

Global Solutions from the World's Largest Data-Focused CRO

Big Data : Hurdle or Solution?

Arnab Goswami and Sampada Angane

Disclaimer



This presentation does not represent views of Quanticate International and belong solely to the author. Authors take full responsibility for the work and errors if any.

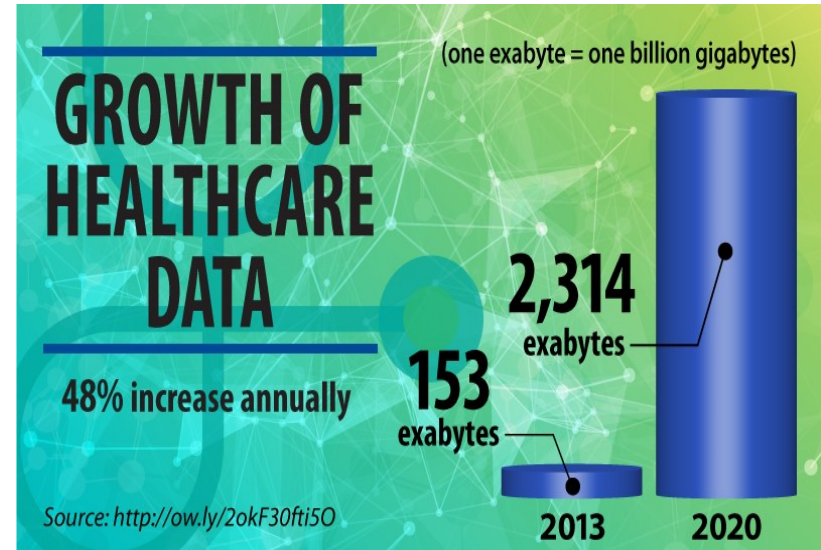
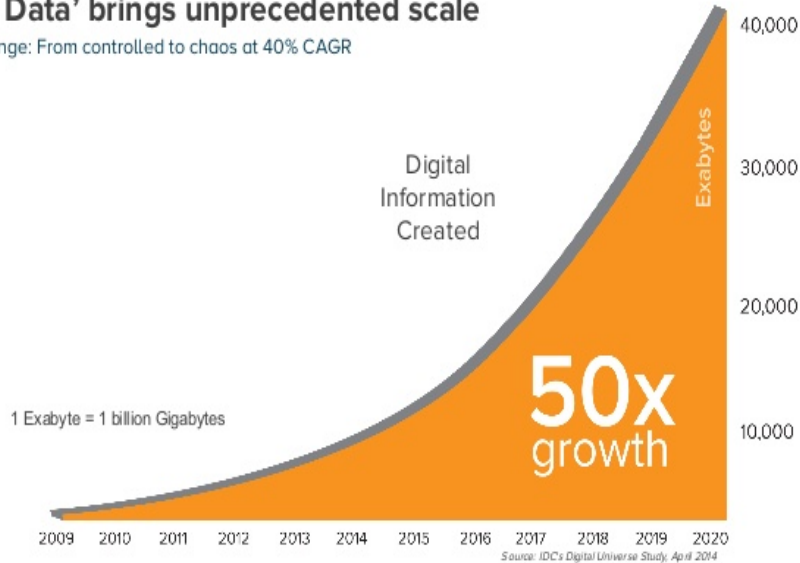
Data Volume



