. All the same, by working alongside our stakeholders when considering the implementation of new IT platforms, we should be able to increase the specificity of the problem we're trying to solve, while at the time getting a clearer picture about what the current processes are before proposing changes. All of this will deepen our understanding of the benefits we can expect, and, perhaps even more importantly, establish firmer criteria by which to judge whether the problem was solved.

And yet substantial barriers remain to deep implementation of efficiency metrics, which in itself might hardly be a bad thing. In fact, from some perspectives it is our good fortune that our processes can't be reduced or quantified in terms of efficiency metrics established for iPhones and GM mufflers. The very difficulty of defining a "unit of research" as an input, or of understanding FTE's relation to output in a research scenario, means we can't be graded according to something else. Again, this should renew our focus on what we're actually doing and incentivize us to create our own measurements for research productivity before others do it for us.

One particularly effective way of doing this mentioned by Dr. Ford stems from multi-center trials, where comparative analyses can deepen insight not only into cost and speed of patient recruitment, but into elements spanning the entire spectrum of clinical trial processes and infrastructure. An additional area in which advances have been made and in which further improvement is still possible, in terms of patient safety and cost efficiency, involves further automation of aspects of billing, administration, and similarly centralized tasks, which can relieve lower-level employees of judgment-level interaction with the system.

The brief discussion that followed on Dr. Ford's remarks largely revolved around frustrations resulting from various research IT implementations. These included declining productivity stemming from CTMS-related double entries, increased complexity of the research participant experience, as well as increased expense stemming from insurance clearance. Taken all together, these frustrations pointed to added costs that have not yet been balanced out by promised increases in efficiency. In fact, adding some weariness to this picture, Dr. Ford referred earlier in his remarks to clinical IT's adoption of large EHRs as a good indication of where research IT might expect to be in ten years. Other than gains in patient safety through better management of medications, his sense was that it was still very hard to show any gains in efficiency or productivity based on clinical IT's implementation of large EHRs.

And so the opening remarks to this year's IT Roundtable painted a decidedly mixed picture of the potential for gains in efficiency and productivity, surveying plenty of possible pathways to improvement while pointing to only occasional evidence of real success.

Update on the Maturity Model
Peter Embi, MD, MS, The Ohio State University

Building on territory covered in Dr. Ford's opening remarks, Peter Embi, M.D., M.S., renewed a discussion covered in past IT Roundtables by introducing the topic of the maturity model. Asking both what a maturity model is and what motivates its adoption, Dr. Embi elaborated the grounds for a maturity model's value, while also focusing on why research IT has entered a phase in its implementation of IT-related processes that makes a serious discussion of a maturity model essential.

Even more, Dr. Embi stressed that, based on insight gleaned from previous years' discussions, it was now becoming clear that not beginning to specify and elaborate a vision of a mature IT infrastructure for an academic health center would likely prove harmful in the near future. The harm done could reveal itself in the areas of access to funding, support, and resources, as well as in the subtler and more pervasive confusion of having missed the opportunity to establish a deep comprehension and plan for how IT fits into the overall enterprise.

Highlighting instances of maturing activities by referring to the capabilities and functionalities that have improved over the past decade, as well as the standards that are beginning to emerge, Dr. Embi emphasized the now necessary role of governance in bringing together this emerging suite of best practices. Once carried through, such objective assessments of the prevailing technologies and benchmarks would benefit organizations and grant them the confidence of knowing that their approach and priorities are correct, as well as leading ideally to improvements in research efficiency and productivity. That, after all, Dr. Embi said, is the whole point.

Moving on to some nomenclatural distinctions that arise in the area of maturity models, Dr. Embi provided definitions of maturity and maturity index, while also expanding on the difference between a deployment (or adoption) model and one focused on maturity. Maturity refers to the degree of formality and optimization of an organization's culture and processes, while a maturity index measures an organization's capacity to deliver a given service in terms of factors like culture and process. A deployment-focused model, on the other hand, which complements a maturity model, measures the degree to which an organization has deployed or adopted a technology related to a delivery service.

Maturity models can be broken down according to these levels:

- Level 1: Initial……Processes poorly controlled and reactive;
- Level 2: Managed……..Processes characterized but still reactive;
- Level 3: Defined……Processes characterized and proactive;
- Level 4: Quantitatively managed……Processes measured and controlled;
- Level 5: Optimized……..Focus on process improvement.

The characterization of deployment models has slight differences:

- Level 1: Absent/ad hoc……..No IT solution available or addressed in improvised way;
- Level 2: Repeatable……Established capability, but mostly informal;
- Level 3: Defined……Standardized capability with documented procedures or capabilities;

- Level 4: Managed……Managed capability with measured results using measured performance indicators;
- Level 5: Optimized……Performance measured, regular reassessments to improve practices, and manage risk.

