



Dan Munro. (<http://www.forbes.com/sites/danmunro/>) Contributor

I write about the intersection of healthcare innovation and policy.

FOLLOW

PHARMA & HEALTHCARE (/HEALTHCARE) 4/08/2014 @ 9:38PM | 3,630 views

Big Pharma Opens New Chapter On Big Data Collaboration

[Comment Now](#) [Follow Comments](#)

In the course of one short week, no less than 3 different models have emerged for sharing big data in the pharmaceutical industry.

The highest profile of these – called Project Data Sphere ([PDS here \(https://www.projectdatasphere.org/projectdatasphere/html/home.html\)](https://www.projectdatasphere.org/projectdatasphere/html/home.html)) – was announced earlier today with the official opening of an online resource to share clinical trial data for use in cancer research.

The number of available data sets available today is fairly small (9), but the list of companies committed to providing data is impressive and includes [AstraZeneca \(/companies/astrazeneca/\)](/companies/astrazeneca/) [AZN +1.51% \(/companies/astrazeneca/\)](/companies/astrazeneca/), [Bayer \(/companies/bayer/\)](/companies/bayer/), [Celgene \(/companies/celgene/\)](/companies/celgene/) [CELG +6.63% \(/companies/celgene/\)](/companies/celgene/), Janssen Research and Development (an affiliate of [Johnson & Johnson \(/companies/johnson-johnson/\)](/companies/johnson-johnson/) [JNJ +0.91% \(/companies/johnson-johnson/\)](/companies/johnson-johnson/)), [Pfizer \(/companies/pfizer/\)](/companies/pfizer/) [PFE +1.17% \(/companies/pfizer/\)](/companies/pfizer/), Memorial Sloan Kettering Cancer Center and [Sanofi \(/companies/sanofi/\)](/companies/sanofi/) [SNY +0.95% \(/companies/sanofi/\)](/companies/sanofi/) U.S.

The primary objective, of course, is to accelerate drug discovery and development. From the website, the underlying objective is posed as a question this way: “What if we could share, integrate and analyze our collective historical cancer research data in a single location?” (<https://twitter.com/intent/tweet?>

url=http%3A%2F%2Fonforb.es%2F1oKReEP&text=%22What%20if%20we%20could%20share%2C%20integrate%20ar

Formal registration for professional researchers will be required, but there will be no charge for accessing the research, uploading data sets and access will be available online globally.

Using clinical trial datasets collaboratively is a big leap forward in the cancer drug discovery process. 8.2 million people still die of cancer every year while the attrition rate for clinical testing of promising compounds can be as high as 95%. This could become substantially lower once researchers in both academia and industry share clinical trial data. We're excited to be working with world class partners like SAS, SAGE Bionetworks, academia, many in industry and importantly patient groups to bring this free resource to researchers globally. **Charles Hugh-Jones** – Chief Medical Officer for North America, Sanofi

Also in development are data sets to be provided to PDS by the Alliance for Clinical Trials in Oncology (sponsored by the National Cancer Institute), [Amgen \(/companies/amgen/\)](#) AMGN +1.7% (/companies/amgen/) and [Quintiles \(/companies/quintiles/\)](#).

The technology platform itself was built by global analytics powerhouse [SAS \(/companies/sas/\)](#) who will continue to host the online service, provide analytics software and technical domain expertise.

SAS has been involved with the [CEO Roundtable on Cancer \(http://www.ceoroundtableoncancer.org/\)](#) for some time. As an employer we have a unique opportunity to positively impact the health and well-being of our employees and their families. By donating our analytics capabilities and expertise to the Project Data Sphere initiative, we're also taking concrete actions to aid in research and the fight against cancer. **Jim Goodnight** – Chief Executive Officer, SAS

Of the three different models, PDS represents a kind of “open source” model with significant industry support and (potentially) rich historical data.

