Investigating the Adoption of Big Data Analytics in Healthcare: The Moderating Role of Behavior Towards Change

Shahid Umar

Registration# PIDE2018-FMPHILPP26

MPhil Public Policy

Supervised By:

Supervisor: Dr. Attiya Yasmin Javid

&

Co- Supervisor: Dr. Fazli Hakim Khattak

School of Public Policy,

Pakistan Institute of Development Economics, PIDE Islamabad

Pakistan Institute of Development Economics, Islamabad
PIDE School of Public Policy

<u>CERTIFICATE</u>

This is to certify that this thesis entitled: **"Investigating the Adoption of Big Data Analytics in Healthcare: The Moderating Role of Behavior towards Change"** submitted by **Mr. Shahid Umar** accepted in its present form by the School of Public Policy, Pakistan Institute of Development Economics (PIDE), Islamabad as satisfying the requirements for partial fulfillment of the degree in Master of Philosophy in Public Policy.

Supervisor:

Dr. Attiya Y. Javid,
Professor,
Department of Economics,
Pakistan Institute of Development Economics,
(PIDE) Islamabad.

Co- Supervisor:

Dr. Fazli Hakim Khattak,
National Consultant, National AIDS Spending
Assessment (NASA) UNAIDS Islamabad.

External Examiner:

Dr. Thair Mahmood
Associate Professor,
School of Economics, Quaid-e-Azam University,
(QAU) Islamabad

Head,
PIDE School of Public Policy:

Dr. Iftikhar Ahmad,
Assistant Professor/HOD
Pakistan Institute of Development Economics,
(PIDE) Islamabad

DECLARATION

I, Shahid Umar hereby states that my M.Phil. thesis titled "Investigating the Adoption of Big Data Analytics in Healthcare: The Moderating Role of Behavior Towards Change" is my work and has not been submitted previously by me for taking any degree from Pakistan Institute of Development Economics, Islamabad. At any time if my statement is found to be incorrect even after my Graduation the university has the right to withdraw my M.Phil. degree.

Shahid Umar

Date: _____

DEDICATION

To my beloved grandfather whose love for me knew no bounds and, who taught me the value of hard work. Thank you so much "baba", I will never forget you

TABLE OF CONTENTS

ACKNOWLEDGMENTS

LIST OF TABLES

TABLE OF FIGURES

LIST OF ACRONYMS AND ABBREVIATIONS

ADB	Asian Development Bank
BDA	Big Data Analytics
CPEC	China Pakistan Economic Corridor
DEG	Digital-era Governance
DHIS	District Health Information System
EHR	Electronic Health Records
EMHR	Electronic Medical Health Records
EU	European Union
FDI	Foreign Direct Investment
GIZ	Deutsche Gesellschaft für Internationale Zusammenarbeit
HDFS	Hadoop Distributed File System
HDC	Health Data Cooperative
HIV	Human Immunodeficiency Virus
IT	Information Technology
ICT	Information Communication Technology
IDC	International Data Corporation
KFW	Kreditanstalt für Wiederaufbau
KPK	Khyber Pakhtunkhwa
MTI	Medical Teaching Institute
MRI	Magnetic Resonance Imaging
MoITT	Ministry of Information Technology and Telecommunication
NIH	National Health Institute
NADRA	National Database Registration Authority
OECD	Organization for Economic Cooperation and Development
R&D	Research and Development
RTC	Resistance to Change
RAHS	Risk Assessment and Horizon Scanning
SDGs	Sustainable Development Goals
UK	United Kingdom
USA	United States of America
WHO	World Health Organization

ABSTRACT

The purpose of this research study was to identify hurdles and challenges pegged in the way of inclusive big data analytics and big data health policymaking. The study aimed to propose a data-driven model for big data integration into the health policymaking process. For this purpose, the author adopted the usage of both primary and secondary data through in-depth interviews and content analysis (both quantitative and qualitative analysis). The empirical results showed that there are several hurdles and challenges when it comes to big data integration in the healthcare and health policy-making process such as resistance towards change, data privacy and security, interoperability, expertise and infrastructure, management and governance, data quality and reliability, financial constraints, lack of capacity, government priorities and political will. Based on the empirical findings, the Health Data Cooperative (HDC) model was proposed, which is a key element and main component for public health policymaking and for leveraging the potential of rapidly emerging and all the existing data sources. In order to exploit all the data needs to overcome the technological algorithmic, computational challenges that define the present highly heterogeneous data environment, along with heaps of complex normative, regulatory, governance and policy constraints. Consequently, HDC offers a roadmap to address technological, political and social hurdles and empower healthcare and health policymaking by incorporating and implementing big data. Key recommendations and solutions concerning big data integration in the healthcare and health policymaking process was presented at the end of this study.