After giving examples of several other such models, Dr. Embi elaborated on the maturity and deployment models he and his team have developed in recent years with the help of input taken from surveys conducted at past IT Roundtables. Focusing on four main areas—namely, IT use in research compliance, electronic data capture, research data repositories, and infrastructure—Dr. Embi was able to show increased adoption in all areas based on surveys conducted in 2005, 2007, and 2012. With this evidence of maturation, Dr. Embi further grounded his earlier claim that the time had become ripe for a more formal investigation into governance and best practices in the area of research IT in academic health centers.

Discussion

A question raised in the discussion following Dr. Embi's discussion asked why there was the need for the two axes—maturity and deployment—when it was clear that technological deployment is always already reliant on maturity and embedded in any number of pre-existing processes. In this light, it would seem natural to merge them. It was responded that, in fact, models do exist that have merged them, EDUCAUSE being an example, and yet the situation there involves the existence of both a maturity and deployment index, both of which are then deployed into a core data service. This was then characterized as a maturing of their assessment processes.

In fact, in many instances, it is precisely a matter of multiplying the number of deployment indices, each being dependent on different drivers. It was added that, in general, it is easier to measure a deployment index than a maturity index, since the former often entails concrete questions about particular tools and numbers of users, while the latter strives typically to quantify the substantially vaguer idea of a mature governance structure.

With this potential for multiplying and specifying the concept of deployment, which allows the researcher to introduce different drivers as indices sought for measurement, one gains the capacity to see a broader picture of an organization's strengths and weaknesses. Furthermore, this grants the organization the ability to make internal comparisons between units within the organization, as well as the opportunity to compare themselves in both broad and specific details to peer organizations and competitors.

A question was next raised regarding whether a crucial dimension had been understated in the development of the two axes, which was the importance of people. The commenter mentioned that he could conceive a situation in which exceptional scores in both maturity and deployment were achieved, while overall performance still lagged on account of the lack of right people. Dr. Embi took this opportunity to suggest that one's human talent might also be subject to measurement and therefore indexed and compared to other indices, which was referred to as component-level scoring, and in this way a broader picture of the organization's state of convergence might be achieved.

Another point that was made related to the cost associated with each level in a maturity index. For instance, if it could be shown that all organizations with a level 4 deployment of a particular technology had spent a certain amount of money, other organizations might realize that the achievement of level 4 was simply out of reach, barring significant internal reorganization in other areas. Not just disenchanting, this sort of measurement could be helpful for rationally orienting an organization's resources.

Session 2: Maturity Model

Peter Embi, MD, MS, The Ohio State University
Boyd Knosp, MS, University of Iowa

Deepening the discussion on maturity models begun in the previous session, Boyd Knosp, MS, began Session 2 by speaking about some of the specific challenges to building an effective research IT maturity model. He specified three that he sees as particularly important to understand, both at a general level and in terms of the costs and investments these challenges impose: one, the role research IT has in leading to innovation and sources of discovery; two, the role of compliance and security; and, three, the organizational structure, which would include determining whether an informatics department is needed, or a biomedical informatics department, or an institute.

Following on these three challenges, Mr. Knosp provided three reasons for why maturity models are important: one, they impose goals, targets, and standards for the IT framework that's being built; two, they provide an organization with insight into itself, both in terms of how it compares to others and how people within the organization relate to it; and, three, they provide a powerful tool to engage and communicate with leadership.

Mr. Knosp next elaborated on the development of the particular maturity model he would be asking participants in this year's IT Roundtable to test in a pilot version during Session 2, mentioning in particular the ways in which focus groups, conversations, and several pilot trials had allowed him to appreciate how this maturity index might be used to establish standards of practice while also serving as a point of discussion within organizations.

Moving on to an explanation of their development index, which IT Roundtable participants would also test in a pilot version during Session 2, Mr. Knosp noted that the five levels making up the index were comprised of both a list of the technologies deployed and a self-assessment of their deployment, adding that the index was in this sense slightly more mature than the standard sort of deployment index that didn't extend beyond basic questions about implementation, since in this case a certain amount of qualitative self-assessment was involved as well.

Focusing next on what exactly their maturity index measures, Bill Barnett, Ph.D., elaborated on how their conception of maturity was dependent upon

- policies;
- governance;
- the extent to which IT is a priority;
- the sort of dedicated infrastructure in place for it;
- what kind of leadership, as well as leadership commitments, are there;
- how are we, as an organization, invested in supporting our IT culture;

- integration of clinical IT with research and teaching mission.

