Date	**29.09.2022**
Team ID	**PNT2022MID47721**
Project Name	**Data Analytics For Hospital Health Care Data**
Maximum Mark	**4 marks**

A Survey on Big Data Analytics In Health Care

ABSTRACT

This paper gives a brief introduction about how we can uncover additional value from health information used in health care centers using a new information management approach called as big data analytics, including big data analytics in health sector provides stakeholders with new insights that have the potential to advance personalized care, improve patient outcomes and avoid unnecessary costs. This paper defines big data analytics and its characteristics, comments on its advantages and challenges in health care.

1. **INTRODUCTION:** The healthcare industry historically has generated large amounts of data, driven by record keeping, compliance & regulatory requirements, and patient care. While most data stored in hard copy form, the current trend is toward rapid digitization of these large amounts of data. Driven by mandatory requirements and the potential to improve the quality of healthcare delivery meanwhile reducing the costs, these massive quantities of data (known as 'big data') hold the promise of supporting a wide range of medical and healthcare functions, to derive previously untapped intelligence and insights from data to address many new and important questions. Within

the health sector, it provides stakeholders with new insights that have the potential to advance personalized care, improve patient outcomes and avoid unnecessary costs.

2. DEFINING BIG DATA.

Big data typically refers to the following types of data: • Traditional enterprise data – includes customer information from CRM systems, transactional ERP data, web store t transactions, and general ledger data.

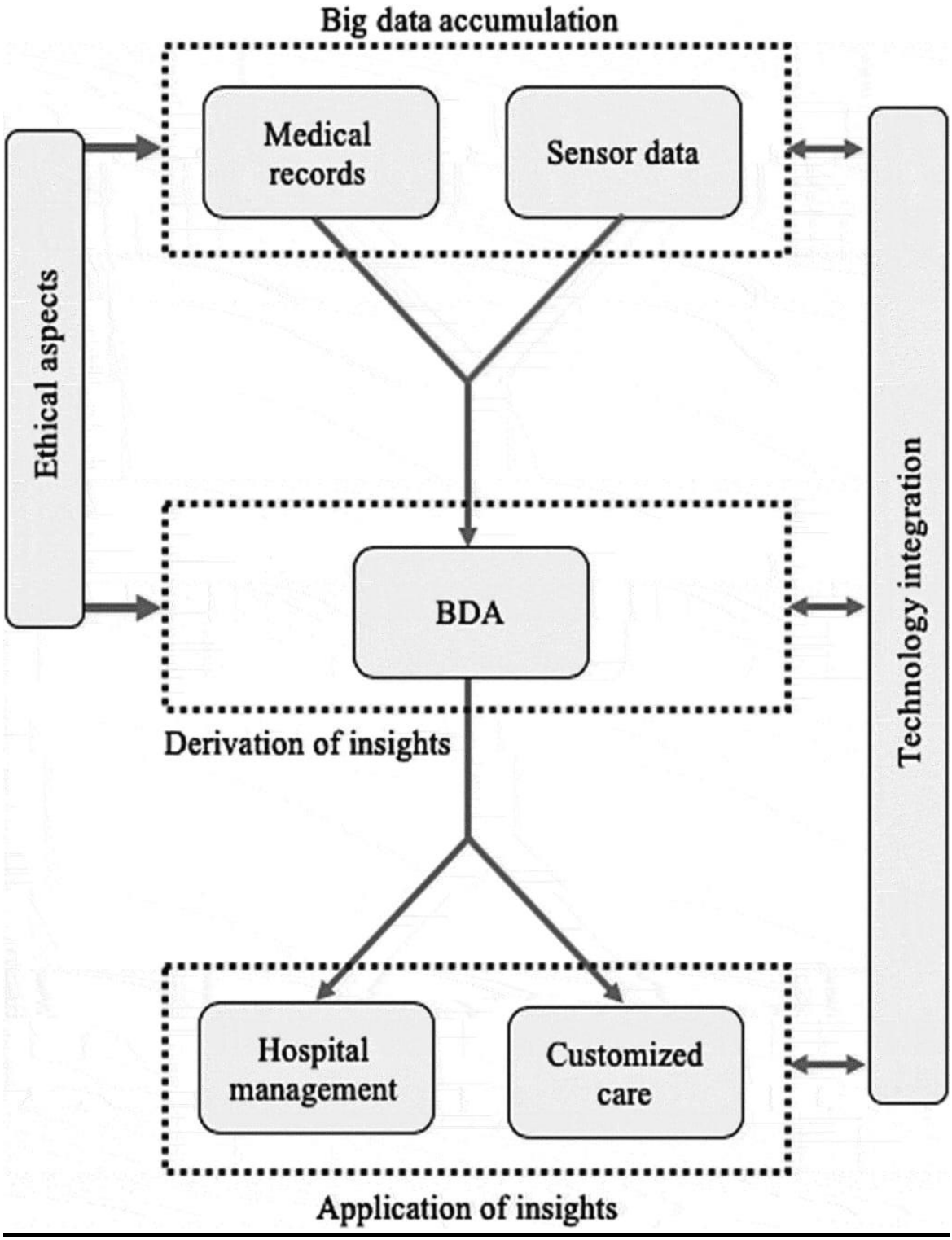
• **Machine-generated /sensor data** – includes Call Detail Records ("CDR"), weblogs, smart meters, manufacturing sensors, equipment logs (often referred to as digital exhaust), and trading systems data.