Key words: *Big data, Big Data Analytics, Healthcare, Resistance Towards Change, Health Policymaking, Health Data Cooperative.*

CHAPTER 1

INTRODUCTION

1.1 Background of the Study

Over the past few years, big data has been a promising trend in the Information Technology (IT) sector. It is a very popular term these days and experts are even considering it an era of big data (Boyd & Crawford, 2012). Large amount of data is produced from different sources such as, location and sensors data from mobile phones and other consumer electronic devices, internet data from web searches and social media data from services like Twitter and Facebook, etc. (Fig. 1). In 2018, the International Data Corporation (IDC) reported that the volume of digital content will grow from 33 zettabytes (ZB = 10^{21} bytes) this year to 175 zettabytes by 2025 (Patrizio, 2018). This explosion and exponential growth of data accompanied with the declining price of computer hardware from the commodification of IT lead to the development of technology that analyze and process big data.
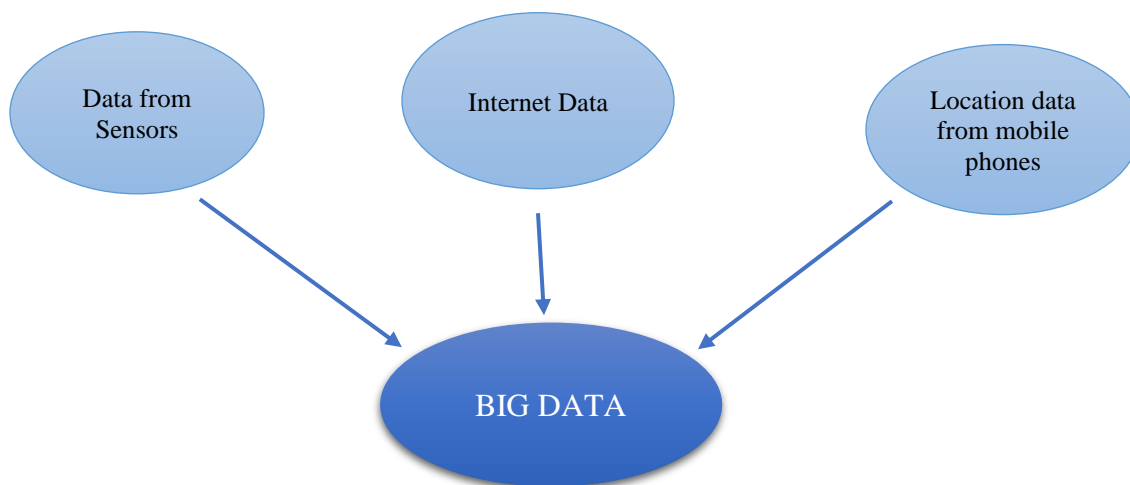
Figure 1: Big data sources (Robin, 2015)

1.1.1 Definitions and Properties

Nonetheless big data has been widely noticed and appreciated, still there is no uniform and universally agreed definition of big data. Undoubtedly, the lack of one uniform definition indicates the interesting landscape of its meaning and operation. In healthcare big data in evolving not only because of its size and volume but also the variety of data types and the speed at which it could be handled. The succeeding definitions would provide a better view of what the concept of big data is. Substantively, the concept of big data has been defined itself since 2001. Certainly, the size is the main characteristic of big data that comes to mind whenever we think about it, but it has some additional properties as well. For instance, a widely used definition proposed by Gartner is the 3V's Model in 2001. It talks about the increase in data's volume, variety and velocity. Hadoop and Mckinsey provided a rather similar definitions of the concept. Apache Hadoop (2010) defined big data as "datasets which could not be captured, managed and processed by general computers within an acceptable scope". According to Mckinsey the term "big data refers to datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze" (Mckinsey, 2011). The concept was redefined by Gartner in 2012 as, "big data is high volume, high velocity and high variety information assets that require new form of processing to enable enhanced decision making, insight discovery and process optimization" (Beyer, 2012).

1.1.2 Big data analytics

Analytics is defined as the process of examining historical data to uncover unknown correlations, hidden patterns and research potential trends. With the help of analytics data can