Maturity Index Pilot Survey Discussion

After the IT Roundtable participants had taken the maturity index pilot surveys and voiced various concerns about the high level of maturity implicit in various questions, as well as the unlikelihood using the index to create an organization-wide level of consensus, Mr. Barnett suggested that the survey had been designed with at least some of those confusions in mind. Insofar as its intention was to initiate internal discussions based on divergent responses to questions, his feeling was that increasingly valuable areas of discussion might be opened up in that way, areas focusing on institutional research priorities and how IT might play a role in them and what sorts of investments might be made given already-existing institutional strategies. The goal, he said, was to force such honest conversations, and even if this results in not getting the kind of answers you'd like to hear, at least you'll be getting clearer messages from the research team and CFO.

Dr. Embi added that, as a Chief Research Information Officer in his present role, he knows that a few years ago his organization lacked the level of maturity to begin filling this sort of survey out, but now with the correct team in place his organization possesses the sort of organizational vision necessary to answer these sorts of questions.

Mr. Barnett acknowledged the tremendous heterogeneity of the research IT situation and the challenges of ever implementing a standardized suite of maturity assessments in this context. Again, the survey might be better seen as a tool to explore the organization's diversity and reflect on its core drivers, potential examples of which being:

- proving to NIH that we're mature enough to be given a grant;
- focusing on regulatory requirements;
- making a particular statement to OCR or CMS.

If organization-wide unity remains out of reach, the goal of at least increasing levels of alignment would surely be helpful.

Further comments focused on what was called the subjective or highly-opinionated aspect to the survey, making it difficult, for instance, for a CRIO not to "strongly agree" with a particular survey question lest s/he be seen as not doing a good job. In response, Mr. Knosp stated that a future goal was to make the survey more objective, which could involve introducing external third-party validators, or more finely-tuned objective units of measurement, all of which would be important for a fundamental driver of the project in general, i.e., capturing a picture of one's position in the competitive landscape.

Deployment Index Pilot Survey Discussion

The IT Roundtable participants next took the technology-focused deployment index pilot survey, whose object was to gauge the level of deployment of IT systems within research institutions.

Opening comments revolved around the degree to which the question of penetrance was insufficiently broached by the survey. This could be a matter, for instance, of having a great data capture tool whose use is not required across the organization, leading some groups to operate at highly managed levels, while others persist at low ad hoc levels. On the other hand, the use of certain poorly-funded, not-very-good tools might be mandated across an organization and managed at a high level, but their low quality would beg the question of why they're being so highly managed in the first place.

It was suggested that one might have two separate indices, one gauging types of tools deployed and another oriented to levels of penetrance reached across the organization. In this way, an over-all maturity index would register the extent to which these levels converged across the various deployment indices.

Mr. Knosp raised the issue of actionablity stemming from the self-assessment received by an organization after taking a survey and acquiring this picture of its maturity and deployment. How do you get to the next level? he asked, acknowledging that from within level 2 it is often impossible even to imagine what level 3 looks like. On the one hand, organizations—or units within an organization—lagging at level 1 or 2 might look to organizations scoring in the higher levels of maturity and pattern their development accordingly. On the other hand, if there is a broad level 1 or 2 stasis across an organization or even landscape, this might indicate the need for the introduction of a particular program or new team to serve as a sort of bootstrapping mechanism.

This fed into a consideration of the sort of thinking required to develop maturity indices in the first place and the sorts of questions that might be asked that would usefully target particular areas of potential maturity. Taking the areas of subject recruitment, registries, and secondary use of clinical data as examples, Mr. Knosp reflected on the challenge of developing precise questions that could singularly illuminate levels of maturity within these separate areas, leading to actionable advice and even a clear picture of the cost of getting to level 5. This sort of inquiry he referred to as the maturing of the maturity index itself.

A commenter raised the basic issue of determining the key structural characteristics required to take the first steps towards truly stratifying the situation of research IT among academic health centers, mentioning these sorts of questions as merely broad brushstrokes contributing to the sort of picture that might ultimately be desirable:

- how many faculty members do you have?

- what's the degree of medical school and health system integration?
- how many CIOs do you have?
- what percentage of your research budget is from NIH?

Dr. Embi admitted that the high-level paper on the research IT maturity index he was presently involved in writing remained at its high level precisely because these sorts of data were still unavailable, making the formalization of the context in which they might be acquired the essential next step to take.

Session 3: Data Warehousing or Population Management Tools? The Next Generation

Update on Research Data Repository Tools
Shawn Murphy, MD, PhD, Massachusetts General Hospital

Shawn Murphy, MD, PhD, introduced the discussion of research databases by stressing that finding patients for clinical trials is the most important thing, noting experiments don't work unless you have enough patients. Thus, tools that go to improving recruitment strategies are crucial to efficiently using—or even reducing—the billions of dollars that are used in recruitment each year.

Dr. Murphy then went on to describe the high throughput methods they have at Partners for supporting translational research, which operate according to a program that involves finding sets of patients from medical record data. Investigators explore phenotypes of these patients using tools with a translational team. Often it goes into distributive networks across institutional boundaries, selecting phenotypes for public health and hypothesis testing; and precision medicine is delivered into clinical care as a result of many of these investigations.

