

Single Day Event

Leveraging Big Data for Efficient Clinical Trials

Prerana Jain,
Alcon



Disclaimer

The opinions expressed in this presentation and on the following slides are solely those of the presenter and not necessarily those of *Alcon*.

Alcon does not guarantee the accuracy or reliability of the information provided herein.

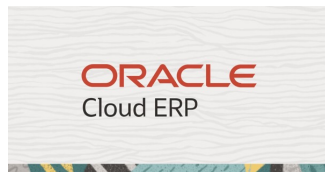


Introduction

What is Big Data?

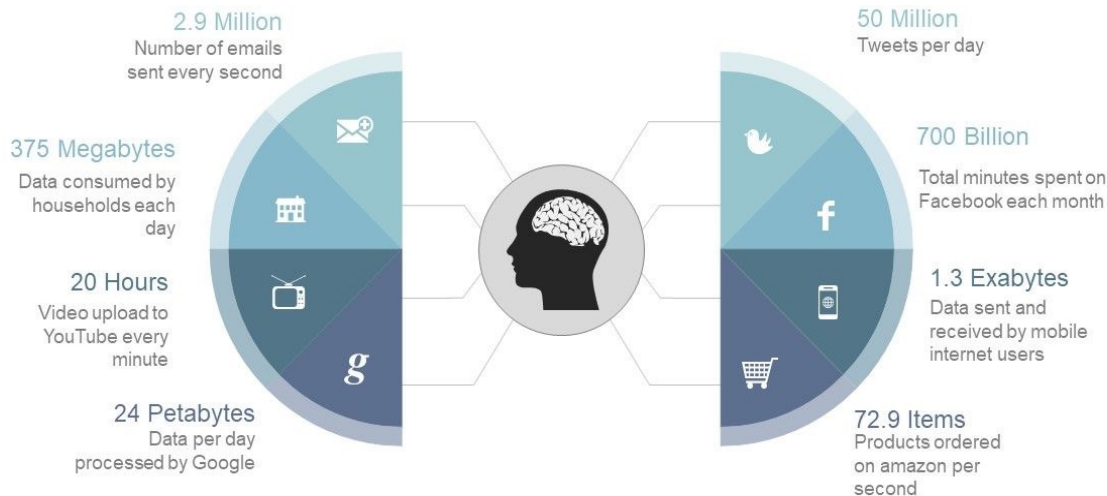


Where is it stored?





How Big Is Big Data?



Source: www.company.com

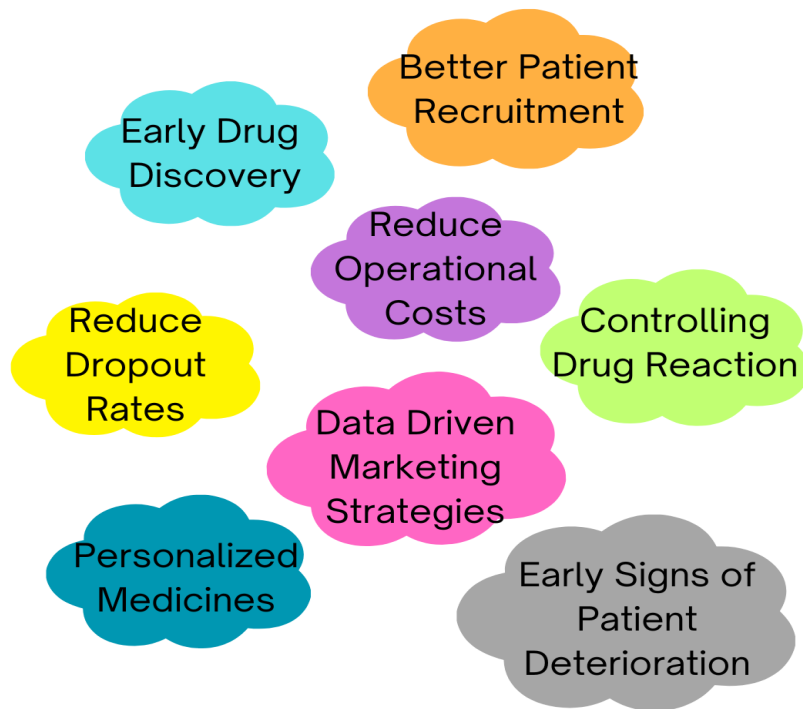


Big Data in Clinical Trials

Sources:

