TO STUDY THE ROLE OF BIG DATA AND DATA SCIENCE IN CRITICAL CARE UNIT

Dissertation Submitted to

The IIHMR University, Jaipur

for the partial fulfillment for the award of the
degree of Master of Business Administration (MBA)
in
Hospital and Health Management

by

Dr. Varda Dixit

Under the Supervision and Guidance of

Dr. Sandesh Kumar Sharma

2022

TO STUDY THE ROLE OF BIG DATA AND DATA SCIENCE IN CRITICAL CARE UNIT

Dissertation Submitted to

The IIHMR University, Jaipur

for the partial fulfillment for the award of the
degree of Master of Business Administration (MBA)
in
Hospital and Health Management

by

Dr. Varda Dixit

Under the Supervision and Guidance of

Dr. Sandesh Kumar Sharma

2022

DECLARATION BY STUDENT

I hereby declare that this dissertation titled **To Study the Role of Big Data and Data Science in Critical Care Unit** is the bonafide record of my original research work. I declare that it has not been submitted to any other university or institution for the award of any degree or diploma. Information derived from the published or unpublished work of others has been duly acknowledged in the text.

Name of Student: Dr. Varda Dixit

Date: 28/06/2022

m16
LABS

Internship Completion Letter

Internship Completion Letter

Date : Mar 22, 2022 to 19th July, 2022

Subject : Completion of Internship Letter

To Whomsoever It May Concern,

M16labs certifies that Dr. Varda Dixit successfully completed the internship program at our organization as a Business Development Intern.

During this time, Dr. Varda worked on the development of key health Information Technology related business development projects.

By working on these projects, Dr. Varda acquired critical Professional skills.

Dr. Varda displayed professional traits during her internship period and managed to complete all assigned tasks as requested. She was hardworking, dedicated, and committed. It was a pleasure having her with us in this period.

Yours Sincerely,

Anil B

Founder & CEO

Meta16labs Healthcare & Analytics Private Limited

Reg Off: Tech Precision Solutions No 8/2 Ulsoor Road, Bengaluru, Karnataka 560042

m16labs.com | metahos.com | contact@m16labs.com | 07941056988

CERTIFICATE

This is to certify that dissertation titled **To Study the Role of Big Data and Data Science in Critical Care Unit** is a record of the research work undertaken by **Dr. Varda Dixit** in partial fulfillment of the requirements for the award of the degree of MBA (Hospital and Health Management) from the IIHMR University, Jaipur under my guidance and supervision and his/her approach to the research study has been sincere, scientific, and analytical.

Dr. Sandesh Kumar Sharma Faculty Guide IIHMR University, Jaipur	Dr. Mahender Kumar School Dean IIHMR University, Jaipur
Date:	Date:

ABSTRACT

TITLE: To Study the Role of Big Data and Data Science in Critical Care Unit

STUDY OBJECTIVES:

- To study the role of big data in Critical Care Units
- To study the benefits of big data in Critical Care Units
- To study the challenges of big data in Critical Care Units

METHODOLOGY:

A descriptive study conducted through secondary research or desk review using publicly available data sources, to identify the role of big data in critical care along with its benefits and challenges. Data has been taken from consulting reports, few studies done on Critical Care in Asian environment and from countries like USA, Spain, Germany, Japan and Korea who already have well established systems which have proven to be more effective in terms of clinical care outcomes.

CONCLUSION:

We can improve our descriptive and predictive analytics based on real data thanks to big data. It won't eliminate the need for clinical trials, but it will help us design them better by precisely characterizing the questions and calculating sample sizes and event rates. It will be crucial to protect concerns like data security, governance, and confidentiality as big data becomes more widely used. Big data have the power to individualize medicine and bring together various sources of data about each individual patient, which can greatly enhance medical care and lower expenses.

ACKNOWLEDGEMENT

I would like to acknowledge everyone who played an important role in completion of this dissertation.

First and foremost, I would like to thank my mentor at M-16 Labs, Mr. Arnav Mahurkar for his constant support and guidance in these 2 months.

Secondly, I would like to express my gratitude to Dr. Prahlad Rai Sodani, President of IIHMR University, Jaipur for encouraging me to take this dissertation.

Also, I am extremely thankful to my faculty guide Dr. Sandesh Kumar Sharma for his valuable guidance, keen interest, and encouragement throughout the completion of this project.

This internship has not only enriched me professionally but also on a personal level.

I would like to thank my father, Mr. Sanjeev Dixit, and my mother, Mrs. Pomica Dixit for giving me this opportunity of pursuing a career of my choice and being so supportive throughout my education even in the time of this ongoing pandemic.

I would also like to thank a few friends who encouraged me throughout this time and helped me on every obstacle. I would like to thank Rishabh and Eshita for being my stressbusters, motivators, and proof-readers.

I would like to show gratitude to all my family and friends for their support.

Dr. Varda Dixit

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

- **EHR:** Electronic Health Record
- **ICU:** Intensive Care Unit
- **CIS:** Clinical Information System
- **AUC:** Area Under the Receiver-Operating Characteristic Curve
- **MIMIC:** Multiparameter Intelligent Monitoring in Intensive Care
- **NLP:** Natural Language Processing
- **BDA:** Big Data Analysis
- **ML:** Machine Learning
- **AI:** Artificial Intelligence
- **NHIS-NSC:** National Health Insurance Service-National Sample Cohort
- **APACHE:** Acute Physiology and Chronic Health Evaluation
- **IDC:** International Data Corporation
- **EMR:** Electronic Medical Record
- **MPM:** Medical Practice Management Software
- **IoT:** Internet of Things
- **IOM:** Institute of Medicine
- **HITECH:** The Health Information Technology for Economic and Clinical Health
- **SAPS:** Simplified Acute Physiology Score
- **DCDM:** Dynamic Clinical Data Mining

TITLE: TO STUDY THE ROLE OF BIG DATA AND DATA SCIENCE IN CRITICAL CARE UNIT

1: INTRODUCTION

Big data is a term used to describe a collection of unstructured and structured data that is gathered by businesses and can be used for machine learning applications, predictive models, and other sophisticated analytical tasks.

