Bio·IT World Briefing On:

Leveraging Clinical and Healthcare Data

This Briefing On features a collection of articles from *Bio·IT World* that shed light on the growing value of integrating clinical and research data, and the steps that innovators are taking to leverage those data.

Leveraging Clinical and Healthcare Data

In this era of post-genomic, personalized medicine, there is growing demand to integrate and leverage the mass of research and clinical data. This has enormous implications not only for the individual patient and the future of health care, but for organizations who see the value of improving diagnostic decisions and finding more rational ways to drive drug development by incorporating patient data.

This Briefing On — "Leveraging Clinical and Healthcare Data" — features a collection of articles from *Bio•IT World* in the past 12 months or so that shed light on the growing value of integrating clinical and research data, and the steps that innovators are taking to leverage those data. John Quackenbush, a cancer researcher at the Dana Farber Cancer Institute, is collaborating with Oracle and others to build a 21st century medical database that integrates genomic information with clinical data, with the goal of improving cancer diagnosis and tailoring treatment for individual patients. Quackenbush said he wants "to build the tools that will allow me and everybody else here to show up on the front page of the [*New York*] *Times*."

Another laudable effort from a major pharma is discussed by Merck's Martin Leach and Ingrid Akerblom, who, working respectively on the research and clinical IT side, are facilitating Merck's efforts to leverage clinical information from its collaboration with the H. Lee Moffitt Cancer Center in Tampa, Florida, among other initiatives.

A new world of genomic data is about to become part of the health care ecosystem. Dietrich Stephan, the founder of the Ignite Institute for Individualized Health, describes his plans for a new health care ecosystem. We also present new commentary on the health care ecosystem from PricewaterhouseCoopers, following on the heels of its timely report on personalized medicine. And we include coverage of the 2010 Bio-IT World Expo keynote talk by John Halamka, CIO of Harvard Medical School, and a key advisor to the Obama Administration's health-IT initiatives.

Rounding out this report, we spotlight Oracle offerings in the clinical space, including electronic data capture and a new product in clinical development analytics.

Our thanks to Oracle for underwriting this Briefing On report.

Kevin Davies PhD
Bio•IT World

Integrating Clinical and Genomics Data

(Originally published March 2009)

How does a physicist wind up at the vanguard of translational medicine, bridging genomics, bioinformatics, and IT in an effort to shed light on cancer biology? That's among the chief responsibilities of John Quackenbush, professor of biostatistics and computational biology at the Dana-Farber Cancer Institute (DFCI) in Boston. An affable personality with slightly greying, shoulder-length hair, Quackenbush is a theoretical physicist by training. He rose to prominence during an eight-year stint at The Institute for Genomic Research (TIGR), founded by Craig Venter, developing and sharing a range of software tools and databases for microarray analysis (see, "John Q: Life After TIGR").

"People think I love building databases; actually, I hate it," he says from his roomy office. "I'm never going to show up on the front page of the *New York Times* with the headline "Quackenbush Builds Integrated Database." If I get there, it'll be because of the discoveries such integrated databases allow me to make. So you could say I want to be able to build the tools that will allow me and everybody else here to show up on the front page of the *Times*."

Whether Quackenbush's efforts become fit to print remains to be seen, but the work he is spearheading at DFCI will likely be felt far and wide in the field of translational medicine.

Bringing Bioinformatics to Cancer

In 2002, Quackenbush was considering leaving TIGR, which had begun focusing on microbial sequencing and annotation, whereas his interest was increasingly turning to the clinical space, beginning with a collaboration with Timothy Yeatman at the H. Lee Moffitt Cancer Center in Tampa, Florida.

Back then, the trouble facing interdisciplinary scientists like Quackenbush was that no one knew where they fit in the traditional hierarchy of academia.

"We really want you here, we just have to figure out where to put you," Quackenbush would hear. "People were very excited about my work in genomics, but they didn't know quite what to do with the bioinformatics part."

He interviewed at several places, even turning down a position as professor of urology at the University of British Columbia in Vancouver, before accepting an offer from the Department of Biostatistics and Computational Biology at DFCI, moving to Boston in 2005. (He also holds a faculty appointment at Harvard's School of Public Health.)

Quackenbush calls DFCI "one of the most progressive places I've seen in terms of thinking about ways to advance science. And I can honestly say it's the least pathological place I've ever worked." New in Boston and driving to work one morning, he spotted a girl selling lemonade by the roadside. She told him one of her classmates had been treated at "the Farber" and her class was raising money for the Jimmy Fund [DFCI's charity]. The level of community support for and patient involvement in cancer research is incredible, he says.

Before hiring Quackenbush, DFCI had recognized that even with genomics becoming democratized, there were opportunities to do new things that were cross-disciplinary. The institute decided to adopt an entrepreneurial model, with the idea of establishing and awarding five years of start-up financial support to research centers that would work across different departments.

"My message during [interview] presentations was consistently about data integration and its value in propelling science forward, which really resonated here," Quackenbush says. Aside from his own research, he has devoted much of the past three years to building the infrastructure necessary for his other mandate: creating a Center for Cancer Computational Biology at DFCI. That could be viewed as a service, but "it's a service to allow me to the things I want to do," he says.

Quackenbush reasoned that the success of such a center would require integrating genomic information with clinical data, as one step toward improving cancer diagnosis and tailoring treatment for individual patients. There would, however, be stiff challenges in linking microarray data with not only clinical information but also public archives such as GenBank, OMIM, and HapMap, while ensuring quality control and reliability.

"Web services are all very well, but you're relying on someone else to maintain the data and not change their protocols," he explains. "Even for GenBank, where things are supposed to be fairly stable, you

frequently see them violating their rules for data entry and standards." Cloud computing could not possibly work in this space, he adds, given the confidential nature of much of the data. Rather than build a large web services model, Quackenbush elected to integrate all this information in a database unique to DFCI.

Around then, Quackenbush crossed paths with Edie Weller, a senior research scientist in his department, during a faculty meeting. Weller, the lead statistician for multiple myeloma, was trying to merge data from different sources—relational databases, raw text files—a nightmarish and time-consuming process involving many Excel spreadsheets. It was particularly frustrating, when designing whole-genome gene expression studies of chemotherapy response for this disease, that she and her colleagues couldn't obtain immediate access to data on their own patient samples, even for information as simple as sample storage location.

"I knew there had to be better ways of merging information and allowing investigators direct access to the data," Weller says. "So although I was initially hesitant to bring it up at the meeting, I finally described how we were linking our data, to

John Q: Life After TIGR

John Quackenbush made his foray out of physics in 1992, when he became intrigued by an initiative from the National Human Genome Research Institute (NHGRI) seeking experts outside biology to work on the Human Genome Project. He spent two years working on the physical map of human chromosome 11 at the Salk Institute, before being hired to set up large-scale sequencing at Stanford's Human Genome Center. When promotion prospects dimmed, Quackenbush headed east to Maryland and Craig Venter's The Institute for Genomic Research (TIGR) in 1997. "The mandate for me at TIGR—going beyond the genome and establishing a microarray laboratory— was really my growing interest," he says.