- Biomarkers
- Genomic Sequencing
- Electronic Health Record
- Claims data
- Patient registries
- Payer records
- Pharmaceutical Research
- Wearables
- Medical Devices
- Clinical trial data

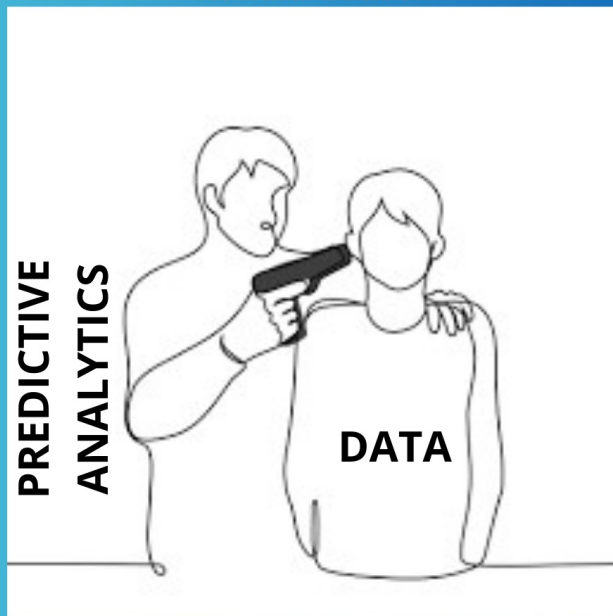
Uses:



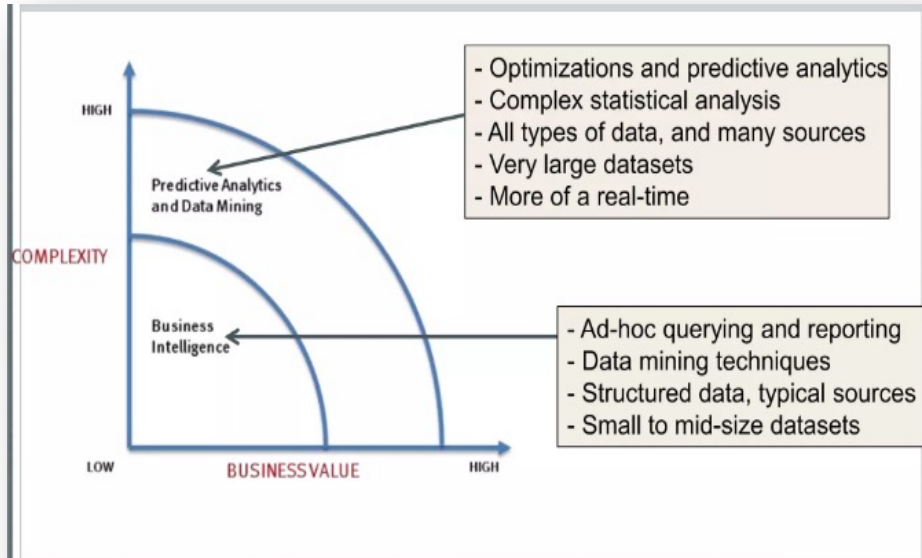
How can big data help us?