The digitization of healthcare systems, which has produced a flood of massive clinical data, is accelerating the development of data science in medicine. Critical care is particularly pertinent to data science, a branch of research devoted to the primary extraction of knowledge from complex data. The quantity of data in the ICU, the need for improved evidence-based care, as well as the complexity of serious illnesses make the use of data science methodologies and data-driven research particularly tempting to critical care clinicians. Even with the rising number of research and studies in this area only a few data science projects have effectively integrated data-driven solutions in the ICU. However, emergency physicians need to be aware of the opportunities and difficulties posed by big data, given the anticipated expansion in this field.

The digitization of the healthcare system is transforming how we practice medicine and conduct clinical research. Electronic medical records (EHRs) are becoming more and more common, opening the door for big data analysis, and bringing it to the patient's bedside. Critical care units in the healthcare system are a particularly striking illustration of how data science is being used to enhance patient care. There is little concrete evidence of the procedures carried out in the ICU, and clinical variability is considerable. Additionally, the complexity of catastrophic illnesses undermines the conventional reductionist approach to medical study. In other words, it is unlikely that research on single-path biomarkers or single-agent interventions will support the clinical reality of the intensive care unit. Intensive research necessitates an integrated approach that considers both the complexity of severe illnesses and the computational techniques and algorithms that make it possible. This method also generates a lot of data, which is then transformed into digital form. Fresh data is continuously produced using bedside monitors, drug pumps, EHRs and ventilators. With most recent developments in molecular diagnostics, these data will soon be pulled from a wide range of genomes, transcriptomes, microbiomes, and other biological systems.

As data rapidly permeate most parts of clinical research and, ultimately, clinical treatment in ICU, healthcare workers need to get familiar with the advantages and hazards of these approaches, big data definitions, algorithm types, applications, obstacles, and its future in critical care.

2: LITERATURE REVIEW

S.No.	Study Title	Authors	Key Learnings
1.	"Big Data and Data Science in Critical Care"	"L. Nelson Sanchez-Pinto, Yuan Luo, and Matthew M. Churpek"	• Algorithms of data science can be classified into supervised, unsupervised, Semi-supervised (hybrid of the two), and deep learning algorithms (defying the first two; focus on learning data representations. • The following are examples of how data science is used in critical care: I. Prognostic Models and Predictive Models II. Phenotyping and Clustering III. Non-traditional data type applications, including image analysis, physiological waveform analysis, and natural language processing. • Data science challenges in critical care: 1. Despite of enormous research and studies very few projects have been successfully implemented in data-driven systems of the ICU. These leads to a level of mistrust of the clinicians in these systems. 2. Overfitting is the phenomena where a model performs well on training data but performs poorly on real-world data. 3. When digital technology is used incorrectly, it can harm people. Data

			scientists, clinical domain experts, medical informaticians, and implementation science professionals can all help to lower this risk by taking a rigorous approach to review and implementation. 4. The perceived loss of autonomy in the face of more complicated computing systems is another concern among therapists.
2.	"Big data and machine learning in critical care: Opportunities for collaborative research"	"A. Núñez Reiz, Miguel Ángel Armengol de la Hoz, Rodrigo Octavio Deliberato, Arturo González-Ferrer, Julio Mayol, and Joan Cháfer"	- Critical care involves big data analysis (BDA), which calls for high-quality data input from clinicians and nurses as well as other elements like machine learning (ML). - When data and machine learning are combined, artificial intelligence (AI) is created. This technology recognizes patterns in data that comprises either categorization or prediction and applies software output to extremely particular tasks, such as automatic medical image interpretation. - Technology plays a big part in the care provided in the ICU with data that has machine learning applications in areas like NLP, pharmacovigilance, and decision support. - Teams including members with clinical skills, such as epidemiologists, biomedical researchers, and computer science backgrounds, such as data scientists and

			statisticians, are assembled with specific goals because BDA and ML are two distinct fields of work. - In the future, computers that manage data integration and analysis will increasingly support clinical decision-making. Therefore, the practitioners will have to learn a critical care specialty that leverages data to tailor care and enhance population health. - Collaboration with other non-healthcare data science experts is essential to making the most of all the data that is routinely gathered throughout the course of care without violating patient privacy rights.
3.	“Big data in healthcare: management, analysis, and future prospects”	“Sabyasachi Dash, Sushil Kumar Shakyawar, Mohit Sharma & Sandeep Kaushik”	- Data science works with a variety of topics, including data management and analytics, to acquire deeper insights into enhancing a system's performance and offerings. - The three dimensions of data growth are “volume, velocity, and variety (the 3 Vs)”. Big data refers to information whose volume is vast. Velocity provides an indication of the rate of data collection, making analysis simpler. Variety, on the other hand, makes comments on the various kinds of organized and unorganized data that organizations and systems might gather. Examples include log files, music, video, or transaction-level data.