At TIGR, Quackenbush quickly recognized that there was a woeful lack of tools for collecting, managing, and analyzing the reams of genomic data being amassed. "Our first publication on gene expression in colon cancer included nine arrays, and it was a year's worth of work just to analyze and generate that data," he says. "It really opened my eyes to the challenges and problems with assumptions people have made about biological systems."

Quackenbush recalls early microarray experiments showing that expression of cyclin A1 was a much more appropriate choice of housekeeping gene than the traditional GAPDH, which fluctuated sharply. "What you start to understand," he elaborates, "is that assumptions in biology are often based on little more than gut feelings or historical approaches, and there's nothing better than data to drive a real understanding of what's going on."

Piles of data are essentially worthless without proper management and analysis tools. Given his physics background, however, Quackenbush was comfortable proceeding where most genomicists feared to tread. He continued to write his own data analysis software, creating databases and a variety of open source software tools to help manage the voluminous data being generated at TIGR.

While there, Quackenbush also participated in the scientific workgroup that put forth the MIAME (Minimum Information About a Microarray Experiment) standards. The goal: to allow uniform recording and reporting of microarray data, with the overarching purpose of facilitating the development of databases, public repositories and data analysis tools. It might not be perfect, he says, but these standards have proved handy over the years, especially for finding and correcting errors in published data.

For example, at DFCI he and his colleague Aedin Culhane recently refuted another group's claim to have identified a lung metastasis signature in breast cancer. "Some of the genes they found resonated with us, so we compared their samples with gene signatures in our database and showed that all of the lung metastasis samples fell into the basal-like subtype of breast tumors," he says. "Such tumors are known to have the highest propensity for metastasizing to the lung. What we recognized, looking at this paper, was that they were really suffering from confounding facts. They weren't predicting lung metastasis; their signature was much more highly predictive of the basal subtype than anything else."

Then and now, Quackenbush's creed is that his software tools must be available in the public domain. "It's my mantra: If you're creating tools, they have to be useful, and they have to be used," he says. "If they're not either useful or used, the overall impact is going to be small; ditto if they're just one but not the other."

Quackenbush was considerably irked, then, when TIGR decided to go with licensing agreements for said tools instead. He remains convinced that attempting to write and market software tools in the genomics space is scarcely a winning proposition. Most of the companies that started out along this path have since gone belly-up. "This was hardly in the spirit of what we were trying to do; we were working to advance the science, rather than create tools," he adds.

This led to his staging—along with two like-minded TIGR colleagues, Steven Salzburg and Owen White—what they humorously called the Open Source Revolution, in 1999. "We decided that if just one of us did it, he'd probably be canned; if all three were involved, [TIGR] couldn't do anything," Quackenbush grins. The trio's efforts to release software to the public domain were mostly welcomed at TIGR, since this eliminated the cost of prying licensing agreements from potential users, and increased the number of successfully funded grants. Nor did it hurt that TIGR was then experiencing a lull from soap-operatic drama, with Venter occupied at Celera Genomics.

But Venter eventually returned, and between the ensuing chaos and TIGR's shifting climate, Quackenbush decided to make his escape. He joined DFCI on March 14, 2005—a date he remembers well for two reasons. His son Adam was born exactly one year later, and March 14, as good geeks know, is also Pi Day.

John. He looked at me like I was crazy."

"It was nuts, madness on multiple levels," Quackenbush recalls. It also clearly illustrated the need for merging different data sources together in cancer research. The multiple myeloma researchers invited him to use their case as a framework for creating a data integration warehouse that could potentially be extended to other types of cancer.

Helping Hands

Oracle, with its expertise in capturing and managing clinical data, came to mind immediately as a potential partner in this data integration venture. "There was no point in reinventing the wheel," Quackenbush says. After submitting a proposal for one of the enterprise software giant's commitment grants, he was quickly offered $1 million spread over two years rather than the three he had requested.

And it was about more than just a financial grant. "We also volunteered technical and subject matter expertise to jump-start Quackenbush's plans for data integration at DFCI," says Vijay Pillai, director of strategic planning and business development at Oracle's health sciences division.

Quackenbush, Weller, and Joseph White, the lead database developer in Quackenbush's group, attended half-a-dozen workshops over a couple of months, led by Steve Jepsen, Oracle's senior director for health industries. "We focused a lot on data security and scalability [during the workshops]," Pillai says. "In moving beyond multiple myeloma, you want to be able to adapt to such growth very dynamically, rather than rebuild your environment. From a security perspective, although clinical studies in different therapeutic areas might reside in one data layer, you want to be sure that the investigators can still only access information they're authorized to. So we helped Quackenbush's group think about these design implications."

The result of these workshops and additional brainstorming was a brand new translational research infrastructure utilizing Oracle's Healthcare Transaction Base (HTB) and Fusion Middleware components. HTB creates an integrated data repository, whereby researchers access clinical and patient sample information via a single platform and can seamlessly connect this with experimental data. The Fusion Middleware suite, on the other hand, lets them get to their data securely from any location. And a third Oracle component, the BPEL (Business Process Execution Language) Process Manager, allows for safe and, if necessary, multiple transfers of complex clinical data across the entire infrastructure.

At the same time, Oracle hand-picked a long-term partner in the intelligent software business—UK-based InforSense—to add what Pillai calls their "great visualization capabilities" to the collaboration.

"When you bring different data sources together, you need not just analytics but visualization tools—such as charts and correlation graphs combining thousands of data points—on top of the base integration layer," he says. "We decided that InforSense's applications could help keep us completely in sync, in terms of data integration and interpretation."

Experts at InforSense suggested ClinicalSense, a web-based tool for clinicians and researchers to get summary statistics about patient populations by fashioning row-and-column matrices out of patient sample attributes. "You could construct a query where 'regimen response' represents the rows, while 'sample count' and 'sample type' are chosen for the columns," explains Mick Correll, InforSense's director for clinical solutions. "This would result in a cross-tabulation matrix showing the number of available samples broken down by type—tissue or blood, for instance—and grouped according to the patient's response to a particular regimen." In other words, users can build more sophisticated queries by defining a hierarchy of attributes, which then enables them to "drill down" into the results matrix, further stratifying the population.

"It provides, I think, a very rich and interactive web experience that makes the data come alive in the hands of clinicians and researchers," Correll says.

ClinicalSense leverages InforSense's next-generation business intelligence platform, besides providing an advanced clinical data model. It's both easy to use and intuitive, thanks in large part to direct feedback from clinicians throughout the product development process. "The multiple myeloma study at DFCI is precisely the type of problem ClinicalSense was built to solve," Correll says. "It provided the right balance between out-of-the-box functionality, and flexibility that will enable it to adapt to changing needs in research."