If you torture the data long enough, it will confess

- Ronald Coase



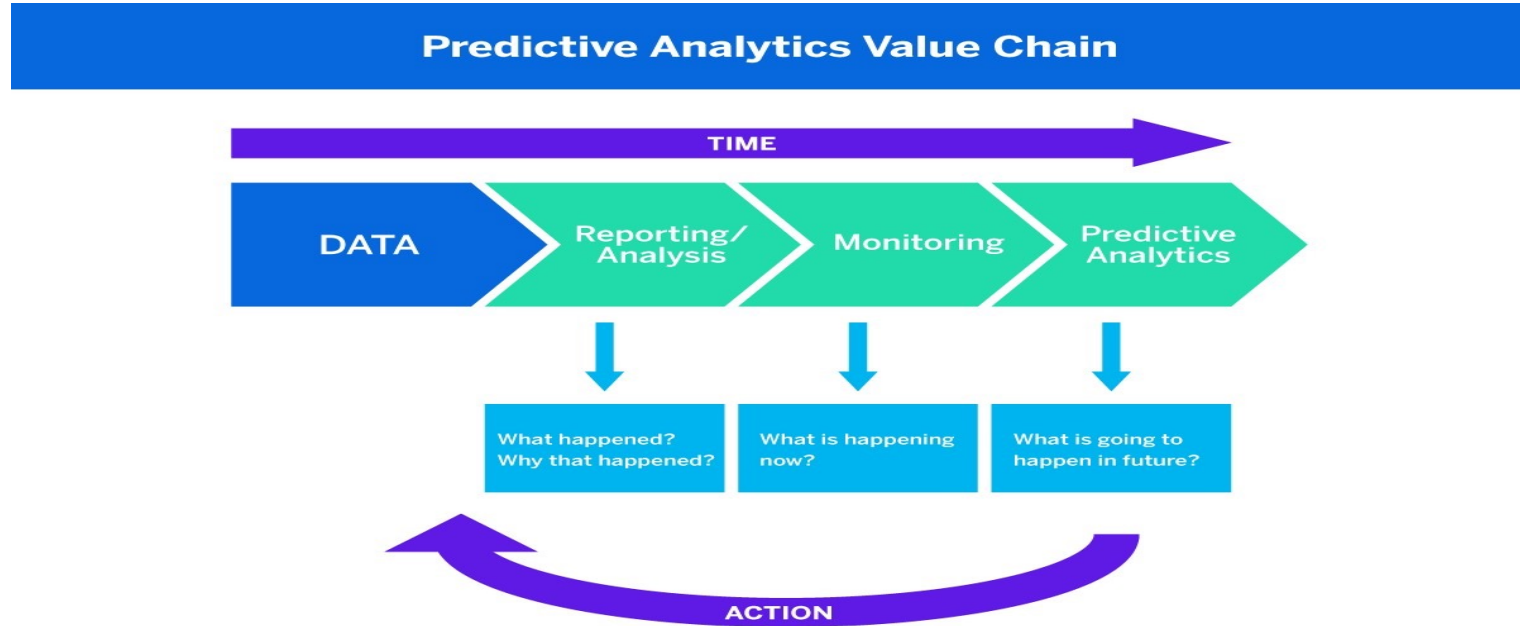
What's driving Big Data?



- Big data is a pool of information which is vast and complex
- Big data analytics helps organizations harness their data and use it to identify new opportunities, thus building Business Intelligence
- Predictive analytics models are designed to assess historical data, discover patterns, observe trends, and use that information to predict future trends



