ORACLE
Health Sciences

UNIFYING ACTION, ACCELERATING RESULTS.

BUSINESS BRIEF

Interpreting Big, Real World Data – The New Clinical Data Scientist Role

With cloud technologies becoming commonplace to store and manage big, real world amounts of clinical (genetic, BP, temp, etc.) and medical (EMRs, EHRs, outcomes, etc.) data; and, the growing popularity of wearable sensor devices to collect and transmit clinical trial patient data from remote locations on a continuous basis, the research world is brimming in terabytes of information.*

But what does one do with all this data? How can we sort through it to find those points that provide additional support for what is known about a drug's effect on a disease? Better still, how can it be optimized to demonstrate breakthrough insights and new patterns in relation to the drug and the disease?

These questions pave the way for the introduction of a new research discipline—data science. In her paper, Michaela Jahn, Global Clinical Data Manager at F. Hoffmann-La Roche, defined it as the application of a team's diverse informatics and analytical capabilities to retrieve and analyze data to support drug project decision making, drug development, and platform development. Data science capabilities include: bioinformatics, imaging informatics, biostatistics, data integration and visualization, text-mining, and information science.

The Clinical Data Scientist

This all sets the stage for the role of the clinical data scientist. Driven by the changes in data management, this role can be seen as an evolutionary step forward for the clinical data manager. Where the data manager was task driven and in reactive mode—organizing and readying data for analysis—the data scientist takes a more active discovery role looking for new anomalies and patterns in the clinical trial data that may suggest additional paths to insight. Additionally, the skills of the data scientist could prove critical for the clinical trial team. His/her discerning data discovery capabilities could not only help the team to identify new data paths to explore, but also to guide them away from unproductive, cost- and time-wasting data directions.

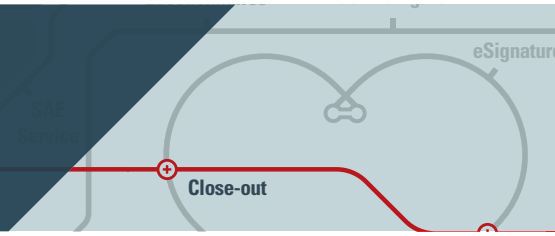
SRINIVAS KARRI
DIRECTOR, PRODUCT STRATEGY
ORACLE HEALTH SCIENCES

Srinivas Karri leads Oracle Health Sciences' Data Management capabilities, and brings years of experience consulting on and managing large data volumes in complex clinical trials

CID/2018/05.24

ORACLE

Oracle Health Sciences breaks down barriers and opens new pathways to unify people and processes, helping to bring new drugs to market faster.



In her paper, Ms. Jahn defines the clinical data scientist as one who would require: comprehensive knowledge of all areas of data management and data delivery, an understanding of protocols, the ability to interpret clinical study data, and knowledge of technologies needed throughout the clinical study, from start-up to close-out. She also recommends that the data scientist be viewed as an equal partner in the study team and have the following responsibilities and understanding:

- Have a clear understanding of protocols, their structure, primary and secondary endpoints (the accurate collection and extraction of data).
- Own oversight of study milestones and what is needed for data delivery.
- Conduct a data risk assessment (what data should be cleaned for a certain therapeutic area and what data can be left as is).
- Understand the basics of statistics and programming.
- Support international standards.
- Understand the basics of the disease area.
- Oversee the external service providers.
- Adapt to new technologies.
- Help clinical scientist understand the data modeling and explore the data.

The paper goes on to identify how the clinical data manager can become a clinical data scientist by evolving or developing skills in these areas.

Glassdoor lists the data scientist as one of the 25 best jobs in America in 2016. Though, a recent article CIO.com observes that currently, not only is there a lack of qualified talent for this emerging role, but also that companies hiring data scientists are still grappling with the most effective ways to utilize their skills. The amount of data collected in clinical trials is only growing, which increases demand for the skills of the data scientist. In clinical trials, it will fall to the clinical data scientist to see new patterns and find new relationships in this data that eventually can save more lives and achieve better patient outcomes.

Finally, it will be interesting to combine the clinical data manager's deep experience, as he/she evolves into a data scientist, and the new, creative – though less experienced – perspectives of millennials coming into the clinical data science role as natives. Where experienced data management professionals have deeper insight into data exploration, millennial data scientists might provide fresh, unexpected attitudes on new data sources and combinations. Hopefully, together, these two groups can optimize clinical R&D data discovery in ways only imagined today.

Data

science capabilities: bioinformatics, imaging informatics, biostatistics, data integration and visualization, text-mining, and information science

Where experienced data professionals have deeper insight into data exploration, millennial scientists might provide fresh, unexpected attitudes on new data sources and combinations.

*The human genome typically includes a few gigabytes per person. Also, simply taking blood pressure (BP) three times/day in a two-year, 500-subject, clinical trial results in two million data points. BP is two data points (Systolic, Diastolic); therefore, 2 x 3 per day x 365 days x 2 years x 500 patients = about 2 million.

ORACLE