Big Data: The 'BIG' explosion

'Big Data' brings unprecedented scale

Challenge: From controlled to chaos at 40% CAGR



Big Data in Clinical Trials

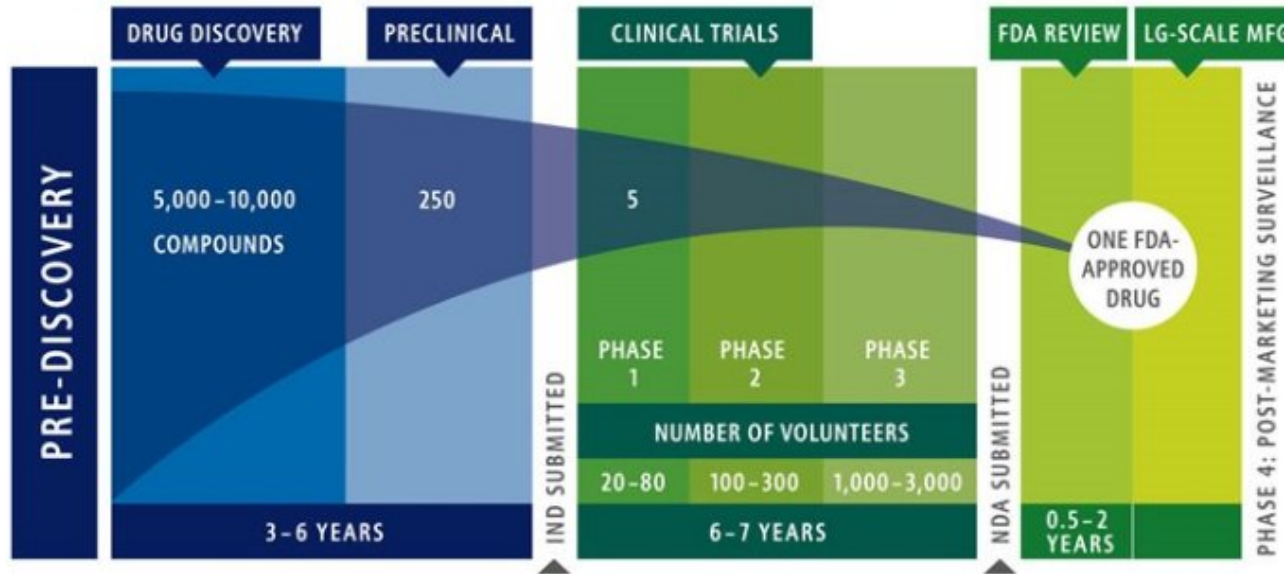
- ☐ New architecture
- ☐ Techniques
- ☐ Algorithms
- ☐ Analytics

- Patient recruitment
- Identify new, targeted therapies
- Identify adverse event responses



