



# BUILDING BRIDGES

## The Paradox of Big Data in Pharma

Aruna Chakraborty



# Agenda

- The Story so far
- Entry of AI, a Panacea?
- Changing the clinical trial paradigm
- Exciting prospects
- I rest my case
- Share your thoughts...



The Story so far...



## The state of clinical trials

UNDERSTANDING A PATIENT'S JOURNEY FROM DIAGNOSIS TO ENROLLMENT TO MONITORING

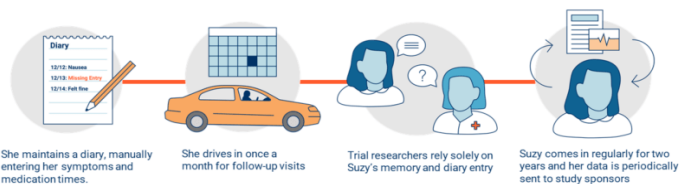
### Finding a Clinical Trial



### Enrollment



### Adherence Monitoring



‘Eroom’s Law’, which is the reverse of the ‘Moore’s Law’ concept in computing. It states that the number of drugs being approved decreases year on year while the costs increase.

The law is attributed to four main causes.

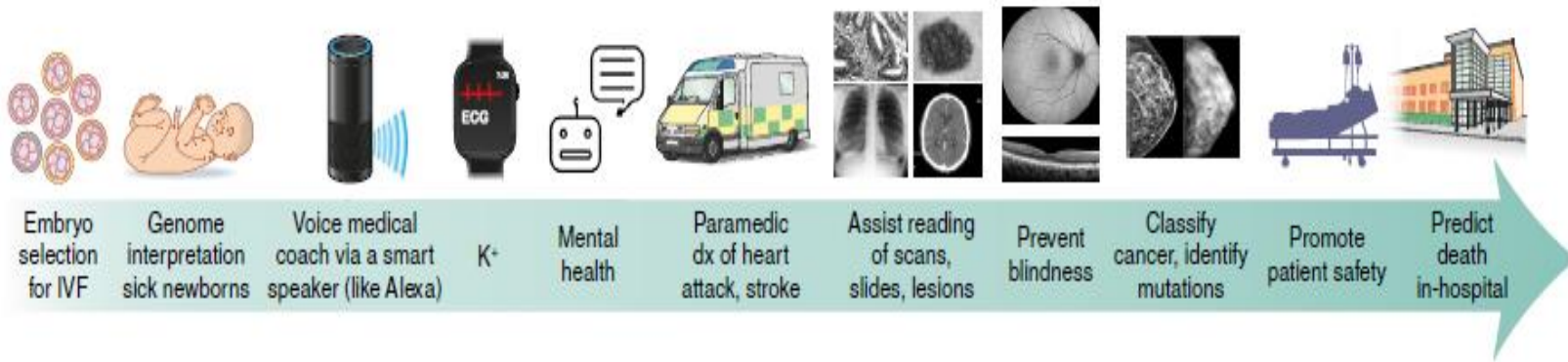
- the 'better than the Beatles' problem
- the 'cautious regulator' problem
- the 'throw money at it' tendency
- the 'basic research—brute force' bias

Medicine is at the crossroad of two major trends. This exemplifies a paradox

1. failed business model, with increasing expenditures and jobs allocated to healthcare, but with deteriorating key outcomes, including reduced life expectancy and high infant, childhood, and maternal mortality.
2. generation of data in massive quantities, from sources such as high-resolution medical imaging, biosensors with continuous output of physiologic metrics, genome sequencing, and electronic medical records.



Entry of AI, a Panacea?



Machine learning  
 Deep learning  
 Natural language processing  
 Robotics and the internet of things

- In AI - “Garbage in, means garbage out”
- Algorithms in dermatology that diagnose melanoma but lack inclusion of skin color and the use of the corpus of genomic data, which so far has seriously underrepresented minorities.
- Access limitations characterize not only private companies but some government-sponsored Big Data initiatives (ex. Genomics England, Danish Biobank, etc.).