Institutional Barriers

Like most large-scale collaborations, this one wasn't without its hitches, particularly with regard to the people responsible for Information Systems (IS) at DFCI. "With these folks, whenever you ask them a question—no matter how benign—their first answer is always 'No,' since no access is the most secure access," Quackenbush says, only half-jokingly. "The word 'fragmentation' has been used to describe this whole problem of having data in different places; I like 'Balkanization' instead, because not only are the data being broken apart, there are all these people actively fighting against integrating it. We spent more time and effort negotiating transferring data into this warehouse than we did actually building the warehouse."

"There was a lot of confusion with IS about the scope of this project; how it fit or conflicted with other IS initiatives," Weller says. "With how quickly research moves, we felt it was imperative to have individuals who understand the biology, as well as systems aspects, working on the project. Once we discussed this in detail and described our data security model to the IS, things were much easier."

Nevertheless, Weller adds, such regulatory issues—especially those involving redistribution of data collected from different hospitals—are hardly minor. "I think the time we spent sorting these out will benefit not only our myeloma project, but other initiatives too," she says.

Upon overcoming these hurdles, the collaborators rolled out their prototype for an integrated data warehouse in May 2008. The warehouse's full production system has been up and running since November, after two training sessions—one for statisticians, data managers and researchers; the other for clinicians—to teach them the art of accessing and querying the database.

"In both cases, I think the new system was well received, and the feedback we've had since has all been very positive," says Correll, who led the training. "I'm sure modifications will be necessary as it moves forward—this is research, after all—but we're clearly on the right track."

Several of Quackenbush's colleagues, including a few skeptics, were invited to sit in on both sessions. "I remember Beverly Ginsburg-Cooper [senior vice-president for research at DFCI] grabbing me by the sleeve, five minutes into the second ses-

sion," he recounts with a smile. "This was after she had watched us go from nothing to a group of patient samples with certain clinical characteristics based on karyotype and trial response, to their gene expression profiles, to a set of genes correlating with response, to PubMed records describing the genes—all as ad hoc queries." It dawned on the skeptics how this would excite young people doing research at DFCI. "This system presents information in a way they're comfortable with; they feel invested in the process and better able to participate in data analysis, to see how they can drive things forward," Quackenbush adds. Ginsburg-Cooper even called it "transformative for research."

DFCI's first integrated data warehouse has been constructed architecturally so there's a path to move forward, Quackenbush says. The institute will pour $8 million into his cancer computational biology research center over the next five years, which he considers "less than we need, although the center's built on a model I think we can expand." He is thus seeking additional funds to accomplish his goal of moving the data integration system beyond multiple myeloma, with breast cancer as the next likely candidate area.

As well, Quackenbush recently applied for a grant that, if approved, will include funding to create a pilot implementation for data from the Nurses' Health Study (NHS) at Harvard, the largest and longest-running investigation of factors influencing women's health. And he has been communicating with multiple myeloma researchers at the University of California, San Francisco, about the possibility of a mirror installation at their end to facilitate data sharing between both groups.

"I don't know where this will go next, to be honest, but it's likely to go somewhere," Quackenbush says. "Our successful collaboration with Oracle and InforSense has put us in a position to think about reaching beyond DFCI and gradually pulling in a lot of Harvard's multi-institutional spores and their data collections. So there'd be some method to the madness."

He greatly appreciates how DFCI nurtures the importance of continued research. "You always hear about Harvard eating its young," he muses. "I came here a little worried that I was going to face all these prima donnas. It's not to say that there aren't those with pretty big egos, but there's a really high level of collaboration here, which is both astonishing and impressive."

Quackenbush embraced interdisciplinary research early on, and now observes that many scientists in his area of interest are spanning their boundaries, which he definitely endorses. "A computational model is just that; a model plus validation is a discovery," he remarks. "People are really trying to drive the latter, rather than being held captive to someone else's experiments. It isn't true of everyone in the field, but I think it's an emerging trend; a systems biology approach that is evolving naturally out of genomics and bioinformatics."

Merck's Informatics Mission

(Originally published May 2008)

After five years at the helm of Merck's basic research IT group, Ingrid Akerblom calls her move to the clinical side "quite an eye opening experience." Akerblom has a Ph.D. in biology from University of California, San Diego and the Salk Institute, and later joined Incyte Pharmaceuticals as its 50th employee and "annotation guru," eventually leading informatics. She was then recruited by Merck — ironically the only pharma not to buy the Incyte database. She joined Merck in November 2002 — just over a year after the acquisition of Rosetta Inpharmatics. Akerblom worked extensively with the Rosetta IT leaders, helping to integrate systems around target identification and chemistry systems.

Assuming her former role is Martin Leach, a Brit who spent nine years leading IT and informatics at Curagen, spanning corporate IT, basic, pre-clinical, and regulatory informatics. But he also brings experience in regulatory and clinical IT areas gained during a two-year stint at Booz Allen prior to joining Merck last year. That clinical insight could prove useful even as he refocuses on basic research, and complements Akerblom's background in basic research as she transitions to the clinical side.

Kevin Davies spoke to Akerblom and Leach about their complementary roles and mutual understanding of the needs of both basic research and clinical teams, which could pay big dividends for Merck.

Bio•IT World: **Ingrid, how did your move to clinical IT come about?**

Ingrid: After five years in [research IT], with the last few years including leadership of Merck's Biomarker IT, it made sense to bring some of this expertise into the Clinical IT areas in order to meet the growing need to marry up clinical and discovery research information. In terms of how we operate, there has also been a significant evolution of the IT teams and operating model. At that time it was fully vertical, integrated. I had all the developers on my team; we had all the support on my team. Now, Martin and I really lead more of a client services team, where we have account managers, program managers, and business analysts, people with business expertise and technology expertise. Most of the delivery is done through shared services in the corporate area. And even within the Merck research labs IT, for some of the more innovative types of things that go on in research that aren't found in the other divisions like manufacturing and marketing. But we're evolving to a fully shared services model, which has its benefits, especially in clinical, with large projects.

Martin: For my groups, I have scientific computing, which is predominantly in the bio space, some cheminformatics, biomarker IT, of which there's a lot of collaboration with Ingrid. And drug discovery services, focused on lab automation, capture and management of biologic and chemical data and information, all of the things that make the basic research lab tick.

How useful will your mutual cross training prove, do you think?

Martin: One thing important to note is the new head at [Merck subsidiary] Rosetta. Kathleen Metters, the worldwide head of basic research, recently appointed Rupert Vessey to be the new head of Rosetta ... Rupert is a former clinical therapeutic area head, [so] basically we have a clinical leader heading up Rosetta, which is predominantly working on genomics, proteomics, genetics, and causal networks. So it's not just the IT with that cross-pollination — we've got two people from IT, from a clinical point of view and a basic research point of view, interacting with a former clinical person who is heading up the genomics space.