The process so far has involved using a unique database system where compliance with HIPAA regulations proceeds through two steps. Dr. Murphy said that it was possible to query 6.7 million Partners patients and 2.5 billion diagnoses, using EHR data, genomic data, nHealth data, and imaging data. Once the patients are found, you advance to the second step, where you get the identified data that you then feed into your SAS in order to perform your investigations. Security is preserved by adding patients to the database through Partners's HIPAA notification, thereby anonymizing the data as a query tool, which allows Partners to fulfill the goal of granting pervasive access to the data while not violating HIPAA.

Dr. Murphy added that these tools are unique for research because they're sufficiently straightforward and intuitive for all clinicians to use.

Next, Dr. Murphy demonstrated the ability special tools have to obfuscate data, thereby thwarting attempts at re-identification. By adding noise, queries come back with slightly different results each time. Only if done repeatedly does a query start to converge, at which point the user gets locked out.

Through supporting something like a biobank, where genomic data get integrated, it's possible to create tools that allow people to investigate phenotypes very specifically. The point is you can use these same tools across various databases—all open source and with many different implementations—and then you can put all these things together in a big data commons. Not just the EHR data, but much of the biobank and genomic data gets included as well, all of which then gets overlaid with informatics tools, such as natural language processing, which go to create things like validated phenotypes.

In the end, you gain the capacity to make very accurate diagnoses—much more so than can be found through the EHR data. The point here is that what we want for research is often something that's quite unique. By putting unique databases together and performing these tool overlays, it is possible to deliver very complex queries for research, while respecting the privacy of the patient and complying with regulations.

The Research Case for Population Health Management Systems
Peter Winkelstein, MD, MS, University at Buffalo

Peter Winkelstein, MD, MS, began his remarks by reminding those present that last year they had started thinking about the world outside of the research endeavor, particularly in terms of the changes coming to clinical IT with the introduction of population health management tools. He specified that these were the sorts of tools now being rolled out in the clinical environment to do care management at the population level—sometimes within a hospital system, sometimes within a community—with goals that tend to focus on a host of cost management measures now receiving attention. Cerner's Healthe Intent and Epic's Healthy Planet were cited as examples of these continuously-improving vendor-produced tools.

Dr. Winkelstein next specified that his intention was to pose the slightly controversial question regarding how population health management tools might fit into research activities. Where do they fit? Do they fit? Is there something about them suited for deployment on the research side?

Fundamentally, he said, a research data repository and a population health management system use one data source while trying to do two different things. EHRs are being used in the clinical environment to do clinical things, while researchers are trying to take that data and do research. And there are lots of problems with that, he said, as everyone knows, not the least of which is that the population health systems are designed entirely around patient care, meaning research comes along as a second thought. Health systems buy EHRs to do documentation and billing, along with some clinical decision support, and so hospitals aren't buying EHRs for research-related reasons.

Next, Dr. Winkelstein outlined the set of functional requirements that we look for from our research data repositories and similar tools:

- we look for aggregated data;
- we look for de-duplicated data;
- we want the data to be normalized and put into some common data model;
- maybe we ask for some cleaning, some filling in of the gaps;
- there can also be de-identification involved, depending on what the IRB protocol is and research needs are.

Dr. Winkelstein's controversial suggestion, which was made, he stressed, for the purpose of debate and exploration—and not for its truth-value in the present—was that, from the perspective of this set of functional requirements, there is no difference between what we're asking of our research data repositories and what we're asking of our population health management tools. The latter are able to do everything just outlined in order to do the care management they're designed to do. He proposed, therefore, that population health management tools are effectively research data repositories, i.e., they subsume the functionality of the research data repository.

The first and most obvious feature of population health management tools that make them potentially superior to research data repositories is that the research institution doesn't have to pay for them. Not only does the clinical side have a lot more money to begin with, but they also have the obligation to make the investment, given their ongoing effort to reduce unnecessary hospitalization use in the Medicaid population. Thus, if the institution has it already, is there a way in which research can use it? To what extent can I use it as my research data repository? And can I do that less expensively than supporting my own research data repository?

Dr. Winkelstein presented some advantages to approaching the problem in this manner:

- population health management tools are vendor supported, so if a problem emerges, it's only a phone call that's required;
- their scope can be enormous, i.e., the entire Medicaid population in an area, which is often larger than that covered by internally-developed research data repositories;
- they come pre-assembled, whereas researchers typically rely on in-house developers, who build the architecture, etc.;
- these tools are very tightly coupled to the clinician's experience, so they're designed to connect back to the EHR, both to send signals back and to work on the clinical side, which addresses one of the key problems we have, namely, that the quality of EHR data tends to be low from a research standpoint, precisely because physicians lack an incentive to input high quality data for research. However, population health management tools can give them feedback which can improve the data and its potential for use in research.