## The current limitations:

- Accuracy
- Bias
- Privacy and security
- Distrust of black boxes
- Poor data hygiene
- Not everyone is a data scientist
- Doubled-edged sword of regulatory culture
- AI Training Complications



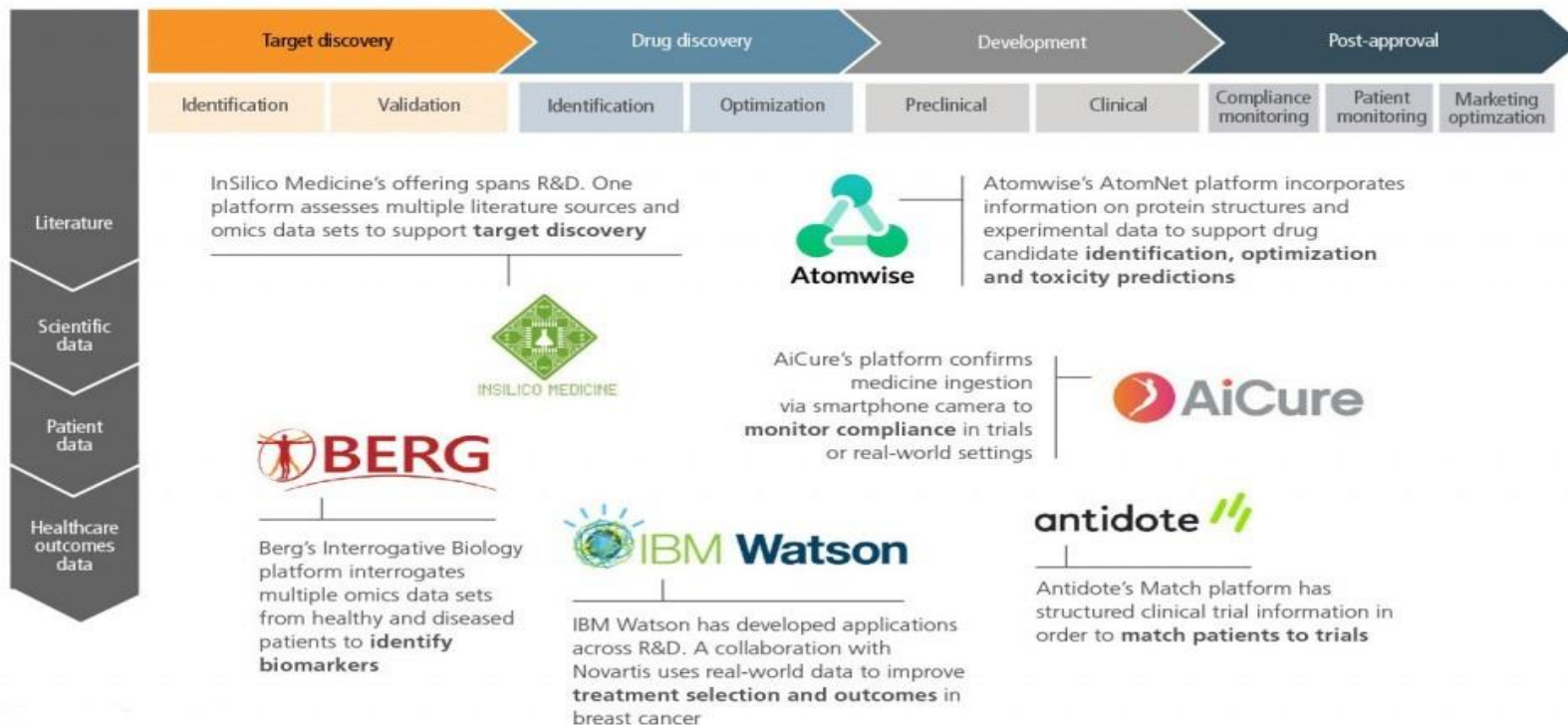
# Changing the Clinical Trial Paradigm.