Ingrid: We've invested a lot in some core platforms; we need to start translating that into results in the clinic at some point. And so having people who have an understanding of what does that really take to help inform the earlier research directions, the platform directions, is a key theme...

When I was in Martin's position, it was very difficult to get the clinical IT teams to focus on longer term strategic projects, even short-term partnerships that weren't about a late-stage trial. Because at the end of the day, that's what they work for, right? They've got to get those trials filed. But when we think about the future, we want to have our data more integrated, and we weren't really getting a lot of traction. So one of the attractions for me moving to this position was someone who has that background will keep their eye on that ball and it won't be all about late stage. That's already proving true — there were a couple of times this year where there were scheduling conflicts between critical projects on both sides. And in the past, I know

which one would have gotten dropped — it would have been the basic research project.

How is Rosetta working at Merck today?

Martin: There's Rosetta Inpharmatics, which is the part of Merck that's doing molecular profiling and genetics research, and then we have Rosetta Biosoftware, that is part of Rosetta Inpharmatics that makes and sells software products such as Resolver, Elucidator, and Syllego, which we also use internally at Merck.

At Rosetta Inpharmatics, I work closely with scientists working in the bioinformatics and the pathways space, who have taken a biological point of view to integrate information. One approach is trying to integrate as much information that is accessible, assay data and so on, for when scientists pull up a gene or target. They have developed a target gene index (TGI)… With a given gene, you can see all the relevant information. I think most pharmas have attempted that. I'd say it's the depth of the information and integration with some of the chemical space that is different than what I have seen at other pharmas. This depth of integration within TGI is still growing… We do data integration and data management within basic research IT, and we provide some of the core services needed to do it from a research point of view.

How do you interact on a more day to day level?

Martin: We have some very high level, strategic, long-term projects that we're working on. We have a large number of folks from my camp working with a large number from Ingrid's camp around the IT needs and implications with all the different clinical data, as well as sample data, and access to this information that's needed to enable translational research. So we have joint projects, very strategic, they have visibility all the way up to the MRL leadership.

In terms of some of the things being done at Rosetta, again, it crosses into the basic and clinical space, and we work together on making sure the right people are engaged in either basic or clinical IT. [Between the basic and clinical IT teams] we are very collaborative in terms of key strategic hires.

Ingrid: We're getting much closer to actually using genetics in our trials, based on the technology set up by our Seattle genetics group and the whole genome analysis group (See, "Merck Ties Gene Networks to Obesity"). We have a project team meeting with Martin, our business and information architects, and Rosetta Biosoftware together with clinical franchise and regulatory leaders, to talk about what is the actual proposed data flow and architecture for moving genetics data from research systems into the clinical systems. Having formerly been in basic [research], it's a lot easier to really see how that all fits together and how to move this data into the clinical systems now.

The Rosetta Biosoftware Syllego system that is being used by the FDA, is something we're looking at — How does that fit into the clinical architecture? We have a clinical warehouse, where should the genetic data go? Should it be Syllego for raw data and CDR for metadata? Again, it's moving into reality now, so understanding what that means and being on the clinical side I think is going to make it a lot easier to easily assimilate that type of data into the mainstream clinical systems.

My Basic IT team worked with the Imaging Research team to put in place an imaging platform with IBM, and Martin's team is continuing this work, that's working well in the early development and research space. Now I want to say look, we could save a tremendous amount of money if we move that into the late stage. But how to do that where every investigator now has to learn that system?… Do we show it through our portal or does it come in through EDC or on its own? So there are all these support issues once you start thinking about really getting out into the clinic with some of these newer things.

How are you handling the surge of data, especially related to genomics?

Martin: Where we are doing work on pharmacogenomics and genetics in the clinical space, there is so much data. For example, one of my team had to secure an additional 100 terabytes (TB) on the East Coast to just accommodate one experiment they were doing! Soon, I'm going to be playing around in the petabytes… At the moment, we need to keep the raw data because there's no clear guidance from the FDA as to what you need to keep. It's going to literally swamp us working in this space until we get better guidance around what data we need to keep versus could keep. One of my [team's] projects is basically a storage strategy this year because if it's 100 TB this year, it's probably going to be a couple of hundred TB next year…

We all [in the industry] have data and document retention policies, but what tools do we have to really monitor and manage that? If I've got a couple of hundred TB that's going to come around in the next couple of years, how do I know what to purge five years from now? Where are the tools to do that really large data management and purging? In the current file sharing landscape we have millions of files that normally have to be managed through retention policies. That's a challenge in itself. What is developing is managing a fewer number of files but with a large overall volume.

How do you view translational medicine?

Martin: In two parts. The first part is increasing the clinical context of basic research experiments, using clinically relevant samples with their clinical information, allowing you to "translate" additional research measurements on the samples with a clinical context. So that enables the research, but then as you get into the pharmacogenomics space, where you're looking at genetic information to segregate populations for responders and non-responders, that's then taking basic research discoveries and really applying them into the clinical space. So I sort of see translational medicine as that mix of pharmacogenomics and biomarkers and everything rolled into one.

Ingrid: I agree. One of the key areas is clearly samples, whether you're doing proteomics, gene expression, genetics, or potentially looking at what populations eventually could respond to your drug. Samples are at the center of that and so we have been actively pursuing better informatics around that in order to make it clear what samples are available from what trials, which are consented, and what can we use them for. We already have siloed platforms to show that data, we need to integrate it more than it is… We have a new standards-based clinical warehouse that went into production last

year, where we're really planning to have all the patient data — whether it's through collaborations or Merck trials — in one place so that it's more available for our future data mining and understanding what types of patients and associated samples we have.

Martin: We have a major strategic collaboration with the H. Lee Moffitt Cancer Center [Tampa, FL] (See, "Cancer Center Builds Gene Biobank," *Bio•IT World*, June 2007.) We get different types of cancer samples and those samples go to Rosetta [for] expression profiling. Moffitt uses that expression data internally for their research, and we get clinical data associated with the samples, as well as the expression profiling data, and we get to use that at Merck... This is a major collaboration driven by Stephen Friend [senior VP of Oncology] and the Oncology franchise. I think it's a landmark in how we approach translational medicine at Merck... Data from this collaboration was the first clinical data from oncology that made its way into Merck's clinical data repository (CDR).

So we have clinical data securely flowing directly from Moffitt through the firewalls, etc., into Merck's CDR meeting all compliance needs. And that data through web services is then shared to Rosetta and other places so that it can be integrated with expression profiling data. We've really embraced industry standards to make that happen. This really has been breaking down silos — it's very hard to find a clinical group that opens up web services where that information is then accessible to basic research. I think that in itself was groundbreaking at Merck. We've tried looking around to other pharmas, like are you guys doing this sort of thing? Everyone is talking to the standards boards, but I think we've really [made] an investment by implementing some of this work in a real active strategic collaboration...