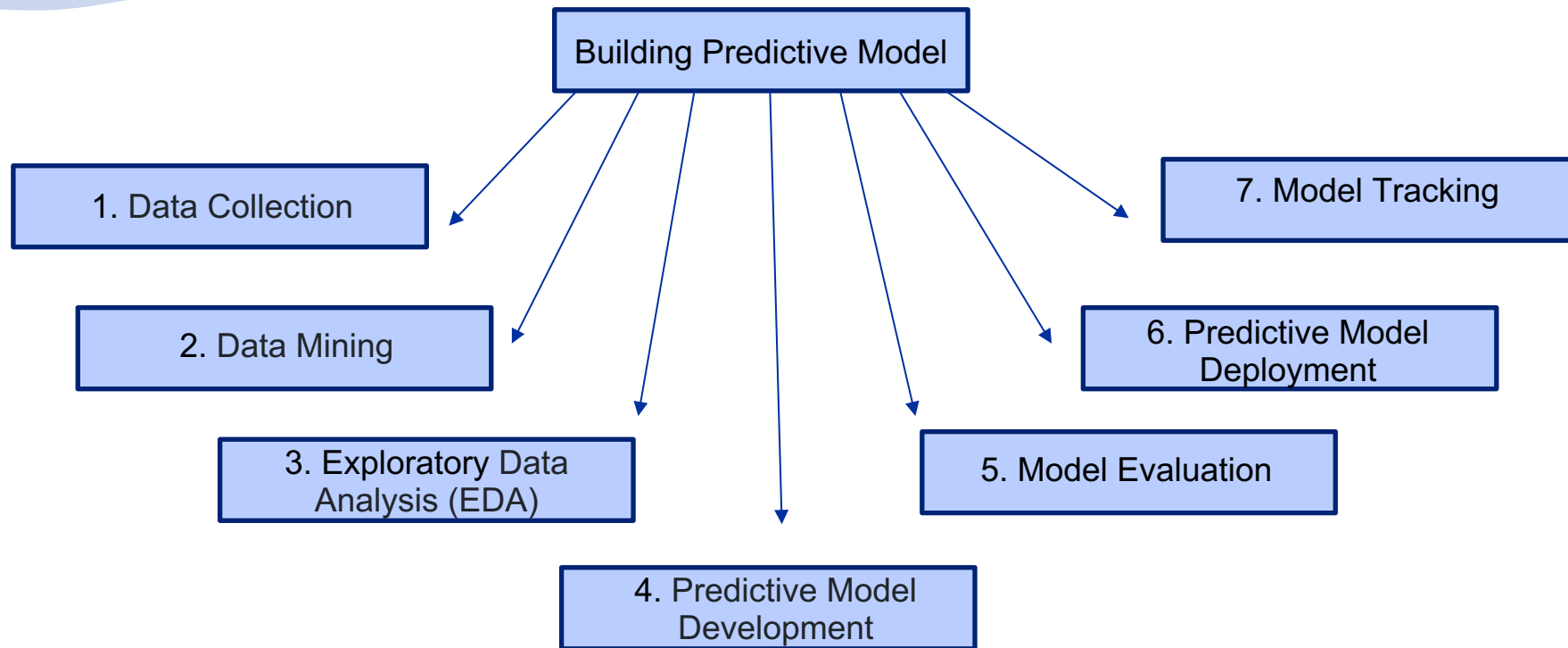
Predictive Analytics



Branch of advanced analytics that predicts future outcomes using historical data combined with statistical modeling, data mining techniques and machine learning.



How to Build a Predictive Model?



Biggest challenge in clinical trials: Patient Recruitment

360 POWERED
POSSIBILITIES

SOLUTIONS SUCCESSES ABOUT RESOURCES

REPORT: The 8 biggest challenges facing clinical trial professionals

AUTHOR
Andrew Burrows
PUBLISHED
30 Nov 2016

2. Site Selection & Recruitment

About 2 million patients participate in clinical trials every year, but to meet U.S. recruitment goals, 6 million patients would be needed. This reality causes delays or budget issues in as many as 90% of trials.

Recruiting and retaining enough participants to complete a trial is among the biggest sources of delays and trial failures.

**DID YOU
KNOW?**

4 common challenges in clinical trials

- Slower patient recruitment than required. One of the biggest challenges is faced just before starting clinical trials, and it's all about patient recruitment. ...
- Compliance with several rules and regulations. ...
- Managing multiple sites. ...
- Preventing professional patients.

