Harlan Krumholz

Straight talk with… Harlan Krumholz

Many big-name pharmaceutical companies have made strides toward data transparency in recent years, but perhaps none is as open as Johnson & Johnson (J&J) is now. On 29 January, the New Jersey–based drug giant announced that it will make all of its clinical trial data available through an academic clearinghouse for scientific information known as the Yale University Open Data Access (YODA) project. Under the agreement—the most comprehensive industry-academic partnership of its kind—YODA will independently review and make decisions about requests to access de-identified, patient-level data from the clinical studies supporting all of J&J's approved drugs.

It's a major step forward for YODA, which was born in 2011 out of an initial project with the global drug and device manufacturer Medtronic, which is headquartered in Minneapolis, to release data from studies of the company's Infuse product, which contains recombinant human bone morphogenetic protein-2, used to promote bone formation. At the helm of YODA is Harlan Krumholz, a physician-scientist who for almost two decades has led the Yale-New Haven Hospital Center for Outcomes Research and Evaluation, which aims to help compare the effectiveness of various healthcare interventions and, ultimately, improve clinical decision-making. Krumholz spoke with **Roxanne Khamsi** about how greater access to data is a boon for medicine.

When you started your medical career as a cardiologist, is it fair to say you didn't think you'd be preoccupied with questions about data algorithms and data sets?

That's true, but there's continuity with where I started. A great influence on me was Lee Goldman [now dean of the faculties of health sciences and medicine at Columbia University Medical Center in New York] because he was asking very practical and pragmatic questions, and he was systematically collecting data that was intended to inform critical decisions that people were making. At that time he was doing it with brute force. I was drawn to figuring out, gosh, how do we better inform these decisions which are being made often too casually and in a data-free environment?

What got you thinking about the need to promote data transparency through a model like YODA?

This first became clear to me in my experience in the Vioxx litigation. I became involved with the Vioxx litigation on behalf of plaintiffs when I realized that there was a lot of data that were out of public view, that only some of the studies that had been conducted on the drug had been published, and that the systematic reviews that had been conducted had been only done on a subset of all of the studies that were available. This just got us thinking about this more broadly.

Why would J&J choose to open its data vaults to you?

Through our discussions, J&J committed to full release of all their data and relinquished all control over decisions about that data. It's demonstrating their full commitment to release and not wanting to be in the middle of any disputes about data access. I think it's sending a message that they are enthusiastic supporters themselves.

How far back will the data access go?

Clearly, all the recent trials should be made available. The question is what if somebody asks for something that was done 20 years ago? This is why I think it's up to us, the independent academic partner, to try to understand what's the cost in preparing it. So if someone has a trivial research question but it would cost a million dollars to prepare the data, you know, is that sensible? We put together a process that both seeks to promote data transparency but also tries to be sensible.

Are you concerned about users re-identifying study participants?

With enough effort you can re-identify almost anything. It's going to be up to us to try to get certification that the database is de-identified. We haven't settled on the final process yet, but what we did in Medtronic was we asked them to take an online short course where we're emphasizing the principles of the data release. One of the principles is that they'll report their results, another principle is that they won't try to re-identify the data, another is that they won't use it for another purpose, another is that they won't give it to another person outside their research group. The whole philosophy is that you declare what you're doing.

Beyond scale, how is this partnership with J&J an evolution from what you'd coordinated with Medtronic?

Medtronic also gave us full jurisdiction over the data, so in that way that's the same. But Medtronic did a project with us in a situation where they were being criticized about a specific product, and the project was about one product, recombinant human bone morphogenetic protein-2. This is a situation where there is no crisis. It's a broad-based release.

Are you in talks with any other companies to expand YODA's role in these data-sharing endeavors?

We are hearing from some companies that there's an interest potentially having an independent academic partner. They're all very preliminary.

How many people are involved in these panels reviewing the data?

We haven't really ramped up the review yet. On the Medtronic project we had maybe six or eight people who were involved. We're going to have to see how we're going to have to scale it. We're probably going to be bringing in people like peer reviewers. So we'll be asking people to help us process these [requests for data]. The office will almost be run like a journal: you'll come in with a proposal, it will be reviewed, you'll get feedback. The only thing is that we expect the acceptance rate to be really, really high.