Use of AI, and the deep-learning subtype in particular, has been enabled by the use of labeled big data. In medicine, impact is at three levels:

- for clinicians, predominantly via rapid, accurate image interpretation
- for health systems, by improving workflow and the potential for reducing medical errors
- for patients, by enabling them to process their own data to promote health.

- In silico drug design
- Selecting precision treatments for individual patients
- Predict study risks and their drivers to enable preventive maintenance and remediation of the drugs
- Medication adherence
- Google's use of AI in pathology has helped with detecting cancer

## AI companies active across the drug lifecycle



| Company          | FDA Approval   | Indication                                    |
|------------------|----------------|-----------------------------------------------|
| Apple            | September 2018 | Atrial fibrillation detection                 |
| Aidoc            | August 2018    | CT brain bleed diagnosis                      |
| iCAD             | August 2018    | Breast density via mammography                |
| Zebra Medical    | July 2018      | Coronary calcium scoring                      |
| Bay Labs         | June 2018      | Echocardiogram EF determination               |
| Neural Analytics | May 2018       | Device for paramedic stroke diagnosis         |
| IDx              | April 2018     | Diabetic retinopathy diagnosis                |
| Icometrix        | April 2018     | MRI brain interpretation                      |
| Imagen           | March 2018     | X-ray wrist fracture diagnosis                |
| Viz.ai           | February 2018  | CT stroke diagnosis                           |
| Arterys          | February 2018  | Liver and lung cancer (MRI, CT) diagnosis     |
| MaxQ-AI          | January 2018   | CT brain bleed diagnosis                      |
| Alivecor         | November 2017  | Atrial fibrillation detection via Apple Watch |
| Arterys          | January 2017   | MRI heart interpretation                      |

- Citizen Scientists - A pioneer in this field has been “PatientsLikeMe”. By allowing patients to publicly track health indicators, data can be aggregated and analyzed by individual patients in an effort to determine their best course of treatment.
- FDA Talks collaboration - Public-Private consortia have arisen to address key non-competitive challenges -- such as the FDA Sentinel Initiative and IMEDS to make progress on pharmacovigilance methodology.



Exciting prospects!



**Your smartphone will see you now**



**Smarter  
Primary Care.**  
Made for people like me.



Has

Perform regular self-checks for skin cancer with your phone.

[CHECK YOUR SKIN NOW](#)

Find out what  
could be wrong

Ada is supported by a sophisticated medical knowledge base, covering thousands of symptoms and conditions. The [Medical Library](#) shares patient-friendly medical information created to help you better understand and manage your health.

WebMD

WebMD, LLC Health & Fitness

PEGI 3

WebMD®

It may be caused by an injury, inflammation or sinusitis (hollow spaces in bones of the face around the nose). It is a very common condition. It may follow a cold or flu. The most common symptoms...

[Read More](#)

Likelihood

9 out of 10 people with these symptoms had this condition.



**Dr.  
Now**

- In 2018 Apple received FDA approval for their algorithm used with the Apple Watch Series 4
- MIT's Clinical Machine Learning Group is using machine learning to understand diseases such as type-2 diabetes, and design effective treatments in the process.
- Blockchain and AI in Pharma

- End of human experimentation is near
- Nano robots in blood
- DIY Biotechnology
- Gamify health



I rest my case...

The scientific theorist is not to be envied. For Nature, or more precisely experiment, is an inexorable and not very friendly judge of his work. It never says "Yes" to a theory. In the most favorable cases it says "Maybe," and in the great majority of cases simply "No." If an experiment agrees with a theory it means for the latter "Maybe," and if it does not agree it means "No."