So the question concerns whether this is where research data repositories belong. And yet here are some reasons for why we don't need to rid ourselves of them just yet:

- the tools that come with a population health management system are designed for care managers, so they don't come with the querying tools Dr.

Murphy had just spoken about. And yet: it's not at all clear that it would be hard to put them on there. Dr. Winkelstein's vendor, for instance, told him that they were interested in linking to I2B2, thereby making I2B2 a sort of front end to the population management tool;

- from a research standpoint, the data model used by population health systems remains ill-described. They throw terms around like ontology, but when such paths are pursued things tend to get fuzzy fast. What is their data model? And how actually do they do their normalization? Fidelity might be lost as they take the data from its native format and stuff it into the common data model, which is to say that none of it has been laid out clearly yet—though, again, that doesn't mean it couldn't be made clear in the future;
- needs of HIPAA, IRB, and de-identification are very tricky in the research world. The regulatory environment of research and population health are therefore deeply different, and as researchers we're required to get all the paperwork in place, while population management systems don't consider any of that;
- in terms of genomic data, the robustness of population health management tools hasn't yet been demonstrated, largely because populations aren't managed in terms of their genomic profile yet. From a research standpoint, however, we're linking our data more and more—it's essential to what we do—while population management tools are still more community-based and not really designed for this national level of linkage—yet.

In these respects, then, population health management systems may still lack the maturity to replace research repositories, even if the financial pressure to use these tools—instead of supporting an entirely separate architecture—means there's a significant incentive to look further into this.

Research Data Repository and Maturity

Dr. Murphy followed Dr. Winkelstein's discussion with an assessment of an institution's access to population health management systems given the maturity of their research repositories:

- at maturity level 1, population management people don't have sufficient trust in your enterprise to come on board. Thus, at this level, it is often necessary to create and/or expand your repository to get something done;
- at level 2, you'll find they have more trust in what you're doing and you'll get more buy-in, but it's at this point that HIPAA regulations start to enter the picture, introducing a difficult situation that's tricky to get around;
- at level 3, as access increases, you discover that the tools used in population management are really complicated and that most clinicians will not be able to use them;
- at level 4, you'll have plenty of buy-in, but interoperability of other big data, like genomics and imaging, isn't so much not thought about, as actively dismissed at the level of population management;

- at level 5, people in your research enterprises start realizing that the phenotypes that are defined by the ICI codes are terrible. Typically, we find that it varies from one diagnosis to another, and this is because the goals on the research side, which rely on the accuracy of diagnoses, differ from those on the clinical side, which often reduce to issues of billing. Another way of characterizing the difference is: encounter-based versus patient-based.

Dr. Murphy added that everything at present is going towards the science of determining the phenotype from EHR data, which is not easy, but research portals present a real opportunity to optimize phenotypes.

Dr. Winkelstein responded by pointing out that extracting phenotypes from this underlying data is already being done by research, so, from an architecture standpoint, why not let the population health management systems do the data aggregation, and then research can run phenotype algorithms. Why does it mean that we have to have separate systems?

Discussion

The first comment agreed with Dr. Winkelstein that his thought experiment was worth pursuing, but he also saw several additional roadblocks to employing a population management tool in lieu of a research data repository at this point. First, he saw the attribution problem within EHR data as posing a difficulty; it seemed unlikely that researchers would be able to use it without feeling forced to question the results. Furthermore, given the desire for knowledge of data provenance within research, he was doubtful that vendors would ever grant sufficient access to their databases, meaning the data provenance would never be disclosed in enough detail to satisfy researchers.

Another comment stated that even if population health management systems aren't ready to replace research data repositories in full, it could still be worthwhile to start looking into the ways in which they might already be useful. In this way, rather than continuing to resort to a completely duplicative infrastructure for research data repository, we could look at the population systems as something we might build on top of and augment from the research side.

A question was raised regarding making clinical notes more meaningful and reflective of a reality that would increase their usefulness for research purposes, diagnostically and otherwise. Dr. Winkelstein responded that the nature of our clinical processes and the systems we have put in place have enabled the increasingly low quality of EHR data, but he added that that problem is independent of moving from a research data repository to a population health management system. Calling it a "source data problem," Dr. Winkelstein pointed out that the source EHR data would be the same in either case. I2B2 was then mentioned as a tool designed specifically to improve EHR data.

A comment was made that suggested the real goal was a merged repository supporting both research and care, which would get lots of feeds, including the population health feed. Columbia and Northwestern were cited as examples of this model.

A basic distinction between research and population health repositories was then offered, basing itself on the idea that the governing driver in population health is revenue, which can't but lead to poor research. Taking the different definitions of asthma offered by population health and research, questions were raised about the extent to which researchers could be satisfied with the 'reasonable definition of asthma' offered by population health systems, as opposed to a more open and contested definition found on the research side. It was added that those 'reasonable definitions' were already in a lot of the tools being used.