		- EHRs are computerized patient medical records that are stored in an electronic system to gather, send, receive, store, buy, mix, and work with multimedia data primarily for the aim of offering healthcare and healthcare-related services. - Advantages of EHR: 1. Healthcare professionals now have more access to a patient's complete medical history, including diagnoses, prescriptions, details about known allergies, demographics, clinical narratives, and the results of multiple laboratory tests. 2. By utilizing automatic reminders and prompts for immunizations, aberrant laboratory findings, cancer screening, etc., clinicians can dramatically enhance their medical practices. 3. Thanks to minimal paperwork, electronic permission, and quick insurance approvals. 4. By reporting illness outbreaks as soon as they occur, they promote reporting of important healthcare quality indicators to organizations and speed up data retrieval. This enhances public health surveillance. 5. Together, EHRs, EMRs, PHRs, MPM software, and many other healthcare data elements have the power to lower

		healthcare costs while also increasing service effectiveness and quality. - - The "Internet of Things (IoT)" is a significant source of data in the healthcare sector since it creates a continuous stream of data while tracking people's health (or patients). - IoT devices may measure and monitor a variety of indicators, and early intervention and treatment may prevent a patient from needing hospitalization, which would result in a considerable decrease in healthcare costs. Data from these devices can be used to monitor one's own health, model the spread of a disease, and identify strategies for containing an outbreak of a certain disease. - The technique of providing healthcare and public health services via mobile devices, particularly in the treatment of chronic illnesses like cancer and diabetes, is known as "mobile health," or "mHealth." - Challenges in Big Data analysis: 1. Large volume of information. 2. Heterogeneity of data - The ML algorithm can transform a medical imaging diagnostic system into automated decision making. - By actively extracting illness indicators from biomedical images, image analysis has an impact on medical care. This strategy transforms how patients are

			identified, treated, and monitored by mining insights from massive amounts of clinical imaging data using machine learning and pattern recognition tools. It focuses on enhancing medical imaging's diagnostic skills for use in clinical decision-making. - Translational bioinformatics is at the forefront of data-driven healthcare and applies bioinformatics tools to turn biomedical and genomic data into predictive and preventative health. - Big data challenges in healthcare: 1. Data storage for huge volumes 2. Cleaning 3. Unified format 4. Accuracy 5. Image Pre-processing 6. Security 7. Meta-Data 8. Querying 9. Visualization 10. Data sharing
4.	"The Evolution of the Electronic Health Record"	"Susan Doyle-Lindrud"	- From a time when a doctor's mind was the only place where knowledge and information processing ability existed, to a future in which information technology would offer knowledge and the processing power to apply it to specific patient data. - Electronic medical records (EHR) have seen tremendous development since they replaced paper-based records.

			- Healthcare providers point out that there are still issues with data exchange between various locations. - Problems before EHRs 1. 1. By making healthcare data easily accessible to authorized users wherever they are, the logistical issues associated with paper recording are eliminated. 2. Reduces the clinical accounting work required to manage patients. - The HITECH Act for Economic and Clinical Health promotes the use of interoperable electronic medical records and the interchange of medical data to lower healthcare costs and raise the standard of care. - When using EHRs, doctors expressed frustration with the system's poor usability, time-consuming data entry, disruptions to in-person patient care, ineffective and unfulfilling work content, inability to share health information, and deterioration of clinical record. EHRs are also more vulnerable to cyberattacks and data breaches. - Increase in standardization and security of the EHR is an urgent need of the hour.
5.	“"Big data" in the intensive care unit.	“Leo Anthony Celi, Roger G Mark, David J Stone, Robert	- A substrate for data that generates, records, and archives individual clinical encounters, enables data scientists and doctors to solve knowledge gaps & errors, and supports

Closing the data loop”	A Montgomery”	crowdsourcing versions for the creation of evidence in clinical practice. - The results of intervention in the ICU are particularly susceptible to change between patient subgroups and clinical scenarios due to the environment's exceptional complexity, which is why there are so few data analysis studies in this context. - A complete and extremely detailed critical care database that encourages clinical research, fosters the creation of clinical decision-support systems, allows testing and completion of algorithms using real-world data is required. This database must be technically feasible. - Critical Care Databases: • ICU patient databases, both for-profit and nonprofit purposes, have been created to assess and compare the severity of ICU patient conditions and outcomes, as well as the costs of care, across participating ICUs. Patient demographics are typically archived, and information particular to the unit and hospital is integrated with data on the underlying sickness, severity of illness, and patient demographics (e.g., length of stay, mortality, and readmission). - Data-driven tools: Example – If ICU resource limitations result in the localization of critical care, MIMIC may assist emergency

			departments. The goal is to reduce patient complaints, labor costs, iatrogenic anemia, increased laboratory costs, and unnecessary tests that cause malpractice due to false positive results. - Artificial intelligence techniques were also used in the MIMIC database to model the interim treatment and physiological response of patients with gastrointestinal bleeding to predict whether laboratory tests have changed significantly from previous measurements. - In a research setting, decision analysis is a data-dependent procedure that methodically evaluates difficult decisions and quantifies the anticipated benefits and drawbacks of various therapeutic approaches. - Cross-validation of models between institutions identifies generalizable and institution-specific results. Most significantly, knowledge discovery will greatly speed up when more researchers start using clinical data mining.
6.	“State of the art review: the data revolution in critical care”	“Marzyeh Ghassemi, Leo Anthony Celi, David J Stone”	- The ability to use data to support decision-making is crucial in intensive care units since the impact of each intervention is influenced by patient and clinical context-specific elements as complexity goes beyond general cognitive capacity. - - The main purpose of all ICU scoring systems, including APACHE, MPM, and

		SAPS, is to evaluate and compare ICU performance. These systems are based on adjusting severity using physiological and other clinical data (For example, by looking at the ratio of actual and expected results). They can also be utilized as a quick indicator of the acuity of a patient. - Current clinical practise provides big data with the opportunity to generate real financial gains through reliable predictive alerts, retrospective report analysis for expensive patients, readmissions, triage, clinical decompensation, adverse events, and treatment optimization for diseases affecting multiple organ systems. - - DCDM (Dynamic Clinical Data Mining) presents EMR users with patient findings identical to the process automatically, making the patient's choice objective. - DCDM aggregates electronic health records of individual patients during treatment. Using modified search engine technology, it queries a universal unidentified clinical database in real time to identify and propose previous cases with sufficient similarity to reveal the current case. Enter relevant decision-making support materials such as interventions and prognosis into the individual patient's EMR. Based on previous treatments and results.