Ingrid: The other important piece in that project that addresses the translational medicine question is that there are joint project teams between Merck and Moffitt clinical and basic researchers, all trying to mine and look at field experiments, build trials, identify new mechanisms, think about the future together. It's a very powerful collaboration, and IT has a seat at that table and is an active participant in those conversations. So I think it's a great area of translation where we really are leveraging clinical data to drive research.

How has informatics evolved at Merck? With budget tightening everywhere, does that impact the build-buy decision?

Ingrid: Ever since I joined, we've been primarily a buy shop, even in basic research. We mostly buy and we try not to customize too much, but you still end up in that space. The clinical systems have been primarily internally built, and now they are mostly purchased, with the exception of the data warehouse, which is based on the Janus data model, but it was still built inside with outsourcing. Where we're trying to find cost savings is in sharing services, particularly around support, maintenance of applications, infrastructure — trying to drive down cost on the maintenance and operations side in order to continue to invest in the new development of strategic applications.

There are innovative areas with many of them in the emerging research technologies where you're doing things that you just can't buy, where faster iterative in-house development is needed, for example we developed MouseTrap to support the management and display of animal phenotypic data. Generally speaking, it's quite a challenge. You've got to really be focused and the business has to partner with you to prioritize... it's critical to have a strong partnership and governance with scientific leaders to assure we are focusing IT resources on the right projects. The other thing is they're also feeling the money pressure. So it's not just IT, it's not just the services anymore, it's everybody really looking at how are we going to contribute to optimizing the bottom line, and how are we going to grow the top line, and let's all prioritize those initiatives together.

What are some projects where you think you're really going to be able to expedite or make better decisions? And what outstanding challenges remain?

Martin: We're in a position now where we know how to generate information for biomarkers, and we know how to collect clinical information. So at least one project this year is, "What is that killer application that you need to integrate the clinical information with biomarker information, so that we really do enable our scientists in their biomarker discovery or validation experiments?" At the moment, we've got bits of the puzzle — genetics being managed in Syllego, expression managed in Resolver, proteomics managed in Elucidator, these all being separate applications and repositories. But what is that killer application that brings it all together and integrates it with clinical, so that you can do some meaningful mining and analysis? That's one of my goals, and I've got some exciting challenges to work through there.

Ingrid: I think that's a shared one, because in the clinical sample area, combining the results data from clinical samples with the associated patient data, what's that platform? I know there are new commercially available things coming out like Azyxxi from Microsoft. So we need to be looking at what's out there, what's the gap, and do we put something together ourselves? We did a pilot last year with an EII platform collaborating with IBM. There was enough productivity gain from that to justify taking EII to the next level which our Innovation IT team is doing in 2008. The whole integration space and then the actual viewing of integrated data in a meaningful way continues to be a major focus.

We're embarking on an electronic medical records (EMR) strategy looking for signal detection among other uses. We're redoing our pharmacovigilance system and approaches. Those are things that are just starting to be reinvested in, figuring out how do we leverage that information, how do we get that connected? There's also appetite for clinical trial simulations across a number of dimensions including enrollment and operations optimization. We just overhauled our entire late-stage development systems in 18 months, so right now we're focused on ensuring that that gets optimized and the value from that investment gets realized.

Martin: Who is going to be the health partner with Merck? Where do we place our bet in key strategic partnerships thinking around EMR data or personal medical record data, and how do we find the best partners to enable translational research? From there it's doing the analysis of who will be the best partner and when will they be mature enough or Merck mature

enough to interact with them.

Another exciting challenge is working with the external basic research team, [Catherine Strader, former head of research at Schering-Plough]. I'm working with her so that we really leverage information from our collaborations. In the past, how information flows in a collaboration has been managed ad hoc. Moving forward we really want to leverage and integrate this information more strategically.

What roles do the senior executives such as Peter Kim and Stephen Friend play?

Ingrid: Peter has a vision. He focuses us all on recognizing that the vast majority of information and innovation is happening outside the walls of Merck. We need to leverage it more by providing platforms that allow deep collaboration with external partners; there also is a focus on combining our own data with publicly generated data for competitive advantage — but holding the line to work pre-competitively where it makes sense. You get that vision through the research strategy meetings.

I think Stephen Friend is clearly a visionary who inspires many individuals at Merck both on the science side and the IT side, a very forward thinker pushing all the teams, Rosetta as well as myself and Martin, to think out of the box.

Merck Ties Gene Networks to Obesity

In March, scientists from Rosetta and Merck published a pair of papers in Nature identifying changes in gene networks associated with obesity. The team, led by scientific executive director of genetics, Eric Schadt, is deploying a more holistic approach to the pathogenesis of common diseases — not merely searching for gene variants, but measuring gene expression in tissues from obese humans as well as mouse models, which is coupled with information on DNA variations and clinical data. Massive computational analysis — the equivalent of 7,000 CPUs — pinpointed entire gene networks perturbed in obesity.

"Common diseases such as obesity result from genetic and environmental disturbances in entire networks of genes rather than in a handful of genes," says Schadt. "The accurate reconstruction of these networks will be critical to identifying the best therapeutic targets."

In one study, Merck researchers and scientists from UCLA identified DNA variations in mouse tissues associated with obesity, diabetes and atherosclerosis. Schadt and colleagues built gene networks and identified the constituent genes implicated in the various diseases, notably three specific genes — Lpl, Pmp1l and Lactb. In a separate paper, Merck scientists collaborated with deCODE Genetics and Iceland's National University to construct obesity expression networks using tissue and clinical data from more than 1,000 Icelanders, in large agreement with the mouse work.

FURTHER READING:

Chen, Y., *et al*. Variations in DNA elucidate molecular networks that cause disease. *Nature*, published online March 16, 2008.

Emilsson V. *et al*. Genetics of human gene expression and gene-gene transcriptional networks. *Nature*, published online March 16, 2008.

Ignition Sequence Starts

(Originally published March 2010)

In 2009, Dietrich Stephan, the co-founder of personal genomics company Navigenics, announced his plans to build an ambitious new academic institute in Northern Virginia—the Ignite Institute for Individualized Health. Ignite is partnering with Life Technologies and acquiring 100 SOLiD 4 instruments, which can deliver $6000 human genome sequences. Kevin Davies spoke with Stephan about Ignite and his choice of sequencing partner.

Bio • IT World: Why so many sequencers and why did you select SOLiD 4 instrumentation?