Following on this discussion, a question was asked regarding the maturity of the learning health system and the relevance of much of the clinical data being collected now for research. Dr. Murphy responded by observing that one way to get at the truth is to take different dimensions of the data, which he referred to as orthogonalizing the data. A problem with EHR data, he said, is that it's eschewed for billing, and there's no way to get that bias out, while physician notes are eschewed for communication purposes, which makes for a great deal of irrelevance, the point being that all data comes with bias—even genomic data, whose penetrance can be very poor. And so the point is that by bringing all these different sorts of data streams together you can cancel the biases out and arrive at something closer to the truth.

Session 4: Coop-etition in Chicago

Justin Starren, MD, PhD, Northwestern University

With Chicago's large number of medical schools, universities, and corporations, all competing viciously and collaborating incredibly smoothly, Justin Starren, MD, PhD, introduced the five members of what he called the Host City Panel, brought together under the theme of *coop-etition*. Dr. Starren also presented a brief overview of the range of integration and collaboration in Chicago among the various bioinformatics communities.

Annette Valenta, DrPH, University of Illinois at Chicago

Dr. Valenta, Academic Director of the Patient Safety Leadership program and founding member of the Institute for Patient Safety Excellence, was tasked with describing and outlining the challenges faced in establishing the Chicago Initiative for Biomedical Informatics and Data Science Training, a "competitive collaboration" established between three Chicago universities—University of Illinois, Northwestern, and University of Chicago—all of which have CTSA sites and all of which train students. While the three compete for faculty and students, their three major areas of collaboration are the Chicago Biomedical Consortium, the Chicago Area Patient Outcomes Research Network, and Regional Chicago CTSA.

Having recruited over twenty new biomedical informatics and data science faculty, while establishing PhD programs at UIC and Northwestern, as well as Masters programs at all three universities, the primary engine behind the *coop-etition* is the Chicago Metropolitan Exchange program, which allows students to take courses at any of the three universities. The core of the curriculum is a two semester course designed and taught by thirteen faculty from the three universities, which is taken by all the PhD and Masters students from among the three universities. Covering an enormous range of material—ranging over knowledge acquisition, ontology representation, database modeling, and qualitative and quantitative methods—the students receive a full overview of the topics everyone needs to know, regardless of where they decide ultimately to do their research or work, whether it's in academia or industry.

Philip Greenland, MD, Northwestern University

Dr. Greenland, Director of the Institute for Public Health and Medicine and Harry W. Dingman Professor of Cardiology at Northwestern, spoke about the Precision Medicine Initiative, which is now called All of Us. This joint initiative between NIH and the White House claims to be part of the "emerging approach" that represents a merger of lifestyle measurement and biologic characterization, instigating the "radical shift" from population approaches to a much more individualized one. So far, Dr. Greenland said, there's been a limited amount of movement forward since July 6, 2016, when the kick-off meeting for the initiative was held, as well as a

missing appreciation for the immense difficulty of organizing a large bunch of people who have never worked together, who don't have a protocol, and who are unclear on what they're supposed to do.

The good news, from the Chicago point of view—largely because the CTSAs work together and get along—is that the three universities—UIC, UofC, and NU—came together to collaborate in what's now referred to as the Chicago Precision Medicine Consortium. One of the interesting components of the PMI is that it's not just healthcare organizations that are expected to recruit the million people who are to take part, but there's an expectation that around 250,000 or so will be recruited by direct mail, advertising, etc.

Also worth mentioning, Dr. Greenland said, is that one of the core principles of the study is that they want it to be "broadly representative" of America in terms of people, geography, health status, etc., which is part of what they're calling the "transformational approach to participation." The beginning of the program will have some data collection, but what that's going to look like is still unclear, though what is clear is that the whole program will be completely open, so that whatever data is collected will be open to investigators anywhere in the world.

Abel Kho, MD, Northwestern University

Dr. Kho, Director of the Institute for Public Health and Medicine and Associate Professor of Medicine and Preventive Medicine at Northwestern, spoke about data networks in Chicago and the attempts being made, through PMI and other initiatives, to reach out into the larger community, by which he means not only into other populations, but into new types of data as well.

A key question in reaching new populations and going beyond the academic medical center regards doing recruitment for clinical trials in a way that engages the community. Dr. Kho mentioned recruitment currently taking place in Chicago through churches, which involves training pastors on the South and West Sides of Chicago to conduct research and recruit for trials.

Dr. Kho also spoke about expanding beyond Chicago into other states and healthcare systems in the tri-state region. This can involve reaching out into independent primary care practices, which typically go unseen by most studies. Data from over 200 practices is being entered into a data warehouse, from where it gets piped down into a population health platform, with the goal of sampling quality measure across the tri-state region.