			- The patient's record is made up of the patient's specific EMR and the relationships between it and the complete population database, which the doctor only needs to see as part of the patient's record. The population-level database's knowledge can improve the fundamental starting point for each individual patient, and clinical decisions can be made by combining and comparing the resulting information. - The use of open data in the healthcare industry necessitates data sharing, which by definition calls for some degree of accountability and interdisciplinary cooperation. This includes data marathons, hackathons, and crowdsourcing. These include some open, thorough, and anonymous EMRs along with coordinated hackathons, crowdsourcing technology, and user feedback gathering from the various sorts of stakeholders that are required. - To address the most urgent issues, such as who oversees shared data, who owns the rights to clinical data, and how to combine heterogeneous and outwardly non-interoperable data, organizations and researchers need to build approaches and processes.
7.	“Critical Care Research Using "Big Data": A	“Kwangha Lee”	- The APACHE score is one of the most reliable prognostic tools for determining

	Reality in the Near Future”		the probability of mortality in critically ill patients. - Recently, physiologic waveform analysis, image analysis, and natural language processing have all been used in data science studies on clusters and phenotypes. - Data from the National Health Insurance Service-National Sample Cohort (NHIS-NSC), which is regarded as typical of the general population, is the foundation for many big data studies on critical care in South Korea. NHISNSC is being used in a number of studies on acute care and critical care facilities, such as Impact of Immunosuppressive Drugs on In-Hospital Mortality, Characteristics & Mortality in All ICU Patients Over the Past Few Years.
8.	“Exploiting big data for critical care research”	“Annemarie B Docherty, Nazir I Lone”	- Big data has the potential to revolutionize the healthcare industry by advancing research methods, accelerating the pace of health improvement, and lowering costs. - In the era of big data, maintaining public trust requires assuring the data's security, governance, and secrecy. - Limitations: 1. Chance – Testing multiple hypotheses and analyzing various subgroups may give more false positives. 2. Error – Big data isn't always collected for research; hence, it can be compromised with respect to quality, accuracy, and completeness.

			3. Bias and Confounding – Data collected from a certain source may sometimes exclude people from different ethnicities, age, background etc. - Ethical and Legal Issues: 1. The biggest concerns in using Big Data for medical research is invasion of privacy and identification of an individual without consent. 2. Anonymizing data may only take us so far as linked datasets might pose the risk of reidentification. Also, insisting on formal consent of big data research may lead to a participation bias which can render any analyses results useless. 3. It is important that we have measures in place to safeguard the patient's privacy before we begin researching with Big Data. - Use of big data in clinical or critical care will continue to grow because of the increasingly available and growing datasets. Combining big and small data also enables the possibility of precision medicine.
9.	"Medical Big Data Is Not Yet Available: Why We Need Realism Rather	"Hun Sung Kim, Dai Jin Kim, Kun Ho Yoon"	- We are a long way from starting to use medical big data for research due it's limitations. - Good quality data is far more important than quantity. Clean and accurate data can give meaningful results using just traditional analytical methods.

	than Exaggeration”		- Collaboration of medical staff and Data Analysts is essential to maintain the quality of data by putting in place certain protocols and guidelines. - Standardization and an international coding system would take us a long way and eliminate challenges like data quality, numerous biases and maintaining operational definitions. It will also enable not only domestic but international partnerships.
10.	“mHealth and big-data integration: promises for healthcare system in India”	“Samaneh Madanian, Dave T Parry, David Airehrour, Marianne Cherrington”	- Mobile computing, medical sensors, and portable gadgets are all integrated into the emerging field of mHealth, which is made possible by mobile devices and connectivity. - Applications of mHealth: 1. Helpful in enhancing patient safety and the standard of treatment while keeping track of health condition. 2. Offers remote support and data gathering that may be expanded and applied across communities to identify common trends and raise healthcare standards. - Benefits of mHealth in India: 1. mHealth could be a relief as India’s population lies predominantly in rural areas where there is a lack of healthcare resources and access, and people struggle to get proper and timely healthcare services.

		2. More than 86.2 percent of the population has access to inexpensive internet-connected devices as a result of the cost decrease of mobile handsets and the extension of their network coverage. 3. Mobile health applications that are accessible to those with low literacy levels are popular since the majority of people in India are illiterate and underprivileged. 4. The internet, smartphones, IoT, sensor, and mobile devices can all be sources of big data in the healthcare industry. Through these sources, information can be extracted and transformed into priceless insights and knowledge acquisition by performing big-data analysis. This results in a more holistic, patient-centric approach that prioritizes treatment quality while balancing pressures on healthcare spending and cost.

3: AIM & OBJECTIVES

3.1 – Aim:

The aim of this study is to analyze how can big data be helpful in management of critical care units in the healthcare industry in a more efficient and effective way. And, to understand the major challenges in harnessing big data for the same?

3.2 – Objectives:

- To study the role of big data in Critical Care Units
- To study the benefits of big data in Critical Care Units
- To study the challenges of big data in Critical Care Units

4: METHODOLOGY

4.1 – Study Design:

Narrative/Traditional Literature review

4.2 – Source of Data:

- Consulting Papers
- PubMed
- Google Scholar
- Government Websites
- Government Publications
- Industry experts' comments

4.3 – Inclusion Criteria:

- All articles written in English language were included.
- All open access articles were included.
- Only published articles were included.
- All articles relevant to our research were included.

4.4 – Exclusion Criteria:

- All articles that did not specify any of the above-mentioned details in the inclusion criteria were excluded.
- Any article that was not included in full text category of articles was excluded.
- All unpublished and unfinished articles were excluded.