Stephan: We're really excited about this partnership. I feel very good about the SOLiD technology, namely its accuracy, price points, and throughput. Having reviewed recent data, it looks like the SOLiD platform is extremely accurate. I really like that, because we don't have to sequence with as much redundancy. We're really creating a strategic partnership with Life Technologies to do co-development around the technology and understanding the interpretation in a clinical setting of the information. So it's more a strategic partnership that swayed the decision…

I think we're finally approaching a price point for full-genome sequencing where we can realistically start to redo all the whole genome sequencing studies (GWAS) that have been backed up for the last five years waiting for this technology. You need some horsepower to do that, and we needed some infrastructure…

We are also … putting the technology into a CLIA environment and using it in a clinical setting. For example, is there a role for sequencing the entire cancer genome out of a diagnostic biopsy from someone newly diagnosed with cancer? Pull out single cells, sequence them, and try to understand if we can prioritize the standards of care, and develop salvage therapies for the 50% of people who will eventually progress… We want to start learning how to apply the technology in a clinical setting and how to interpret it immediately.

Do you really feel it is worth revisiting all the GWAS data at the full sequence level?

I'm of the bent that all these heritable risk factors will be relevant in the aggregate. Common variants were extremely valuable in many cases, not only in risk assessment but as guideposts to uncover rare variants in those portions of the genome. But there's epigenetic modification and copy number variation that we can now capture with the current technologies. We're finally at a point where we can, in a systematic way, go back and rip through all those case-control studies and get close to capturing the totality of heritable risk for complex genetic disease—step 1—and then maybe even start to sub-classify those common, complex diseases, and re-name those molecularly homogenous sub-types which will likely be differentially triggered by environmental exposures and be responsive to different therapies. I believe this systematic discovery strategy will form the foundation for individualized medicine, which should improve clinical outcomes.

What will be the focus of your sequencing efforts in the clinical realm?

It's an extension of Dan Von Hoff's work at Scottsdale Health Care, where he ran a beautiful clinical trial using expression profiles of patients. When someone presented with end-stage cancer and was asked to be assigned to a Phase I trial, Dan would randomly assign a Phase 1 trial, the other study arm he would profile and then intelligently assign to a trial. He's presented beautiful data that he can improve response rates. So [we're] trying to use that same strategy but with the next evolution of technologies to see if we can't improve outcomes with experimental or off-label therapies for those individuals who progress through current standards of care. Our work will begin in a clinical research setting then as a clinical service and hopefully soon make a similar strategy the standard of care to optimize outcomes.

You used the term 'genome institute,' but you're also stressing that Ignite is about *individualized* health?

There's a couple of aspects in the Ignite model that we think are unique… One is the horsepower to drive out pure molecular subclasses of disease as a starting point for personalized medicine. So renaming complex genetic diseases according to new nomenclature around molecular homogeneity. A second is aligning the disease foci to the market needs out there in the world. Basically that means the most prevalent, intractable diseases will be the ones we research.

The third aspect will be that we'll have commercialization infrastructure surrounding the institute so we can make early go/no-go decisions. We can model the economics of new tools and strategies. For example, who is going to pay for this when we turn it on? So if you take a big genome institute, surround it with a commercialization infrastructure, and then place that in close partnership with a health care system, hopefully you can build those trusted relationships to quickly move from a robust discovery around a pathogenic pathway to a little biotech or drug company and then

into a clinical setting quickly… Once you have a little company it often takes years and tens of millions of dollars to educate doctors, get regulatory approval, and change the standards of care and get any of them to use it… We felt a need to compress the cost and timelines.

Finally, we view our project as a national and international resource and look forward to partnering and collaborating extensively. The permanent facility will be located on a campus so that we can expand operations as these collaborations and spin-out companies take root.

What's the status of settling on a permanent location for Ignite?

We're currently in the Center for Innovative Technology building near Dulles Airport, it's a state-owned building (the black upside-down pyramid building visible from Dulles). The state will put in some money to refurbish it, and we'll leave it behind as a biotech incubator for the Commonwealth of Virginia. On the permanent site, we're still evaluating a couple of options through a process led by, in my opinion, the best full-service health care facilities owner representative in the country—Nancy Kelley of Murphy-McManus. It looks like we'll purchase a pre-existing shell so we can go in and very rapidly refurbish floors to grow in real time.

Let's talk a bit about Navigenics. You have a new CEO, Vance Vanier?

Vance is spectacular. He embodies the spirit and values and personality of Navigenics in a perfect way. He's great!

What is the significance of the recent New York State granting Navigenics regulatory approval?

It's a gold seal of approval, if ever there was one… I think that was the most important regulatory thing that's ever happened [to Navigenics], simply because it is so carefully done in New York State. It allows the Navigenics Health Compass to be sold in New York through a physician channel [not direct to consumer]. I'm happy with that—it's a medical diagnostic product. I'm thrilled that they have that kind of channel.

With the Navigenics Health Compass celebrating its second anniversary, what are your thoughts about the evolution of consumer genomics?

It's a great question. I joke that I had my bullet proof vest on for a few years there in the scientific community. That topic has been the singular most interesting development is the scientific acceptance of risk strategy, as far as I can tell. People are now comfortable with the science… The concept is still as fresh today as it was back then—the notion that personalized preventive risk stratification is critical to reducing health care spend and the burden of disease… The Health Compass infrastructure was built to be able to turn on these personalized preventive risk estimates, and over time we would populate the rules engine and the technology engine with more robust things. And that's what we're doing.

You tried to seek areas of common ground with the other DTC companies regarding risk calculations and assumptions?

Yes, the Navigenics team did propose the standards setting exercise, and there has been movement toward convergence around the prevalence estimates around the diseases. Navigenics still doesn't test for quantitative traits and non-medically relevant things—that's an area lacking agreement. We probably don't have perfect convergence now but we're heading in the right direction.

One of your original arguments for Navigenics was the soaring health care cost. How do you regard the ongoing health care debate?

I think it's been a healthy discussion around increasing efficiency and reducing redundancy in delivering health care and I hope we see some progress. I think we can cut tens of percents off our health care spending with reform. But that's not going to change the slope of the burden of chronic disease—it will only push it out by a couple of years. I hope that Round Two follows shortly, where the discussion focuses on truly altering that slope, otherwise the same thing is going to happen. Extending a healthy lifespan and reducing the time you're sick is still I think the key to all this.

Harnessing (and Securing) Meaningful Data

(Originally published May 2010)

"There isn't going to be some massive database in the basement of the White House run by Sarah Palin," promised John Halamka, the CIO of Harvard Medical School, in his keynote at the Bio-IT World Expo. But there will be a "federated mechanism that enables us to send data from place to place for a whole variety of purposes for care and research."

Halamka serves as the Chair of the US Healthcare Information Technology Standards Panel. Of the $30 billion allotted to health care IT in the Obama Administration's stimulus package, most of it will be distributed to hospitals and clinics after they've put health care IT infrastructure in place and are using it wisely. The remaining $2 billion is being distributed by the Office of the National Coordinator for health care IT advances.