Dr. Kho referred to All of Us as a broad picture of the informatics aims for precision medicine, noting that it is a complex, multi-group effort whose goal at the end of the day involves reaching participants in different places. What's exciting is the opportunity to build some central infrastructure, capable of being widely re-used for

other purposes in the future and involving a large central repository with lots of feeders coming in from multiple different sources.

Sam Volchenboum, MD, PhD, MS, University of Chicago

Dr. Volchenboum began by mentioning another regional cooperation in Chicago, which is the Institute for Translational Medicine at the University of Chicago, which is another way multiple institutions in the city have been brought together to collaborate, further pointing to the fact that there's a lot of multi-layer collaboration going on.

Dr. Volchenboum's presentation extended a talk he had delivered the previous night on the requirements of building a data commons, the basic requirements being having a computing cluster that is publicly available, along with data, software, and tools, and to have the data conform to a fair model of data, which means it's viable, accessible, interoperable, and usable. Having oriented his discussion the night before around the issue of data standards, his presentation at the IT Roundtable was a demonstration of what his team built for the neuroblastoma data commons. Dr. Volchenboum said that he sees the result as an opportunity to share data in a way that's never existed before, since it puts into place all the necessary regulation, privacy, and IP issues.

Steven Collens, Matter

Mr. Collens, the CEO of Matter, a healthcare technology incubator and start-up center that started 18 months ago, spoke about how he organizes his thoughts in terms of what's driving innovation in healthcare. Calling it the most interesting, dynamic, and exciting time *ever* for entrepreneurship and technology innovation in healthcare, he went on to mention the three trends he sees as driving innovation in particular—namely, the internet of things, the ubiquity of electronic health records, and the changing regulatory environment.

Asking next where this innovation might come from, he took his initial steps into an answer by saying it wouldn't be large companies, hobbled as they are in their efforts to innovate by a brand to protect and their deeply routinized and longstanding habits. Entrepreneurs, on the other hand, are great at innovation, Mr. Collens said, being good with capital and close to their end customers. Most important, they can test their products before they're fully baked and thereby get crucial feedback, which is particularly hard for a mature company to do.

And yet there are particular challenges that entrepreneur's face in the field of healthcare, not the least of which being the non-intuitive and often very complicated regulatory environment. Unlike other industries, Mr. Collens affirmed, there is a need to be deeply embedded within the healthcare system to grasp the value and importance of your innovation.

Discussing Chicago in particular, Mr. Collens mentioned that there are more Healthcare Industry leaders, more medical associations, more doctors per capita, and a huge amount of healthcare expertise in the city. What Chicago still lacks, however, is what Cambridge, the Bay Area, and San Diego already have, namely, a community in which the established players are connected to entrepreneurs and already engaged in collaboration. The goal of Matter, then, is to reduce the friction separating potential healthcare innovators from already established players in Chicago, thereby bringing together entrepreneurs and industry.

Discussion

The first question was directed to Drs. Kho and Greenland and concerned whether they might speak to the amount of momentum and/or inertia currently driving the All of Us initiative. Dr. Kho admitted that there was more inertia at present than momentum, even if the challenge holding everything up had become clear—namely, how to manage consent. Otherwise, many items seem to be in place and there's a great amount of potential, but it's still too complex to see how everything will fall. Dr. Greenland added that the large size of the study has created difficulties in instantiating deep levels of trust among the participants, a factor compounded by the unrealistic expectations introduced, which had hoped to have the project up and running by the end of Obama's term.

A follow-up question asked whether this confusion at the national level had allowed Chicago at a local level to become better organized, which revealed that indeed there had been a certain amount of unification that had taken place, a crowning example of which being the founding of the Chicago Biomedical Consortium.

Another question directed to Dr. Kho concerned how centralized the All of Us data architecture was shaping up to be, to which Dr. Kho said that it seemed potentially too centralized at the moment, leading to uncertainties over what steps to take next.

In terms of centralization, Mr. Collens was asked what he thought works best as a general rule. He agreed with the consensus emerging in the room, which was finding that too much centralization makes for an unhealthy climate for innovation. He suggested that a decentralized though connected network might be optimal, similar to what one can find in Cambridge, and that what Chicago still needed was a platform or outlet to force interaction between unconnected nodes of the network, which is what Matter sees as its primary purpose—namely, to bring together the right people in the right place who haven't yet managed to connect with one another.

Session 5: Vendor Panel

Introducing the final panel of the 2016 IT Roundtable, Dr. Ford mentioned that one of the traditionally unique aspects of the gathering was the inclusion of a broad range of stakeholders, from pharmaceutical company representatives in past years to this year's vendor panel, all of which he called an indication of the understanding we all have that we'll have to work together to make progress.

John McIlwain, MBA, CEO, Velos Inc.