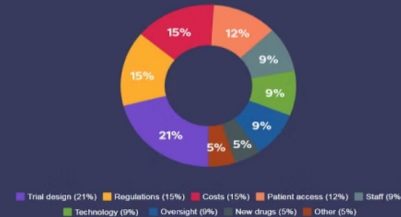
21-Apr-2021



RightPatient

<https://www.rightpatient.com> > professional-patients > 4-...

The biggest challenges facing clinical trial professionals



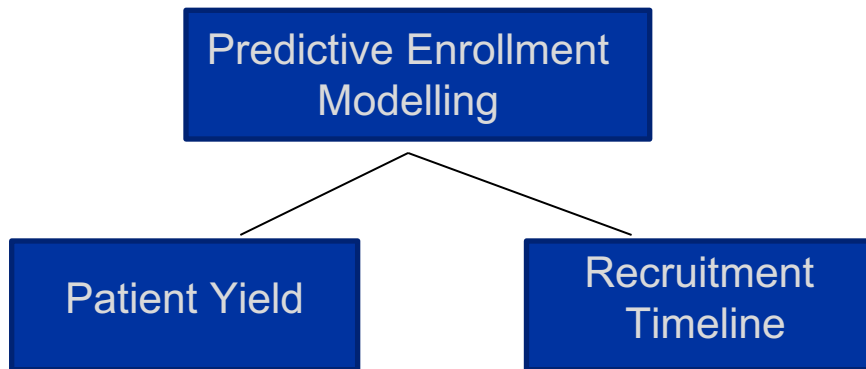


Predictive Modelling scenarios

- **Traditional approach:**
 - Inputs from experts, doctors and scientists who are highly experienced in the area, feasibility survey questionnaire
 - Over optimistic numbers
- **Modern approach:**
 - Predictive Modelling techniques on Big data to find potential patients within a certain age range across a wide geographic area



Predictive Modelling scenarios



- Inclusion/Exclusion criteria, patient motivation, recruitment tactics, development phase, visit schedule impacts patient yield.
- Patients do not enroll in a trial in a linear fashion.



Predictive Modelling scenarios

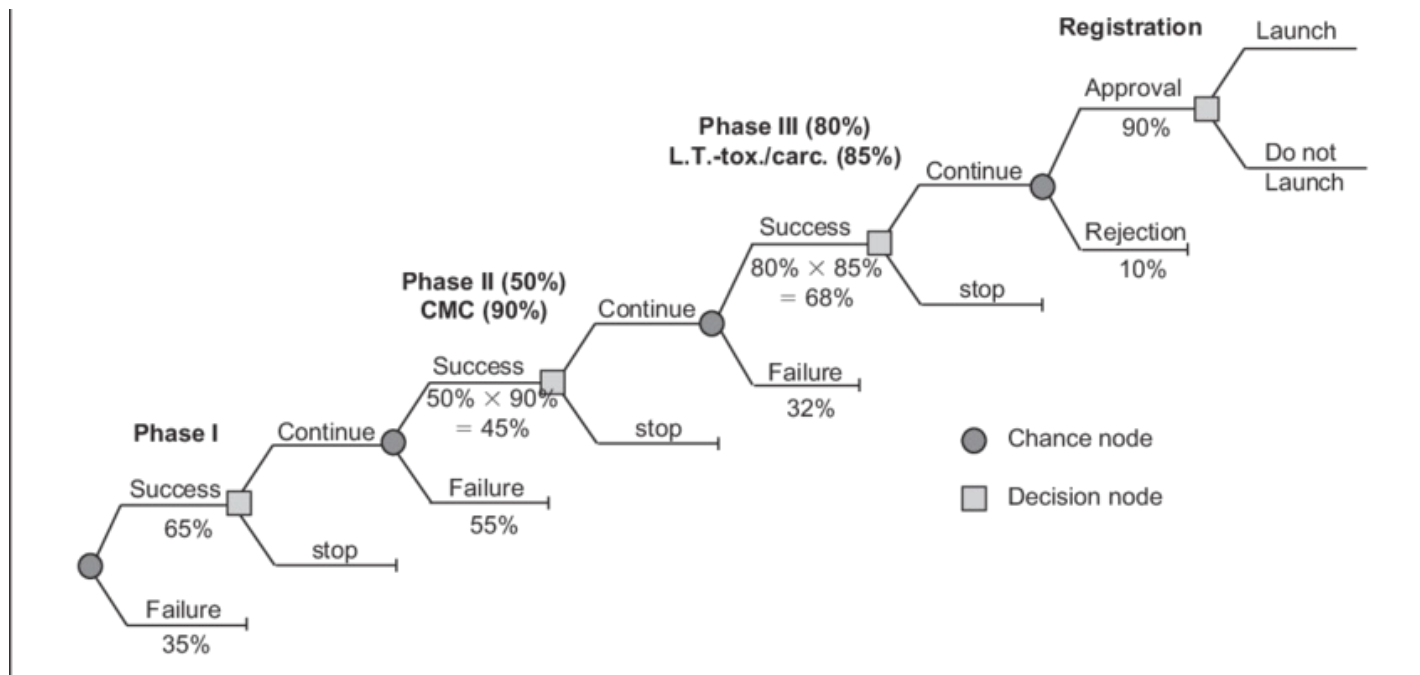
- Predictive model alerts the deviations from the plan before they occur, the study manager and recruitment vendor have the opportunity to adjust recruitment efforts so that enrollment never falls off track
- Using the right predictive model, Patient enrolment can be expedited