• **Social data** – includes customer feedback streams, micro-blogging sites like Twitter, social media platforms like Facebook In fact, there are four key characteristics that define big data:

• **Volume** is the amount of data generated by organizations or individuals. Enterprises in all industries are looking for ways to handle the Every increasing data volume that's being created every day.

• **Velocity** is the frequency and speed at which data is generated, captured and shared. Consumers as well as businesses now generate more data and in much shorter cycles, from hours, minutes, seconds down to milliseconds.

• **Variety** is the proliferation of new data types including those from social, machine and mobile sources. New types include content, location or geo-spatial, hardware data points, log data, machine data, metrics, mobile, physical data.


Big data accumulation
Medical records
Sensor data
Ethical aspects
BDA
Technology integration
Derivation of insights
Hospital management
Customized care
Application of insights

IMPLEMENTATION PROGRESS OF BDA WITHIN THE HEALTH CARE SECTORS.

Process: Disease outbreak BDA used for monitoring of disease networking. An example is Google. Org's use of BDA to study the timing and location of search engine queries to predict disease outbreaks. Research shows that one-third of consumers currently use social networking for health care purposes (Facebook, YouTube, blogs, Google, Twitter). As demands for access to health information from social networking sites continue to proliferate, BDA can potentially support key prevention programs such as disease surveillance and outbreak management.

Process: "HERITAGE HEALTH PRICE" Identify patients who will be admitted to a hospital within the next year using historical claims data. Over $30 billion was spent on unnecessary hospital admissions.

EDA PRESPECTIVES:

"Exploratory data analysis' is an attitude, a state of flexibility, a willingness to look for those things that we believe are not there, as well as those we believe to be there."

FUNDAMENTAL LOGICAL TOOLS IN EDA

```
Import pandas as pd. ( logic sequence)
Import num.py as pd
Import math plot lib.py
%num.py line formation
Import teras #py. libraries
```

CHARACTERSTICS OF BIG DATA:

1. Volume (size)

A large amount of data is a primary characteristic of big data.

2. Variety (complexity)

Big data includes structured, semi-structured, and unstructured data in different formats, such as text, image, audio, video, and sensor data, among others.

3. Velocity (speed)

Big data handles high rates of data inflow and processes the data in real-time.

4. Veracity (quality)

Big data accumulates detailed data that is exhaustive in scope.