– Albert Einstein

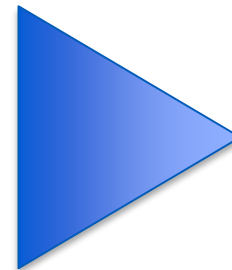
| Mechanistic Biomarker Discovery                                                                                                     | Empirical Biomarker Discovery                                                                                                                                                                                                            |
|-------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <ul style="list-style-type: none"> <li>- Uses biological knowledge at hand, tumor cells shed ctDNA</li> </ul>                       | <ul style="list-style-type: none"> <li>- Don't presume deep biological understanding, and hope to learn through a broad-based discovery process</li> </ul>                                                                               |
| <ul style="list-style-type: none"> <li>- Need to draw an impractical amount of blood to get just a single ctDNA molecule</li> </ul> | <ul style="list-style-type: none"> <li>- Protein profiles should be different in tumor/non-tumor conditions</li> </ul>                                                                                                                   |
| <ul style="list-style-type: none"> <li>- Many misleading driver mutations besides the one targeted</li> </ul>                       | <ul style="list-style-type: none"> <li>- Select a high-content assay (protein array, mass spec, etc.)</li> <li>- Collect "a lot" of samples</li> <li>- Something – lately, machine learning (ML)</li> <li>- Voila- biomarker!</li> </ul> |
| <ul style="list-style-type: none"> <li>- Loss of specificity</li> <li>- Loss of sensitivity</li> </ul>                              | <ul style="list-style-type: none"> <li>- No guarantee of reliability.</li> </ul>                                                                                                                                                         |

# Grounded Data Scientist

- “Instead of designing a (ML-driven) discovery project to *work*, you design it *not to fail* – specifically design around all the ways that it might cheat.”
- “statistical methods will always take the easiest way out to the answer you ‘want’: right for the wrong reasons looks right in your small data set, even if it stops working in a larger set.”
- Mechanism “allows us to define negative and positive controls which provide *invariants* you can enforce on or teach to a model.”