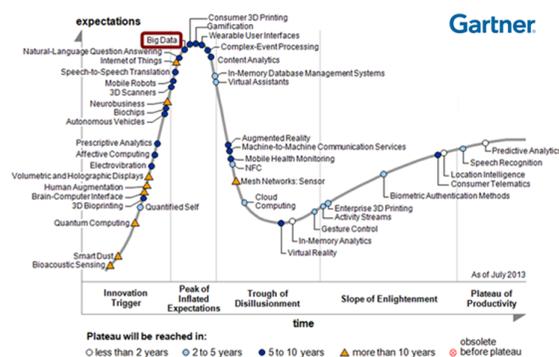
The second model was voted on last week by the European Parliament in the form of legislation that is likely to take effect by 2016 ([here \(http://www.alltrials.net/2014/europe-votes-for-clinical-trial-transparency/\)](#)). The overwhelmingly positive vote (547 to 17) is also aimed at clinical trial data transparency, but only targets new trials after the law takes effect. Assuming it's enacted, the benefits of a legislative model will be some strict conditions.

- Require that all drug trials in Europe are registered before they begin on the publicly accessible EU clinical trials register.
- Require that a summary of the results from these trials is published on the register within a year of the trial's end.
- Require that a summary understandable to a lay person of what was found in the trial is published on the register.
- Establish a new publicly accessible EU clinical trials register, to be set up and run by the European Medicines Agency.
- Impose financial penalties on anyone running a clinical trial who does not adhere to these new laws.

The third approach – announced just yesterday – is the 5-year commercial agreement between Genentech and PatientsLikeMe ([here](http://www.bizjournals.com/sanfrancisco/blog/biotech/2014/04/genentech-patientslikeme-cancer-social-media.html?utm_content=buffer9ba09&utm_medium=social&utm_source=twitter.com&utm_campaign=buffer) (http://www.bizjournals.com/sanfrancisco/blog/biotech/2014/04/genentech-patientslikeme-cancer-social-media.html?utm_content=buffer9ba09&utm_medium=social&utm_source=twitter.com&utm_campaign=buffer)) which is also targeting cancer research as the lead effort. The data in this case isn't clinical trial data, but data provided by actual patients in the course of their current treatment for a wide range of conditions – including cancer.

'It's not just about big data it's about the right data. In the end, everything comes down to whether or not it makes a difference in patients' lives in ways that they care about. Who better to provide that than the patient?' **Jamie Heywood** – Co-Founder and Chairman of PatientsLikeMe

All of which highlights the early nature of big pharma collaboration around big data. Some of this is also clear evidence of Gartner's assessment of Big Data generally which they view as being near the “peak of inflated expectations” on their venerable “Hype Cycle for Emerging Technologies (2013).”



The flurry of data collaboration has – at least in some cases – taken years to develop, but it also speaks volumes about the need to greatly accelerate drug discovery and development. In two of the cases (PDS and Genentech), it's really borne out of economic necessity to change the timeline around the huge expense evident in that lengthy process.

The larger question then seems to be whether the value of a voluntary “open-source” model will trump the regulated or commercial variants. In all cases, however, the clear intent is for patients to benefit directly from accelerated drug development and the use of their de-identified data for broader research purposes. Clinical trial data that's captured and used once is time consuming, expensive and often at risk of being a duplication.

As Gartner suggests, big data may well be at its peak of inflated expectation, but what we're also seeing through other examples (like [Google's Flu Trend here](http://www.theatlantic.com/technology/archive/2014/03/in-defense-of-google-flu-trends/359688/) (<http://www.theatlantic.com/technology/archive/2014/03/in-defense-of-google-flu-trends/359688/>)) is that multiple data sources used collaboratively can produce richer results. Ultimately, the big promise of big data is much more signal – and a lot less noise. Applying a collaborative approach is always an ambitious and risky effort, but it's a horse race in which we all have a stake and it's one with enormous life saving potential.

[Comment Now](#) [Follow Comments](#)