5. Value (knowledge)

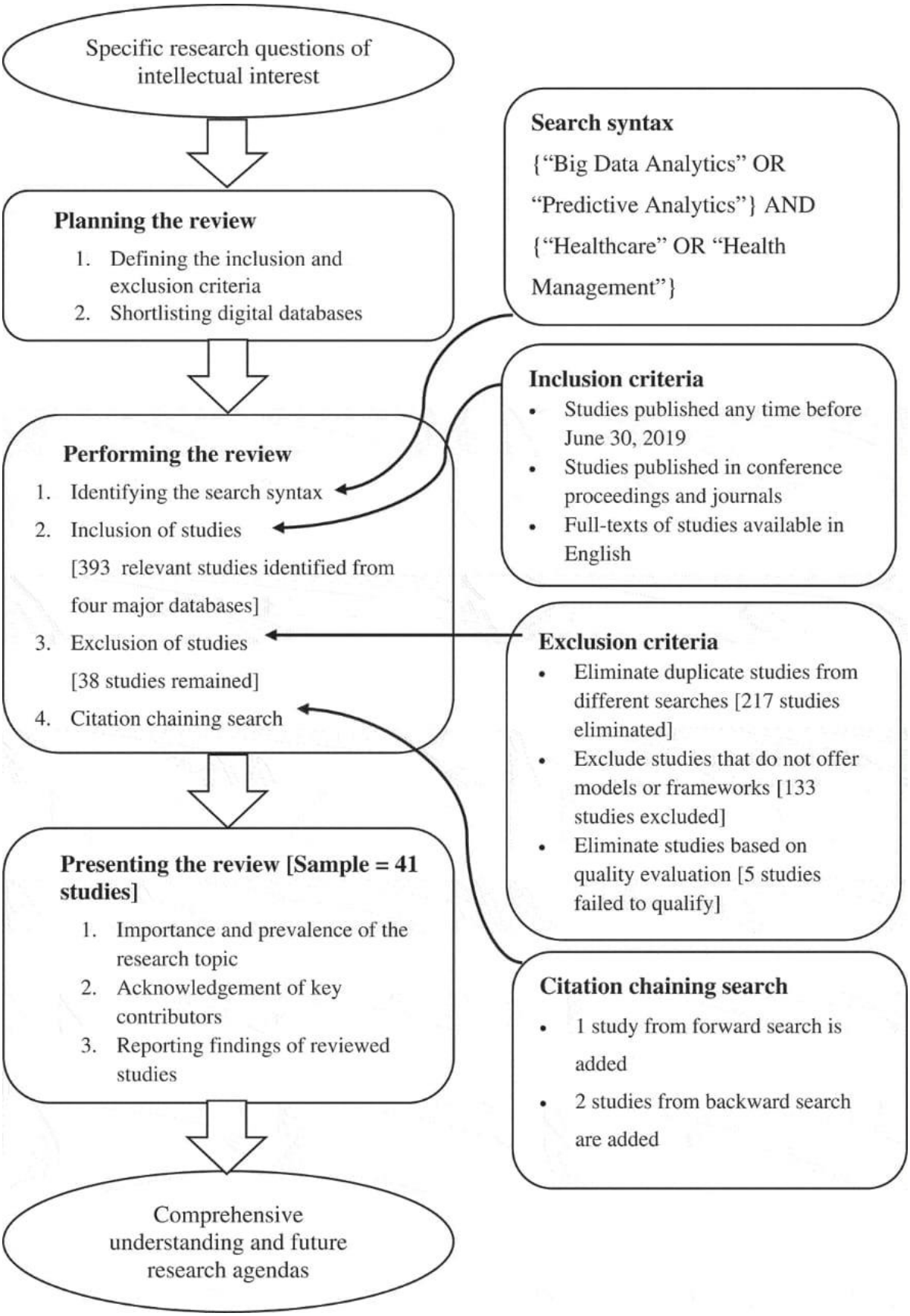
Big data offers in-depth information about a topic of discussion.

6. Variability (flexibility)

Big data provides support for the constantly changing nature of data by offering extensionality (the addition of new data fields) and scalability (expansion in size).

7. Valence (connectedness)

Big data connects common fields to conjoin different data sets.