4.5 – Screening of Literature:

Citations of all the search results were obtained and uploaded to an online citation manager. All the duplicate results that were found were excluded effectively. Furthermore, the search included some articles which were not relevant to the research study and were excluded along with articles categorized as bulletin or review articles.

All the articles which were written in any other language apart from English were also excluded. In addition to this, articles were screened to check their compliance with the inclusion criteria and those that did not seem relevant enough were excluded from the study.

4.6– Data Extraction:

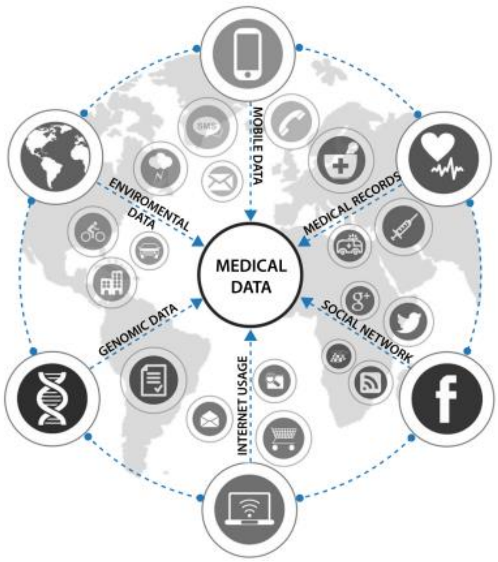
10 articles obtained from 3 different online databases i.e., ScienceDirect, PubMed, Google Scholar, that were related to the topic of research were used to extract information for this study. Due to different style of approaches in different studies, a narrative synthesis was done to establish the status of big data presently in the health industry and benefits it provides for managing a critical care unit and also to find out about the risks involved in the same.

The articles were searched thoroughly for relevant information that fulfils all the three objectives of the study. Under-given results were obtained by the information gathered using these selected 10 articles.

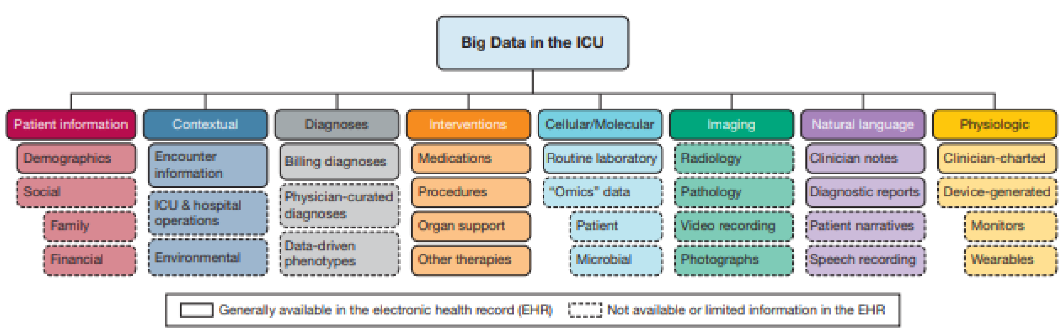
5: RESULTS & DISCUSSION

5.1– The Role of Big Data in Critical Care Unit

- Sources of data

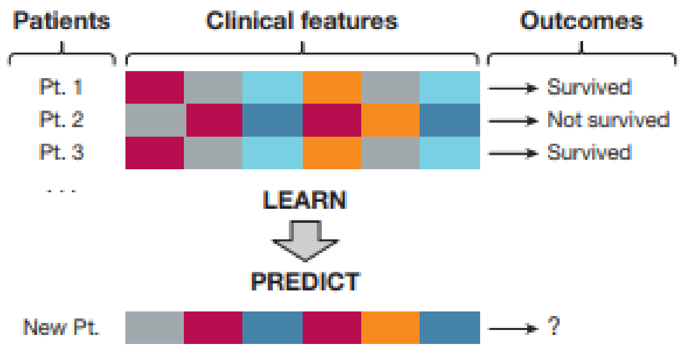


- Electronic Medical Records
- Social Network
- Internet Usage
- Genomic Data
- Environmental Data
- Mobile Data



- Application of data science in healthcare includes:

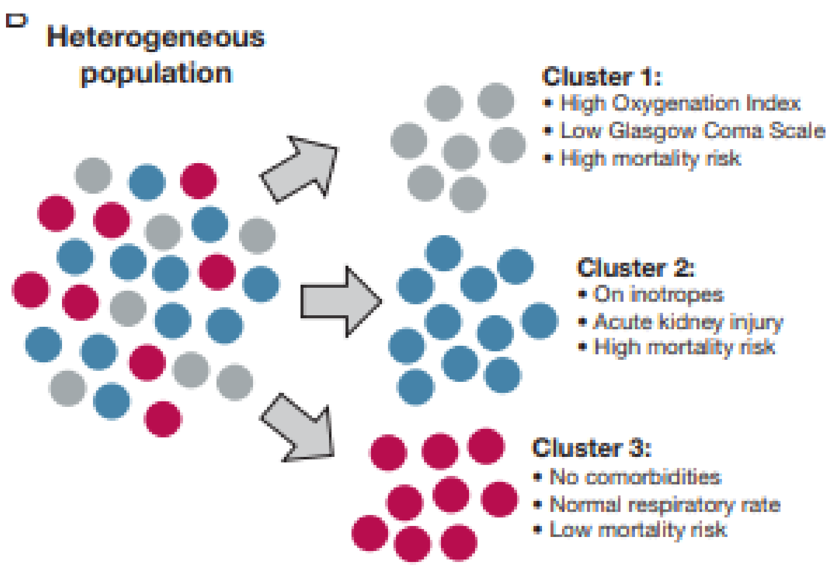
1. Predictive Models & Prognostic Models



   - Predictive models are typically trained to foresee circumstances, occurrences, or actions. For instance, in oncology, biomarker-based models are used to identify cancer subtypes that are amenable to specific treatments.
   - To predict potential disease-related endpoints or outcomes, prognostic models are carefully trained. As an illustration, APACHE scores are used to calculate the probability of patient mortality in intensive care units.