# DATA DREDGING



# Your Thoughts ?

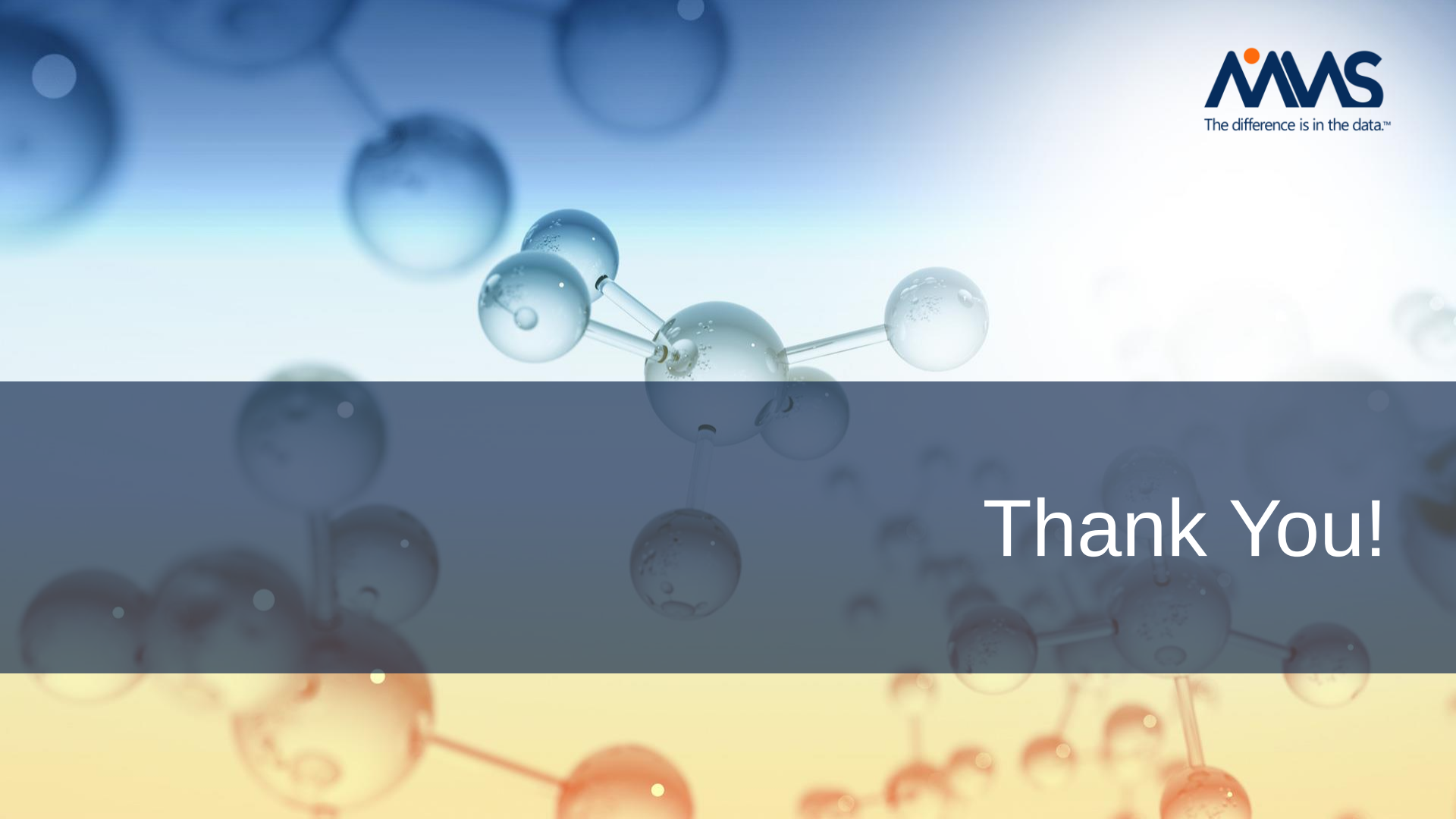




# References

- <https://www.cbinsights.com/research/clinical-trials-ai-tech-disruption/#magic>
- [https://en.wikipedia.org/wiki/Eroom%27s\\_law](https://en.wikipedia.org/wiki/Eroom%27s_law)
- <https://www.pharmaceutical-technology.com/comment/can-artificial-intelligence-help-pharma/>
- <https://www.lek.com/insights/artificial-intelligence-life-sciences-formula-pharma-success-across-drug-lifecycle>
- <https://becominghuman.ai/artificial-intelligence-in-pharma-4608b076a503>
- <https://www.statnews.com/2018/12/06/artificial-intelligence-pharma-health-care/>
- <https://www.gwern.net/docs/ai/2019-topol.pdf>
- <https://pharma.elsevier.com/pharma-rd/ai-life-sciences-good-bad-ugly/>
- <https://ai.googleblog.com/2017/03/assisting-pathologists-in-detecting.html>
- <https://www.decisionskills.com/blog/how-ice-cream-kills-understanding-cause-and-effect>
- <http://www.bbc.com/future/story/20161110-the-real-risks-of-artificial-intelligence>

- <https://www.healthaffairs.org/doi/10.1377/hblog20140710.040035/full/>
- <https://www.fda.gov/Safety/FDASSentinelInitiative/default.htm>
- <http://reaganudall.org/innovation-medical-evidence-development-and-surveillance>
- <https://www.patientslikeme.com/?format=html>
- <http://davidgratzer.com/reading-of-the-week/reading-of-the-week-your-smartphone-will-see-you-now-torous-on-digital-psychiatry-also-the-costs-of-homelessness/>
- <https://www.siliconrepublic.com/machines/ai-pharma-potential>
- <https://www.iotcoresoft.com/iot-knowlegde-center/blockchain-and-ai-impact-on-pharmaceutical-companies>
- <https://medicalfuturist.com/disruptive-technology-pharma>
- <https://www.forbes.com/sites/davidshaywitz/2019/01/03/data-scientist-explains-how-ais-seductive-power-can-mislead-biomarker-researchers/#18e5434e5959>
- <https://www.decisionskills.com/blog/how-ice-cream-kills-understanding-cause-and-effect>



Thank You!