Expedited Enrollment = More months saved = Money saved in operating costs

Expedited Enrollment = Product goes to market faster = More Patent life = Increased Revenue

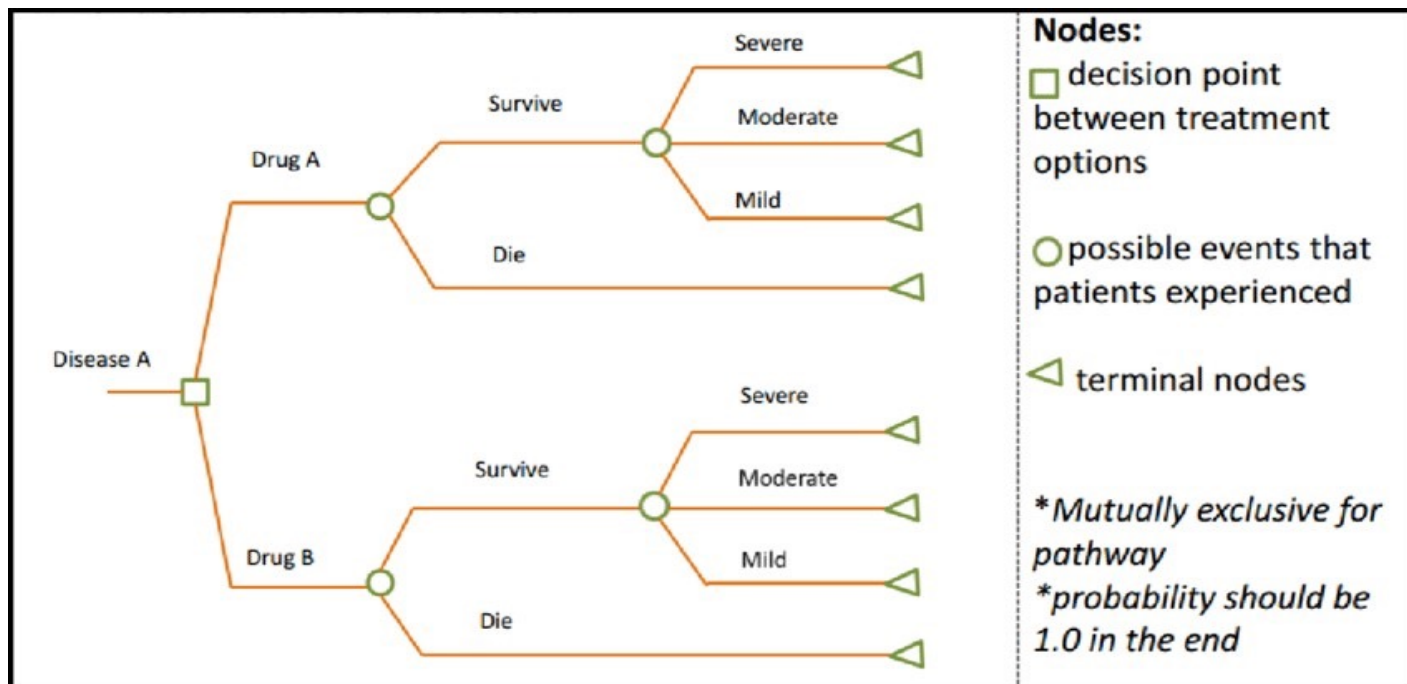
Predictive Modelling scenarios

Problem: To proceed or not proceed with clinical trial for a candidate drug



Predictive Modelling scenarios

Problem: Better treatment decision for a disease





Predictive Modelling scenarios

- **Problem:** Develop a stress prediction tool using heart rate variability data from the Apple Watch electrocardiograph (ECG) sensor
- **Models used:** Random Forest (RF) and Support Vector Machines (SVM)
- **Inputs:** ECG measurements and stress questionnaire data.
- **Results:** Models showed specificity in their capacity to assess “no stress” states but were less successful at capturing “stress” states.
- **Conclusion:** With further development and refinement, Apple Watch ECG sensor data could be used to develop a stress prediction tool.

Predictive Modelling scenarios

Case Study:

(Emily C. Zabor et al. 'Logistic Regression in Clinical Studies' published in International Journal of Radiation Oncology biology physics on Aug 5, 2021)

For patients with recurrence or second primary squamous cell carcinoma in the head and neck after treatment with radiation therapy, some may be offered reirradiation using intensity modulated radiation therapy (IMRT). To investigate factors associated with acute toxicity among reirradiated patients in this setting, a