Having been a longstanding member of the Forum, Mr. McIlwain began his presentation by expressing excitement over the progress made in the research IT community over the past four or five years in terms of systems deployed and his more general sense that there was a great deal of upbeat momentum driving progress forward. Breaking his presentation into two parts, he said he'd have some things to say on a couple tactical things he's seen work and not work, and then he'd talk about some of the opportunities on the horizon, both the easier things to do and some of the more difficult deployments that might yet be worthwhile.

After briefly outlining Velos' footprint and referencing a new specialty EHR-like product designed for the field of cell therapy, Mr. McIlwain noted several drivers that he sees as shaping research IT in the upcoming years. His overall vision focused on increased efficiency through deeper and more integrated connections made within the supply chain, leading to cost reduction and improvement in data standards and renewed focus on one's core competencies. Explaining what he called a typical Velos deployment, Mr. McIlwain spoke to the opportunity of integrating the increasingly complex and differentiated environment of systems into a single platform and interface, while also pointing to the advantage of mandating certain billing, administrative, and other centralized functions with an eye towards ensuring greater safety and efficiency.

Michael Rauwerdink, Director of Product Development, Forte Research Systems

Titling his presentation "The Future of CTMS in an EHR World," Mr. Rauwerdink began by contrasting CTMS systems to EHRs, saying that, despite their different domains and the continued importance of focusing on each system individually, a key for both systems moving forward involved finding areas in which to integrate them. He specified patient safety and increased cost efficiency as two areas in particular that would stand to benefit from deeper integration.

At Forte, a key to such integration involves a standards-based approach, which includes but also goes beyond data standards. Mr. Rauwerdink said that being able to scale and work across organizations requires commonality and best practices on both sides of the table, as well as strong buy-in from both the organization and vendors involved. In the end, all of this leads to a capacity for repeatable integrations, resulting in significant value and economies of scale stemming from

the potential for increasingly far-reaching collaboration. After a summary of several case studies exemplifying successful integrations between CTMS and EHR systems, Mr. Rauwerdink concluded by stressing that a key component to everything he'd covered involved dialogue between the vendor community and the researchers present at the IT Roundtable.

Nancy Smider, PhD, Research Informatics Implementation Lead, Epic Systems Corporation

Giving a high-level landscape of some of the priorities that she sees for implementing EHR systems in research, Ms. Smider began her presentation with a discussion of what she called "participatory science," one aspect of which entails not just engaging patients as participants, but as partners in research as well. Referring to what she called the "recruitment toolbox," she outlined a variety of approaches to recruitment, overviewing everything from recruitment that is sensitive to person-preferences to types of targeted outreach and targeted point-of-care approaches. All modes of recruitment need to be considered if faster and more cost efficient recruitment is to take place.

Pushing the patient-centered idea slightly further, Ms. Smider mentioned the roles patients might play in study design, highlighting patient advocacy groups who take part in reviews and designs of studies as one example already under way. Following this with a discussion of the democratization of research as an extension of the idea of the patient as partner, Ms. Smider mentioned All of Us as well as self-initiated data donation as novel approaches to engagement with the general population in research, re-evoking thereby a central theme to the 2016 IT Roundtable by referencing the blurring of the boundaries between care and research activities.

Highlighting the issue of patient-generated data, whether the data comes from questionnaires or consumer wearables, Ms. Smider suggested that a key driver for research existed in more broadly capturing data from the environment in which people live, which could include elements such as access to healthcare, food, and transportation—each of these being factors impacting health over time.

To conclude, Ms. Smider emphasized the value of collaboration between organizations like hers and research institutions in the development and evolution of standards, which she saw as key to more broad-based innovation, as well as to extending the reach and usefulness of EHRs. Areas requiring considerable attention for future success included identity management in data across databases and deeper considerations of the chasm separating research from clinical, a socio-cultural challenge in which she sees maturity and deployment indexes serving a critical role in the upcoming years.

Discussion

Dr. Ford made a comment about the present-day difficulties of integration between clinical and research activities, mentioning in particular a curious circumstance emerging from the current state of innovation saturation, where at the time of deployment everything might seem to be in place, but with so much changing in so many different areas at once, disruption to an initially integrated process can arise quite quickly out of the gates.

Dr. Embi added that Dr. Ford's point speaks to the need to address governance and social issues simultaneously with the deployment of technology, and the further need to have something like a dedicated unit within research constantly assessing and re-assessing the need for adjustments as divergent evolutions sweep through the clinical and research sides.

Dr. Ford then asked a question about a single IRB policy, adding that the situation at Hopkins, where twenty different IRBs might hold your present protocol, resulted in a great amount of duplication and the need for a shadow IRB protocol. Mr. Rauwerdink responded by saying that the solution so far has involved integration as a way for the CTMS data to participate in the build-up of siloes, but he admitted that there was still progress to be made in that area.