Specific research questions of intellectual interest
Planning the review
1. Defining the inclusion and exclusion criteria
2. Shortlisting digital databases
Performing the review
1. Identifying the search syntax
2. Inclusion of studies [393 relevant studies identified from four major databases]
3. Exclusion of studies [38 studies remained]
4. Citation chaining search
Presenting the review [Sample = 41 studies]
1. Importance and prevalence of the research topic
2. Acknowledgement of key contributors
3. Reporting findings of reviewed studies
Comprehensive understanding and future research agendas
Search syntax
{"Big Data Analytics" OR "Predictive Analytics"} AND {"Healthcare" OR "Health Management"}
Inclusion criteria
• Studies published any time before June 30, 2019
• Studies published in conference proceedings and journals
• Full-texts of studies available in English
Exclusion criteria
• Eliminate duplicate studies from different searches [217 studies eliminated]
• Exclude studies that do not offer models or frameworks [133 studies excluded]
• Eliminate studies based on quality evaluation [5 studies failed to qualify]
Citation chaining search
• 1 study from forward search is added
• 2 studies from backward search are added

What Exactly Does A Healthcare Data Analyst Do?

Most reports that a Healthcare Data Analyst prepares will fall into one of two categories:

- **Financial Analysis**
- **Clinical Analysis**

Some reports will fall into both categories.

For this post, I'll concentrate on work a Healthcare Data Analyst does inside a hospital. This is by no means an exhaustive list, but here are some examples of both categories of analysis.

FINANCIAL ANALYSIS:

Hospitals concentrate heavily on providing high-quality patient care and that is priority 1, but at the end of the day a hospital is a business. And, just like all other businesses, hospitals have to remain financially stable in order to continue to keep the doors open and offer care to patients.

One of the most critical areas of the analyst's work is monitoring the financial outcomes in a myriad of ways. Here are some examples of financially-based analytics

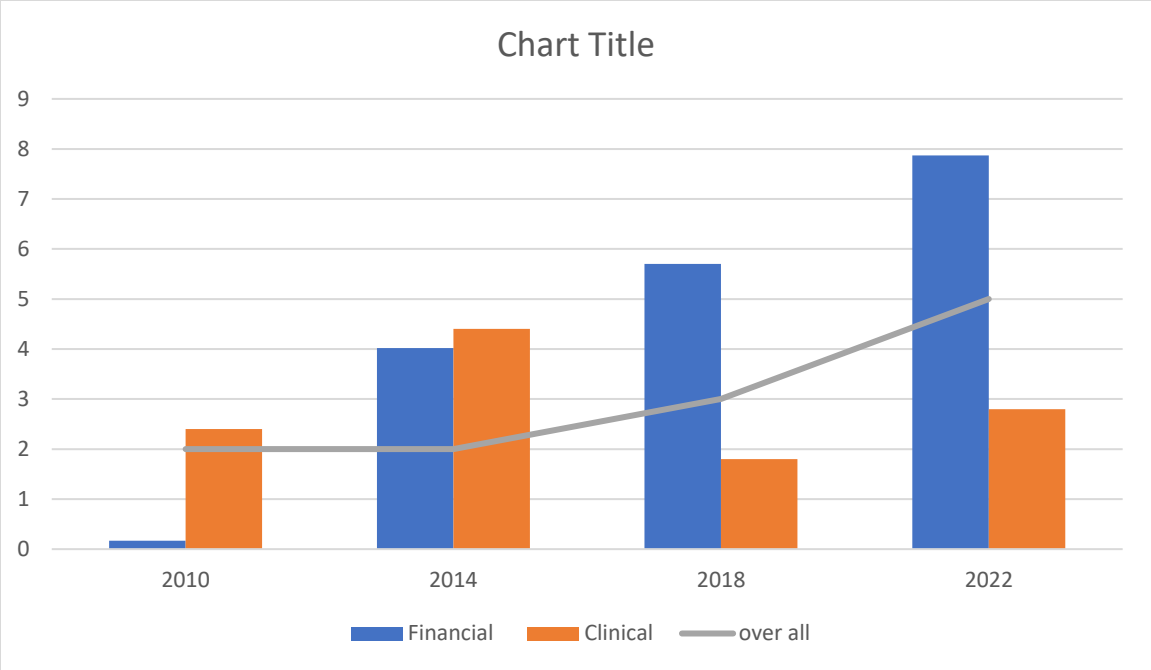
LENGTH-OF-STAY REPORTING

Length-of-stay (LOS) refers to the number of days a patient is in the hospital. If a patient is in the hospital for 5 days, that's reported as a "5 Day LOS".