Bench to Market

Drug Discovery and Development: A LONG, RISKY ROAD



Source: Pharmaceutical Research and Manufacturers of America

Sources of Big Data in healthcare

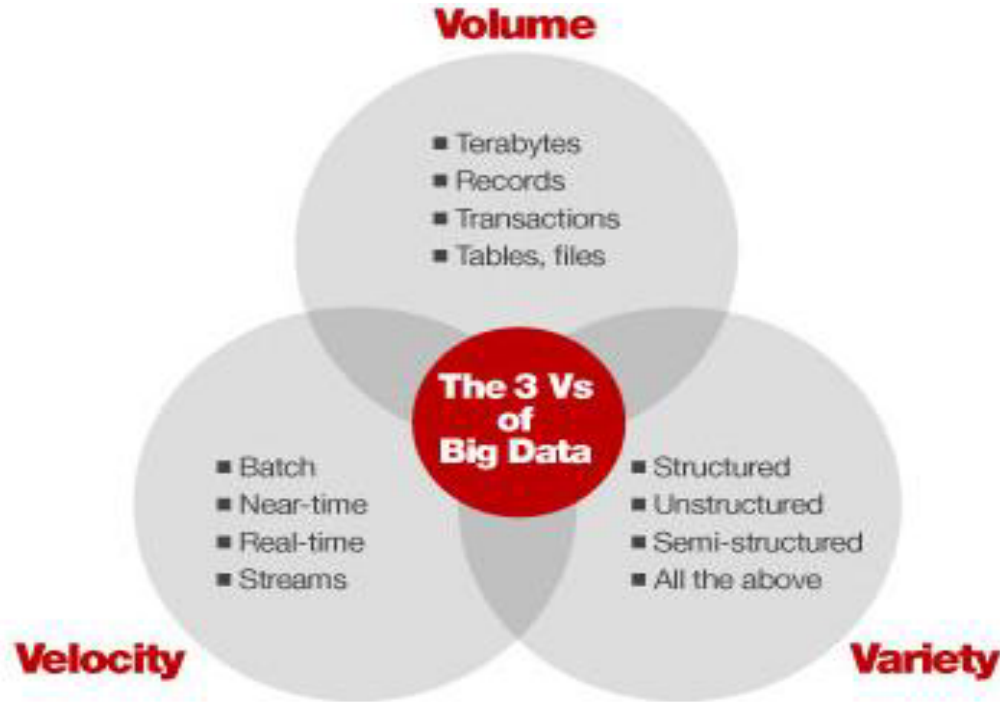


Quanticate

The Clinical Data Experts



Main characteristics of Big Data



Solutions over conventional process

RCT (Randomized Clinical Trial) vs BDT (Big Data Trial)

Randomized Clinical Trial

- Conventional approach
- Interventional and observational trials
- Performed in experimental conditions
- Biological efficacy may not translate to clinical efficacy

Big Data Trial

- Real World Data
- Data collected from EMRs
- Directly reflects clinical effectiveness

AI and Machine Learning

ML is a subset of AI in which computers learn without being explicitly programmed. It relies on neural networks.

- ☐ Site management
- ☐ Patient adherence
- ☐ Remote monitoring
- ☐ Reduce data errors
- ☐ Finding best sample sizes



** Logos of Spark, Hadoop, Ceph, Google and DISCO has been used for representative purpose only*

Predictive Modeling

Conventional patient recruitment

- ☐ Time consuming
- ☐ Resource-intensive
- ☐ Lengthy procedure
- ☐ Many suitable candidates may be missed

Automatic patient recruitment

- ☐ Efficient and effective
- ☐ Takes less time
- ☐ Frees up resources
- ☐ Rigorous in matching patients to protocols



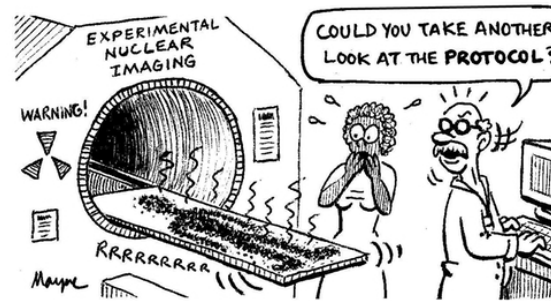
Electronic Health Records (EHR)

- ☐ Demographics, medical history, allergies, laboratory test results.
- ☐ Smart List and Smart Phrase objects.
- ☐ Standardizations
- ☐ Exchange of data
- ☐ Error Reduction
- ☐ Improved quality
- ☐ Observer bias is reduced
- ☐ No data replication



Identify adverse event responses

- ☐ [ClinicalTrials.gov](https://clinicaltrials.gov)
- ☐ Drug-AEs database
- ☐ Social media



Challenges for Big data

- Security
 - ☐ Access
 - ☐ Privacy
 - ☐ Share-ability



Is the lock full proof?

Challenges for Big data (Contd.)

- ❑ Data Classification
- ❑ Data Modelling
- ❑ Cloud Storage
- ❑ Data Accommodation
- ❑ Data personnel
- ❑ Communication
- ❑ Data Nature
- ❑ Technology Incorporation



Challenges for Big data (Contd.)

- ☐ Access
- ☐ Privacy
- ☐ Share-ability
- ☐ Proprietary rights
- ☐ Translation
- ☐ Incentive



Data Security

- ❑ Big data contained subject's personal information and their health history
- ❑ It is important for the database to be protected from hacking, cyber theft and phishing, where the stolen data can be sold for a huge sum.
- ❑ More regional and global makes the job more complicated and have more serious impact on security, standards, language and terminology.
- ❑ The accessibility of the healthcare data need to be consistently reviewed and monitored.