2. Clustering and Phenotyping

   - In critical care, unsupervised learning algorithms have primarily been employed to identify naturally existing patient groupings or clusters that have comparable clinical and/or genomic characteristics. These are frequently referred to as subtypes, phenotypes, or subphenotypes.

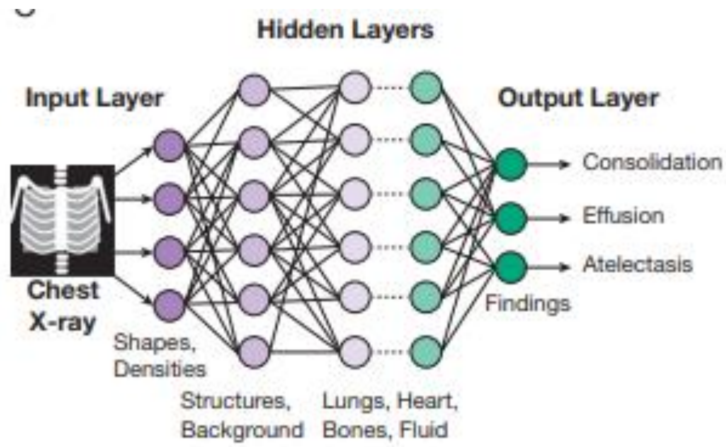


- As an illustration, clustering analysis was used to identify and validate six clinically discernible subgroups of ICU patients who, despite having the same diagnosis, differed greatly in all baseline traits and clinical trajectories.

3. Applications that use unusual data types
   - *Natural language processing (NLP)*

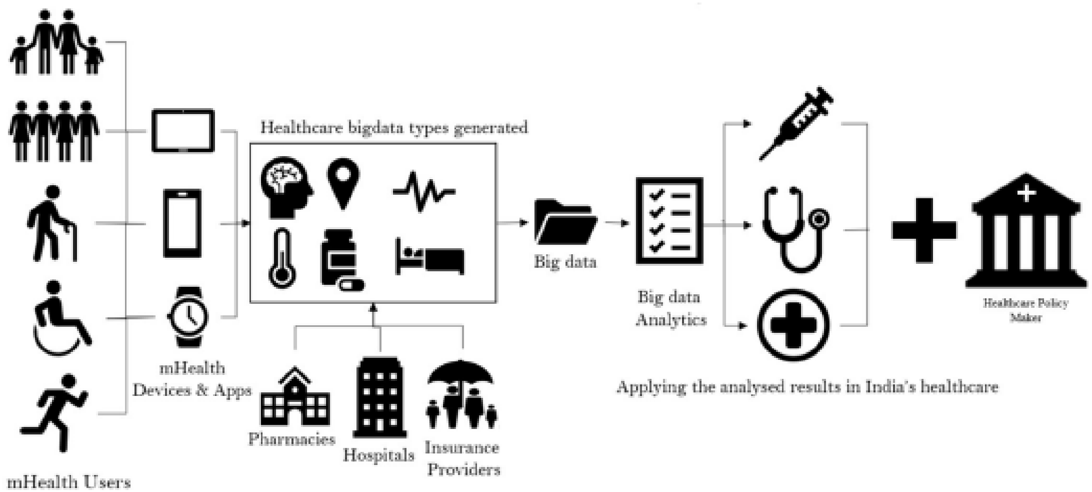
     "NLP can be used to analyze unstructured data such as progress notes, discharge summaries, nursing notes, and diagnostic reports." It extracts textual properties that can be applied to various tasks using task-specific algorithms. Certain types of patients can be identified using NLP.
   - *Physiological waveform analysis*

     "Physiological waveform data refers to information obtained through bedside monitors and wearable technology, such as electrocardiography, photoplethysmography, impedance pneumography, invasive arterial manometry, end-tidal capnography, and electroencephalography." Example: Utilizing photoplethysmography data to detect hypovolemia and integrated physiological data to predict hyperlactatemia are all examples of how waveform data has been used by researchers.
   - *Image analysis*

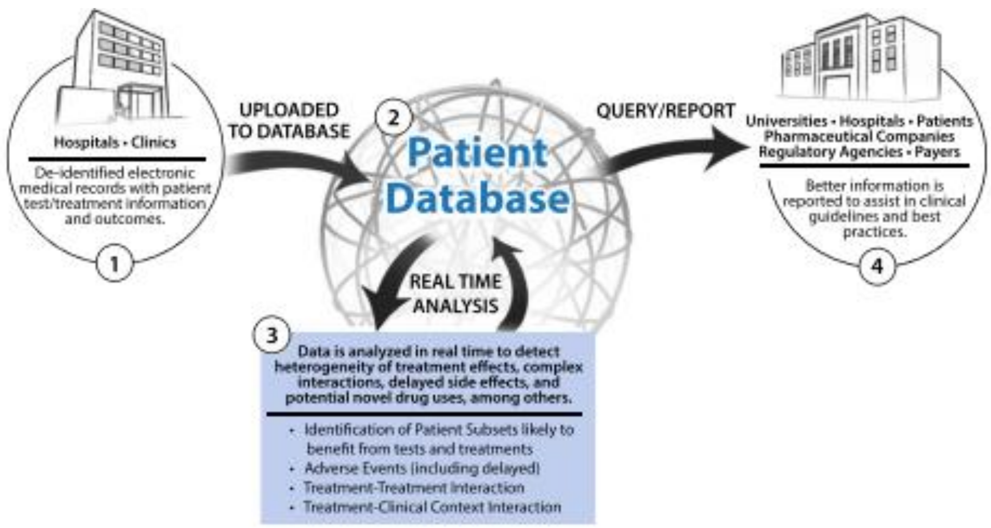


"Deep learning-based advances in image analysis techniques to identify pulmonary pathology in chest radiographs as well as normal and abnormal findings in brain and abdominal imaging may be especially useful in intensive care units (ICUs) where the availability of specialists who can quickly and accurately interpret radiographic images is constrained." However, their efficacy and safety must first be thoroughly assessed before any clinical implementation is taken into consideration.