"Here's the strategy," Halamka said. "Give $2 billion in grants to accelerate the industry. Give the industry a set of standards that are unambiguous for everything from medications, to labs, to quality measurements, to both clinical care and population health… Declare how hospitals and doctors have to use this wisely, and then certify products as being good enough to have the features and functions and capabilities to make this whole thing work."

In the next five years as these standards are put into place, doctors and hospitals will be required to collect "meaningful" data and protect that data. "This is not using a word processor to record data!" Halamka clarified. "This is actually using codified mechanisms so that if you capture medications, problems, allergies, labs, etc. You could use them to inform drug discovery."

Meaningful data include requiring all orders to be electronic; recording medication and allergy lists for all patients; and recording and updating demographics and vital signs in a timely manner—all using consistent and controlled vocabularies.

There are rules that still need to be clarified, Halamka says, but gathering meaningful data will begin to enable smarter health care. Offices and hospitals will be able to mine their data and send targeted wellness reminders for preventative care. Patients will have access to a continuity of care document, delivered electronically, explaining treatment history, prescriptions, and diagnoses.

But with this wealth of data—and wealth of possibilities—comes a huge security responsibility. Halamka stressed that laptops and thumb drives must be encrypted. Patient privacy must be protected.

"I spend about $1 million a year just protecting the Beth Israel Deaconess [hospital] records against the nefarious internet. We're attacked every seven seconds, 24 hours a day, seven days a week," Halamka said. "Half of the attacks come from Eastern Europe; half of the attacks come from Eastern Cambridge. Every September, 1200 new hackers arrive – they're called freshmen!"

Emerging Data

Although about 20% of clinics and hospitals currently have electronic health records, all should by 2015. Public health applications are already beginning to emerge. Health data gathered can be aggregated regionally to look at public health trends or build doctor report cards. The SureScripts (representing pharmacies) and RxHub (representing payors) database includes de-identified drug information on 160 million people that can be used to check for drug interactions.

In another example, the Social Security Administration used to spend $500 million a year getting paper records. After moving to electronic records two years ago, a disability claim that took months to adjudicate can be handled in 48 hours.

There are even opportunities to gather rich patient data in the home. Halamka is testing a bathroom scale that calculates his lean body mass and body mass index and transmits the data via XML in real time to Google Health and Microsoft Health. But although an avid blogger and registered technophile, he said he declined the Twitter reporting feature.

Convergence is Key, says PricewaterhouseCoopers

(Originally published March 2010)

In a report on The New Science of Personalized Medicine published late last year, consultants at PricewaterhouseCoopers (PwC) painted a rosy picture of the prospects for a $232-billlion market poised for annual growth of more than 10%. One of the chief trends that PwC sees isn't growth but convergence.

"We've got three convergence factors," says Michael Mentesana, a partner in the health industry advisory group of PwC's pharmaceutical and life sciences division. "Payer, provider and biopharma all need to work more collaboratively together and supported by the government."

"What we need are fundamental changes in the ways of working. Technology will continue to move, but we need changes in behavior. We need a culture of acceptance of personalized medicine. If we get that, then personalized medicine will stop being a pipe dream happening in small places and really be the catalyst for change." Mentesana says the PwC health advisory practice issued the report "because it drives us internally to think about industry challenges and put out a vision and a point of view"—looking ahead in a holistic sense, providing insights that clients might not see, particularly an increased focus on a customer centric view.

Blockbuster RIP

"There is no doubt that the blockbuster model will no longer be the sole engine that drives big pharma. I mean, there will be more blockbusters, but it will not be the fuel that it has been in the past for the pharma industry. All the low-hanging fruit has been tapped. The advances in 'omics, the human genome map, targeted therapeutics, the continued use of smaller patient populations, orphan drug status, you start looking at more selectivity... The focus has evolved from treatment to prevention and, hopefully, curative therapies."

"Convergence—that's what has to happen," Mentesana continues. He cites interest from pharmacy organizations in diagnostics and delivery of care. "They are taking things out of the hospital setting and into the home setting. That convergence in itself would be huge progress."

Other convergence factors include rising health care costs and our aging population. "As collaboration continues to get more and more closely aligned, you're going to see the emergence of a new hospital structure of the future," tied together with companies such as Google and Microsoft.

Art Karacsony, PwC's director of US pharmaceutical and life sciences marketing, agrees that many non-traditional companies are taking a closer look at personalized medicine, including software vendors, internet providers, and large retail organizations. They will have a role in supporting the move toward personalized medicine to provide more data to make more informed decisions.

Improvements in the ways drugs are developed and managed won't reduce the drug development cycle overnight, because of issues surrounding safety and efficacy. "But they'll get more efficient drugs that are transforming the cost-of-care so the industry will get the premiums they need to fund their R&D ... It's about efficiency, reducing cost of care, and outcomes."

In the Clinic

Mentesana says the prevailing attitude in pharma is to view personalized medicine as "an inevitable conclusion. It will allow us to get to more patient populations, even though they might be smaller. The industry will have higher success rates if the drug or therapy delivers what has been promised ... If this works, the industry can help take the cost of care down."

"From a culture perspective, the pharma companies need to keep working with the FDA to validate more selective biomarkers. These methods need to be accepted by the industry, but also by the providers to implement it in the hospital setting and the payors. That's why the convergence engine is the most important one ... Everyone collaboratively working, in a pre-competitive space, on selective biomarkers ensure no one misses the boat on personalized medicine."

Mentesana sees technologies such as *in silico* modeling as potentially disruptive. The combination of animal models with *in silico* models will be a major scientific innovation and "will push the companies closer to personalized medicine."

"I'm really excited with the idea of having everyone [being able to get their] electronic health record," facilitated by the likes of Google and Microsoft. "I'd love to see a day when the technology is on [my] mobile device so I can update my medical record. Whoever cracks the code, and

gets it accepted and connects it all to the hospital—that will be the way forward."

PwC sees a growth business with the "non-traditional players" in helping them develop new business models and organizations to support personalized medicine. "Some companies are also asking us to help with organizing funding. Or they ask strategy questions, for example: What does personalized medicine mean to us? What should our strategy be? What is the business model for our organization and how do we organize funding? To be fair we are in the infancy of this process."

FURTHER READING:
The PwC report on The New Science of Personalized Medicine can be obtained here: http://www.pwc.com/us/p4.

Oracle's Patti Gaves on EDC and Integration

(Originally published October 2010)

In a recent conversation with industry veteran Patti Gaves of Oracle Health Sciences Global Business Unit, *eCliniqua* was curious about her perspective on the current status of electronic data capture (EDC), the industry's strong focus on integration of electronic point solutions, and the evolution toward eClinical trials. Gaves, senior director of Life Sciences Product Strategy, has more than 15 years of clinical data management experience and has worked in customer implementation and operations management.

eCliniqua: Data suggest that approximately 50% of clinical trials are using EDC today. How has the market changed over the past decade?