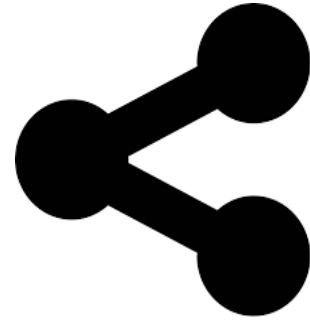
Privacy

- ❑ Access to health care data is plagued by vulnerability due to patient privacy considerations which are protected by federal and local laws of protected health information such as Health Insurance Portability and Accountability Act of 1996 (HIPAA).
- ❑ The fear of litigation and breach of privacy discourages providers from sharing patient health data, even when they are de-identified.



Share-ability

- ❑ De-identification – a complicated process.
- ❑ Informatics for Integrating Biology and the Bedside (i2b2).
- ❑ Coding and the flow of the analysis.



Proprietary rights

- ❑ Ownership of patient data among various stakeholders in the healthcare system including providers, patients, insurance companies and software vendors.
- ❑ In general, the current model is such that the patient owns his/her data, and the provider stores the data with proprietary software systems



Data Classification

- ❑ Big data is massive, less structured and heterogeneous, so there is a need to identify and classify the data so that it can be used effectively.
- ❑ The big data also required to be contextualized or pooled together so that it will become more relevant to specific individuals or groups



Data Modelling

- ☐ Although big data is excellent for modelling and simulation, there is a need to identify, structure and pool the proper relevant data so that it can be used to model the problems, which later can be used for clinical intervention.
- ☐ Without the proper structured data, it is challenging to analyse and visualize the output and to extract specific information or data.

Data Accommodation

- ❑ One robust big data system is required to accommodate all the data and it has to be compatible and simplified.
- ❑ This is to ensure that the users are able to retrieve the information without any hassle.



Alternative for health practitioners?

- ❑ Since big data has the ability to predict future medical issues, big data can pose risk and undermine doctors.
- ❑ The patients too will rely on the technology rather consulting the healthcare practitioners.



vs.



Summary

- ❑ First step is to **introduce standardization** of data and processes.
- ❑ AI is not just for **data analysis**, but also for **data pre-processing/aggregation**.
- ❑ Success will require **process transformation** to reduce variation.
- ❑ **Confidentiality** or competitor-sensitive concerns can be addressed using a **de-centralized** aggregation and analysis approach.
- ❑ **Expect more** analytical functionality and capability **from vendor solutions**.
- ❑ Focus on the **enablers, algorithms and visualization** to optimize data value.
- ❑ **Incentivisation** between AI and Pharma community needs to find common ground to **accelerate** collaboration.

Bibliography

- ☐ <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5583160/>
- ☐ <http://www.worldpharmatoday.com/Articles/overcoming-the-challenges-of-clinical-trial-recruitment-using-big-data-technologies>
- ☐ <https://www.techemergence.com/machine-learning-in-pharma-medicine/>
- ☐ The power of petabytes-- *NATURE* / VOL 527 / 5 NOVEMBER 2015
- ☐ Big data in healthcare—the promises, challenges and opportunities from a research perspective: A case study with a model database--Adibuzzaman *et al.*
- ☐ Big data and clinical research: perspective from a clinician. Zhongheng Zhang *J Thorac Dis* 2014;6(12):1659-1664
- ☐ Big data in designing Clinical trials: opportunities and Challenges Charles S. Mayo *et al.* *Frontiers in Oncology. August 2017. Volume 7*
- ☐ The Usefulness and Challenges of Big Data in Healthcare. *Journal of Healthcare Communications* 2018. Vol.3 No.2:21



Arnab Goswami

E-mail: Arnab.Goswami@qunaticate.com

Sampada Angane

E-mail: Sampada.Angane@quanticate.com