5.2– The Benefits of Big Data in Critical Care Unit



1. Access to a patient's complete medical history, including diagnoses, prescriptions, information about known allergies, demographics, clinical narratives, and the outcomes of numerous laboratory tests, has improved for healthcare practitioners.
2. By utilizing automatic reminders and prompts for immunizations, aberrant laboratory findings, cancer screening, etc., clinicians can dramatically enhance their medical practices.
3. As a result of fewer paperwork, electronic permission and rapid insurance approvals.
4. By reporting illness outbreaks as soon as they occur, they promote reporting of important healthcare quality indicators to organizations and speed up data retrieval. This enhances public health surveillance.
5. Together, EHRs, EMRs, PHRs, MPM software, and many other healthcare data elements have the power to lower healthcare costs while also increasing service effectiveness and quality.
6. Evidence-based medicine is becoming more prevalent in healthcare, and a continual, systematic examination of clinical data is required for efficient decision-making.



7. mHealth devices can detect and monitor a variety of indicators, such as electrocardiography, levels of hypertension and hypotension, and health vital signs. The analysis of this data helps to give appropriate, individualized care and reveals insightful patterns in people's health.
8. Health care errors can be prevented, safety can be improved, and inappropriate care may be avoided with the use of patient data and regular updates to doctors. Giving the person proper advice or even medical prescriptions for their illnesses can be helpful.
9. The use of big data by doctors may make it simpler for them to prescribe medications and diagnose ailments. This could aid in the analysis of prevalent diseases in a region, aid in the monitoring of epidemics, assist in the appropriate response, and help take the required precautions at the appropriate time before the situation worsens.
10. With assistants conducting basic health assessments, storing them in the cloud, and sending them to central digital health teams that quickly assess the information and quickly approve treatment plans before sending prescriptions to caregivers, big data may increase healthcare service coverage to remote areas of India where medical professionals are scarce. This information can be used to extract the pattern of disease, teach it to digital health assessors, and reduce emergency room visits and hospital stays.

5.3 – The Challenges of Big Data in Critical Care Unit

- Limitations:
  1. Chance - Testing several hypotheses and looking at different subgroups may result in more false positives.
  2. Error - Since big data is not always gathered for research, its quality, accuracy, and completeness may be at risk.
  3. Bias and Confounding - Data gathered from one source may occasionally omit individuals of various ethnicities, ages, backgrounds, etc.

- Ethical and Legal Issues:
  1. The biggest concerns in using Big Data for medical research is invasion of privacy and identification of an individual without consent.
  2. Anonymizing data may only take us as far as linked datasets might pose the risk of reidentification. Also, insisting on formal consent of big data research may lead to a participation bias which can render any analysis results useless.
  3. It is important that we have measures in place to safeguard the patient's privacy before we begin researching with Big Data.

- Challenges of big data in healthcare:
  1. Storage of large volume of data
     - Big data storage is one of the main issues, and many businesses do it on-site. Control over security, access, and uptime are just a few benefits.
     - However, scaling it costs money, and maintaining it is challenging.
     - Additionally, because it offers low prices and great reliability, cloud-based storage employing IT infrastructure is popular in healthcare enterprises.
  2. Cleaning
     - Data needs to be cleaned up or scrubbed, either manually or automatically, after acquisition to ensure accuracy, correctness, consistency, relevancy, and purity.

3. Unified format
    - Big data management is too challenging, especially when it is presented to healthcare providers without perfect data organization and with a requirement to codify all information surfaced for the purposes of claims, billing, and clinical analytics.
    - Core clinical concepts are developed by CPT, ICD and other medical coding systems. But they have their own limitations.
4. Accuracy
    - Inaccurate reporting of patient data is caused by poor EHR functionality, complicated workflows, and a flawed understanding of why big data is crucial to capture well.
    - The use of patient self-report questionnaires for their symptoms may enhance the quality of the record.
5. Image Pre-processing
    - Multiple types of noise and abnormalities are common technical problems that affect medical imaging. Medical imaging may be tampered with if handled improperly.
    - You can take a number of steps to accomplish your objectives, including noise reduction, artefact removal, contrast adjusting collected photographs, and image quality adjusting after improper handling.
6. Security
    - HIPAA security regulations were created as a result of security lapses, hackings, phishing scams, and ransomware.
    - These provide businesses with direction for storing, transmitting, authenticating, and controlling access, integrity, and auditing procedures.
7. Meta-Data
    - The meta-data includes the date the data was created, its intended use, who is accountable for it, and how and when it was previously used by academics and data analysts.
    - This enables analysts to replicate earlier inquiries to facilitate scholarly investigation and precise benchmarking. By doing this, the data is made more usable and infrequently used "data dumpsters" are avoided.