Gaves: If you look at where we were ten years ago, when we were trying to sell the benefits and concepts, we are definitely in a mainstream situation. I hear objections to EDC less and less. I was with Oracle when we were introducing EDC, and back then, there were a lot of conversations about process adoption, and what you need to implement EDC trials successfully. When you talk to customers today, especially large ones who have been working with this for ten years or so, they are running about 80% of studies with EDC, and all new trials are running with EDC.

Going forward, what can we expect in the way of EDC adoption?

I see it starting to plateau in certain segments, such as some of the large customers, but there are those segments that are down in the continuum who have yet to come through the whole process. For some, such as academic medical centers, and small regional CROs, they are still coming along. In certain regions of the world, EDC is slow to take up due to some infrastructure problems in parts of Africa, Asia, and the Middle East. It's still paper-bound there, but starting to evolve. Also, in the post-marketing surveillance world, there is still quite a bit of paper.

Do you see clients starting to think about eClinical trials and the integration of point solutions?

Absolutely—this is where we are headed. EDC is about having data sooner and doing something with it, but there is a need to leverage that technology with other solutions. The question is—can you access the data beyond the application in which it resides? The answer is often "no." Oracle is working toward a fully integrated suite so accessing that data from any application will be possible. The cost of ownership around having best-in-breed in every application and trying to keep them hooked together and working is becoming insurmountable. The idea of moving to a more open, less proprietary approach, using Web services and services oriented architecture to facilitate your own product integration and out into the world is really a high priority right now.

What is Oracle's approach to integration?

Oracle-wide, we have a strategy, and accompanying toolset to support integration called Application Integration Architecture (AIA). This approach incorporates standards-based technologies including XML, Web services, and Business Process Execution Language (BPEL). This allows customers to automate business processes across the enterprise using Web services. Oracle uses this methodology to package prebuilt integrations and maintain them across product releases. Customers and partners can also use the toolset to develop integrations independently. For health sciences specifically, applicable industry standards include CDISC, HL7, and BRIDG model.

Are most clients looking to integrate two solutions at a time, such as EDC with a clinical trial management system, or are they looking to integrate many point solutions at once?

We are seeing a lot of customers recognizing that, as they look across the enterprise, they need the same information, and they want to access it while minimizing redundancy. They only want to define things once. This is one of our pain points—pieces of information being consumed downstream by upwards of fifty systems in big organizations. There is a big impact downstream in having the integration right. Historically, integration has been two solutions at a time because of the high maintenance cost associated with this, but that's changing. The uptake is aggressive compared to the EDC timeframe. There will be a lot of movement toward integration and interoperability in the next five years, whereas it took about ten to adopt EDC.

What's driving this rapid change?

It's about an eClinical evolution. It's hap-

pening very quickly because of the climate—economic and political. There are a lot of cost-containment and health care reform pressures that will put a lot of cost pressures on clinical development.

Any final comments?

There is a lot more technology outsourcing to CROs, which represents a shift in segmentation for us. The sponsors are less committed to the technology now because maintaining it is not their core competency. They have investments that they are preserving, but as we go forward, and depending on how their work is allocated—they may care less and less about whose product is being used. As a result, there will be more of a focus on CROs.

Oracle Clinical Development Analytics Could be a Game Changer

(Originally published February 2010)

A clinical trial intelligence tool newly released by business software giant Oracle appears destined to boost R&D productivity and reshape interactions between industry sponsors and their multitude of partners aiding drug development.

For years now, the key players—pharmaceutical companies, clinical research organizations (CROs), and investigative sites—have navigated their way through clinical studies as special-interest camps with "misaligned objectives," says Nick Giannasi, vice president of life sciences product strategy in Oracle's health sciences division. Oracle Clinical Development Analytics, launched in November, is designed to clarify what the targets are and bring visibility to how well they're being met in real time.

The absence of mid-study intelligence has been making rapid, informed decision making difficult and negating many of the benefits of moving from paper to electronic data capture, says Giannasi. Information technology (IT) vendors and consultancies can craft toolkits for the job, but customized solutions are relatively expensive and inconvenient to maintain over the long haul. They don't even make a lot of sense, now that metrics for measuring performance are largely standardized. Out-of-the-box solutions may be the only plausible means to substantive progress in making clinical development programs more productive and efficient.

Oracle Clinical Development Analytics provides actionable, fact-based insights that are easy to understand and specific to the user's role within an organization—be it clinical data managers eying completed-versus-expected number of case report forms across a portfolio of studies or clinical monitors identifying sites where training for a particular protocol has yet to happen, says Giannasi. The first version has 45 embedded data management metrics. Sites and partners can be easily compared against a common set of performance yardsticks. The information can be used to take corrective action, minimizing cost and time losses, as well as establish preferred partnerships with top performers.

Data transparency enhances trust between and among sponsors, CROs, and sites while helping to eliminate "adversarial discussions," says Giannasi. That's a soft but important product benefit, given the often fractious relationship between the three parties. Project managers and clinical monitors empowered with information to do their jobs better should also be easier to retain.

The potential dividends for study monitors are particularly noteworthy. With Clinical Development Analytics, they can use key site performance metrics to help balance their workload and do more targeted traveling, says Giannasi. The application can be loaded on Apple iPhones, which outpace Blackberries in popularity among monitors. Plug-ins will be easy to create for other mobile devices moving forward, he adds, since the software is built on an open IT platform.

Similarly, minimal IT know-how is required for users to add company-specific performance metrics or customize and publish their own dashboards with the software's drag-and-drop interface, says Giannasi.

Oracle Clinical Development Analytics is currently being evaluated by several large sponsors and CROs, says Giannasi. Its major strength is the underlying brand. Oracle is a big player in the software space and the market leader, worldwide across industries, in the business intelligence niche. It is well positioned to access data from its other IT products and has more than 50,000 users of its clinical trial management system from which it can learn and mine strategic development partners. The demand for clinical development intelligence is also expected to escalate in response to growing interest in adaptive trial design, translational medicine, and use of electronic health records in longitudinal and comparative effectiveness studies.

MRI manufacturers using analytic software to help service engineers prioritize visits to imaging centers may be suggestive of the type of productivity gains the new Oracle solution could deliver. Using that as a benchmark, productivity gains among monitors using Clinical Development Analytics could be in the 10%-25% range with a positive ripple effect on site output, says Giannasi. Over the next six to 18 months, Oracle plans to build case studies demonstrating the product's contribution to reducing overall drug development cycle time.

Ultimately, the software may help lessen the number of monitors and sites needed per study, as well as reduce the need for the support of external patient recruitment companies, says Giannasi. "At the moment, enrollment is more art than science."