multicenter retrospective analysis was conducted among 505 patients who were initially treated with ≥ 40 Gy of radiation to the head and/or neck and then underwent re-IMRT to ≥ 40 Gy after development of recurrent or second primary squamous cell carcinoma.

Predictive Model Development

Table 1 Univariable logistic regression results

Characteristic	N	OR	95% CI	P value
Age (per 1 year)	462	1.01	0.99, 1.03	.5
Sex	462			.2
Male		—	—	
Female		1.37	0.84, 2.21	
Charlson comorbidity score	459			.6
0		—	—	
1		1.30	0.74, 2.23	
≥2		1.25	0.71, 2.14	
Kamofsky performance status	422			.5
>70		—	—	
≤70		1.23	0.64, 2.29	
rT stage	462			.002
All others		—	—	
rT4		2.02	1.29, 3.17	
rN stage	409			.4
N0		—	—	
N+		1.24	0.78, 1.97	
Systemic chemotherapy	461			.052
No		—	—	
Yes		1.75	0.99, 3.23	
Re-IMRT dose	462			.4
<66		—	—	
≥66		0.81	0.50, 1.30	

Abbreviations: CI = confidence interval; OR = odds ratio.

Predictive Model Development

Table 4 Full multivariable logistic regression model

Characteristic	OR	95% CI	P value
Age (per 1 year)	1.01	0.98, 1.04	.4
Charlson comorbidity score			.6
0	—	—	
1	1.36	0.69, 2.61	
≥2	1.30	0.67, 2.46	
Karnofsky performance status			.4
>70	—	—	
≤70	1.37	0.67, 2.70	
rT stage			.008
All others	—	—	
rT4	2.08	1.21, 3.56	
rN stage			.2
N0	—	—	
N+	1.40	0.83, 2.38	
Systemic chemotherapy			.046
No	—	—	
Yes	1.94	1.01, 3.95	
Re-IMRT dose			.4
<66	—	—	
≥66	0.80	0.46, 1.36	

Abbreviations: CI = confidence interval; IMRT = intensity modulated radiation therapy; OR = odds ratio.