8. Querying
    - The query tool might not have access to the whole data repository if there is improper compatibility between the datasets. The different parts of the dataset should also be correctly connected or linked and accessible. If not, it might be impossible to paint a whole picture of each patient's health.
    - You can query huge datasets and relational databases with SQL if you're satisfied with the precision, comprehensiveness, and uniformity of your data.
9. Visualization
    - Charts, heatmaps, and histograms that accurately display contrasting statistics and clearly label information make it much simpler to understand and effectively apply the information while also reducing the chance of confusion.
10. Data sharing
    - The largest obstacle to data sharing is thinking of data as a product that might provide you an edge over your competition. Therefore, information exchange across various EHR systems may be purposefully blocked by manufacturers as well as providers.

6: RECOMMENDATIONS

These few suggestions can help produce better results if big data must be included in patient care of critically ill:

- Both data scientists and clinicians must be familiar with the analytical methods employed by data scientists to do their jobs effectively. Clinicians and researchers must cooperate and preserve healthy skepticism up until that point. Interdisciplinary communication is made possible and necessary by this framework.
- Creation of a clinical informatics system that uses crowdsourcing to launch, enhance, and expand evidence-based and clinical decision-support systems. Knowledge creation becomes routine and thoroughly integrated into clinical workflows in this collaborative medical culture.
- Long ago, healthcare should have adopted the use of individual data to benefit the population and population data to benefit the individuals, but due to issues with privacy and security, data ownership, etc., this has not yet happened. The identification and expansion of best practices through this kind of work can significantly improve human health and the efficient use of medical resources.
- Institutions and researchers must develop strategies and procedures that handle some of our most urgent issues, such as who manages shared data, who owns clinical data, and how to combine disparate and outwardly incompatible data sources. Big data needs to be used differently than traditional data utilization, i.e., cooperatively. The big data hoopla can be transformed into valuable insights by fostering a culture of openness and reproducibility.

With bigger clinical datasets, more readily accessible genetic databases, connections to data outside the healthcare sector, and active collaboration with informatics professionals and data scientists, big data is becoming more prevalent in critical care research and clinical practice. Big data's time in healthcare is just getting started.

7: CONCLUSION

We may be entering a new era of critical care, one in which clinical decision-making is increasingly supported by computers that integrate and analyze data. To customize treatment and navigate crucial care disciplines that enhance population health, doctors must embrace the power of data. To fully comprehend all the usage data that is frequently gathered during the course of care, without jeopardizing the patient's right to privacy, collaboration with other non-health care experts is crucial in the field of data science.

Clinicians and systems driven by data must collaborate. Large volumes of data are evaluated by computer systems and made available to doctors working at the bedside as manageable, interpretable, and useful knowledge that improves clinician decision-making. Clinicians contextualize and oversee the implementation of the diagnostic and therapeutic recommendations made by predictive models. With a team-based and rigorously scientific approach, false alarms are reduced, and the system is continuously improved. Data science has the power to transform. There is still work to be done, but it is very likely that this situation will occur shortly.

Based on actual data, big data enables you to improve descriptive and predictive analysis. Clinical trials are still necessary, but this can make them better by better identifying the research topics and calculating sample sizes and event rates. It is crucial to secure data security, governance, and confidentiality as big data becomes more commonplace.

By personalizing healthcare and merging various data sources for every patient, big data has the potential to enhance healthcare and drastically lower expenses.

8: REFERENCES

1. Sanchez-Pinto, L. N., Luo, Y., & Churpek, M. M. (2018). Big Data and Data Science in Critical Care. *Chest*, *154*(5), 1239–1248. https://doi.org/10.1016/j.chest.2018.04.037
2. Organizing Committee of the Madrid 2017 Critical Care Datathon, Núñez Reiz, A., Martínez Sagasti, F., Álvarez González, M., Blesa Malpica, A., Martín Benítez, J. C., Nieto Cabrera, M., Del Pino Ramírez, Á., Gil Perdomo, J. M., Prada Alonso, J., Celi, L. A., Armengol de la Hoz, M. Á., Deliberato, R., Paik, K., Pollard, T., Raffa, J., Torres, F., Mayol, J., Chafer, J., González Ferrer, A., … Sánchez García, M. (2019). Big data and machine learning in critical care: Opportunities for collaborative research. *Medicina intensiva*, *43*(1), 52–57. https://doi.org/10.1016/j.medin.2018.06.002
3. https://journalofbigdata.springeropen.com/track/pdf/10.1186/s40537-019-0217-0.pdf
4. https://www.ons.org/pubs/article/236236/preview
5. Celi, L. A., Mark, R. G., Stone, D. J., & Montgomery, R. A. (2013). "Big data" in the intensive care unit. Closing the data loop. *American journal of respiratory and critical care medicine*, *187*(11), 1157–1160. https://doi.org/10.1164/rccm.201212-2311ED
6. Ghassemi, M., Celi, L. A., & Stone, D. J. (2015). State of the art review: the data revolution in critical care. *Critical care (London, England)*, *19*(1), 118. https://doi.org/10.1186/s13054-015-0801-4
7. Lee K. (2018). Critical Care Research Using "Big Data": A Reality in the Near Future. *Acute and critical care*, *33*(4), 269–270. https://doi.org/10.4266/acc.2018.00346
8. Docherty, A. B., & Lone, N. I. (2015). Exploiting big data for critical care research. *Current opinion in critical care*, *21*(5), 467–472. https://doi.org/10.1097/MCC.0000000000000228
9. Kim, H. S., Kim, D. J., & Yoon, K. H. (2019). Medical Big Data Is Not Yet Available: Why We Need Realism Rather than Exaggeration. *Endocrinology and metabolism (Seoul, Korea)*, *34*(4), 349–354. https://doi.org/10.3803/EnM.2019.34.4.349
10. Madanian, S., Parry, D. T., Airehrour, D., & Cherrington, M. (2019). mHealth and big-data integration: promises for healthcare system in India. *BMJ health & care informatics*, *26*(1), e100071. https://doi.org/10.1136/bmjhci-2019-100071

THANK YOU!