Clinical analytics has emerged as a significant area of focus for IT leaders among healthcare providers, who are moving toward adopting accountable care. Many of the leaders have said that they are prioritizing clinical analytics over other systems in their organizations.

CLINICAL ANALYSIS:

Clinical analytics is a field that makes use of real-time medical data to generate insights, take decisions, increase revenues, and save on costs. The implementation of clinical analytics in organizations has led to reduced medication errors, improved population health, and cost savings for many organizations. The rapid advancements in key technologies and adoption of electronic health records (EHRs) has led to the growth of clinical analytics in recent years



CHALLENGES:

Big data analytics is only in the early stages of growth, and existing technologies and approaches are incapable of addressing the issues associated with big data. Big data may be regarded as large, challenging structures. Much analysis would also be necessary in this area to resolve the healthcare system's problems.

Data management, security and privacy issues

Issues such as data integrity and privacy lead to poor data management Privacy violation and discrimination. Disclosure of Personal Health Information is also a major risk.

Technological Issues

Lack of required infrastructure cannot produce safe conclusions Socio inequality, as data are only open to a small elite of technical specialists who know how to interpret and use it, and to those who can employ them.

Skilled Resource set

There is a need to have a Data scientist and Data analyst to perform big data analysis. There is already a huge shortage in the required skill set for Big Data Analytics.

Data Ownership

There is a lot of big data flowing which includes genomics, remote sensing, social media, mobile app and many other data types.

Healthcare Models

There is a need to have sufficient business case evidence in health to measure investment return.

Limited awareness and support

It will cause lack of funding and awareness. Dependency on private funding will support a few big players that will further lead to international economic competitiveness. Funding models have to be revisited to ensure better care.

CONCLUSION AND FUTURE SCOPE

The current study intended to address four research questions related to the application of BDA in healthcare. These questions have been answered following a standard protocol for reviewing resources from key databases. The prior literature on the application of BDA in healthcare has focused on five main themes, namely health awareness, stakeholders of the healthcare ecosystem, hospital management practices, specific medical conditions, and healthcare service delivery through technology use. The study has identified the gaps in the existing literature and provided an actionable research agenda for future research on the utilization of big data in the healthcare sector. However, despite the significant contributions of this current study, it suffers from three main limitations: first, book chapters, magazine articles, and thesis studies have been excluded from the scope of this study; second, journal articles and conference studies not available in English were not considered; third, studies not available in the four databases were not reviewed unless they appeared in the forward and backward searches. Future research is invited to overcome these limitations. Also, we recommend that scholars study the application of BDA in services provided by, for example, banking and financial institutions, media and broadcast channels, and the travel and hospitality industry by adopting the protocol followed in the current study. Similarly, the application of new technologies, such as blockchain, cloud computing, and machine learning, in healthcare provides promising avenues of exploration. We conclude this SLR with a call for theory development regarding the specific applications of BDA and the general integration of technology in the healthcare sector.

REFERENCE:

https://www.tandfonline.com/doi/full/10.1080/17517575.2020.1812005

Big data analytics in healthcare: a systematic literature review

Sayan tan Khanra ORCID Icon, Amandeep Dhir ORCID Icon A. K. M. Najmul Islam &Matti Manty maki ORCID Icon

https://www.researchgate.net/publication/334262484_A_review_of_the_literature_on_big_data_analytics_in_healthcare

A review of the literature on big data analytics in healthcare

July 2019Journal of the Operational Research Society 71(1):1-19 Follow journal

DOI: 10.1080/01605682.2019.1630328

https://www.emerald.com/insight/content/doi/10.1108/MD-01-2020-0035/full/html

Big data analysis for decision-making processes: challenges and opportunities for the management of health-care organizations

Simone Fanelli, Lorenzo Pratici, Fiorella Pia Salvatore, Chiara Carolina Donelli, Antonello Zangrandi

https://www.ijert.org/big-data-analytics-in-healthcare-a-review

Authors: Prableen Kaur

Paper ID: IJERTV10IS060198

Volume & Issue: Volume 10, Issue 06 (June 2021)

Published (First Online): 23-06-2021

ISSN (Online): 2278-0181

Publisher Name: IJERT