Predictive Model Development

Variable Selection:

- Akaike's information criterion (AIC) to compare different models according to model fit.
- The AIC is a measure of prediction error, so smaller values indicate lower prediction error and therefore a comparatively better model.
- In case of insufficient number of events to support all of the variables (priori clinical interest) -> compare different models and decide what subset of variables to include.
- One common approach is to split variables into subsets based on type and compare models that include different subsets

Predictive Model Deployment

Predictive Analytics Tools in Market



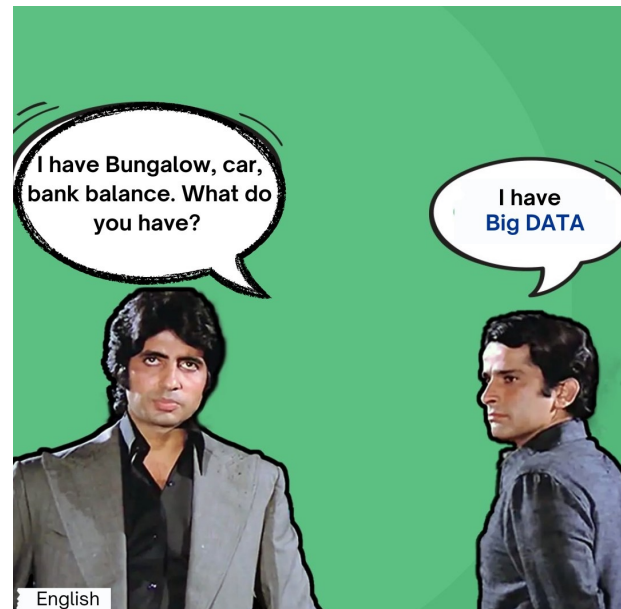


Pros & Cons of Big Data

Pros	Cons
Improves decision making	Cybersecurity risks
Reduce operational costs	Talent gaps
Increases productivity	Compliance considerations
Greater Innovation	Quality & Accuracy of data
Data-driven marketing strategy	Cost and Infrastructure issue
Increases agility	
Recognizes early signs of patient deterioration	

Conclusion

By leveraging Big Data effectively, businesses can gain **valuable insights**. Businesses that use it effectively hold a potential **competitive advantage** over those that don't because they're able to make faster and more informed business decisions.



- **Anastasiya Zharovskikh** 'How AI is Improving Predictive Analytics in Healthcare' 24 Sept- 2020
- **Patrick Schober** '*Anesthesia & Analgesia*' Published online 2021 Jan 14
- **ProjectPro** 'Predictive Modeling Techniques- A Comprehensive Guide [2023]' 15 Jul 2023
- **IBM Newsletter** 'What is exploratory data analysis?'
- **Abhishek Sharma** 'Random Forest vs Decision Tree | Which Is Right for You?' Published On May 12, 2020 and Last Modified On July 31st, 2023
- **Emily C. Zabor et al.** 'Logistic Regression in Clinical Studies' published in International Journal of Radiation Oncology biology physics on Aug 5, 2021
- www.researchgate.net

Q & A

Thank you!



**The Global Healthcare
Data Science Community**

Contact Channels

📞 UK +44 1843 609600
✉ office@phuse.global
🌐 phuse.global

Social Media

🐦 [@phusetwitta](https://twitter.com/phusetwitta)
📘 [/phusebook](https://www.facebook.com/phusebook)
▶ [/phusetube](https://www.youtube.com/channel/UCphusetube)
🌐 [/company/phuse](https://www.linkedin.com/company/phuse)