generate value for the world economy, driving innovation, improving productivity, growth and efficiency (See Figure 2). Big data in this regard enhances existing analytics.
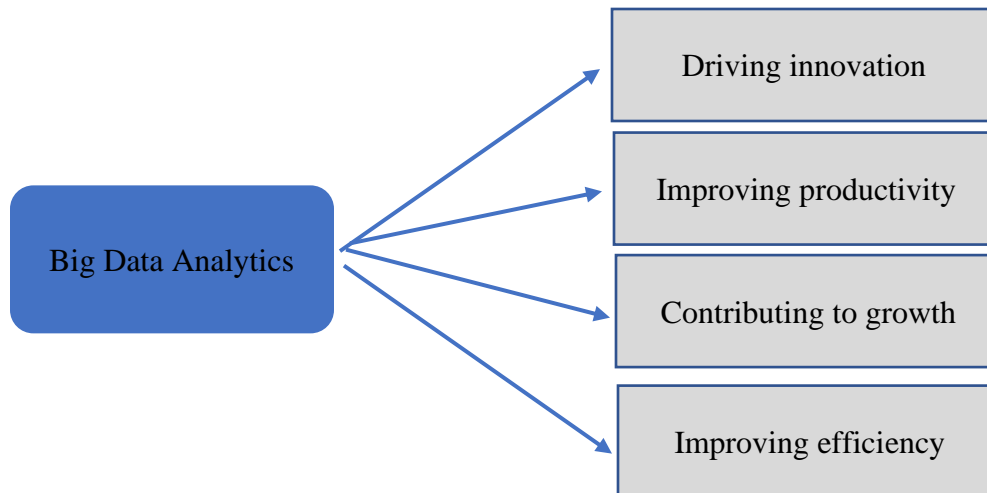
Figure 2: Values generated from big data analytics (Sippe, 2015)

A definitive and real-world example of big data analytics is the Google's Flu Trends. Back in 2008, Google introduced Google Flu Trends. It is a web service that "precisely estimate the current level of weekly influenza activity in each region of the United States" (Richterich, 2016). Google quantified that they found a close relation between the number of people who search for flu-related topics and how many people actually are affected by flu. Today, Google Flu Trends can estimate influenza in eighteen countries through the aggregation of search queries.

1.1.3 Big data in healthcare

Like any other sector, health sector also generates large amount of data on a daily basis by maintaining records and patient care. This clinical data is stored by the healthcare providers in their electronic medical health records (EMHR). In addition to EMHR, there are three other main sources of data in the healthcare sector, which include: pharmaceutical R&D data, data of costs

and claims of service that was provided and patient sentiment and behavior data which describes patient preferences and activities (Figure 3).
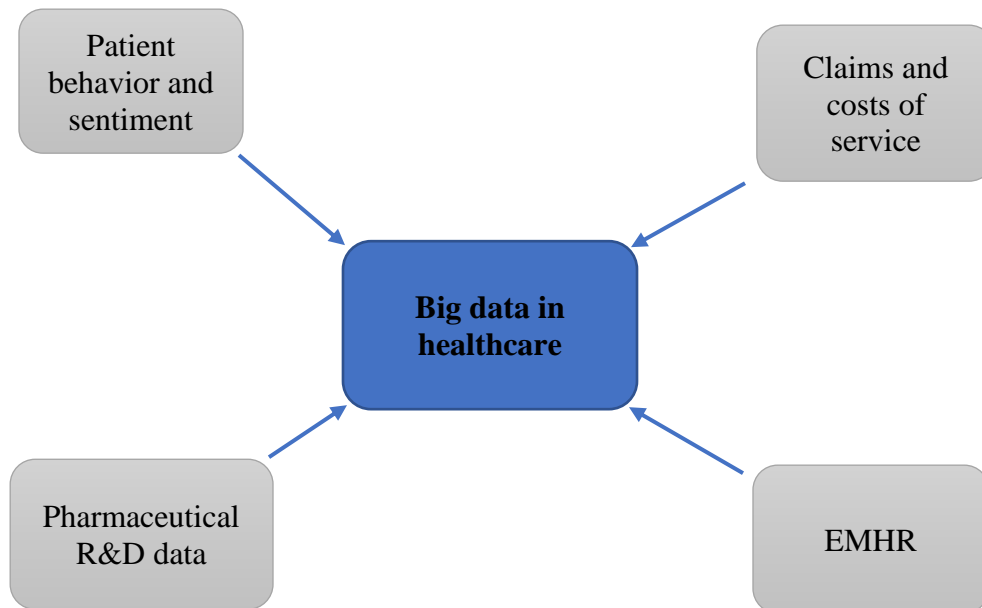
Figure 3: Sources of big data in healthcare (Groves et al., 2016)

Similarly, big data is also available in shape of doctor's written prescriptions and notes, medial images such as MRI and CT scans outcomes, drugstore documents, laboratory records, insurance files and other administrative data, medical journals and social media data from tweets and updates on webpages. By understanding trends and stencils within data, big data analytics appears to save lives, avoids unnecessary costs and improves patient outcomes. By description, "big data in healthcare refers to electronic health datasets so large and complex that they are difficult to manage with traditional software, hardware, data management tools and methods" (Ragupathi & Ragupathi, 2014). Those digitized data also hold the promise of supporting a wide range of medical and healthcare functions. Consequently, big data analytics application in healthcare take benefit of extracting insights from data for better and improved decision-making

purpose. Analytics of big data is the method of examining huge amount of data, from diverse data sources and in numerous formats, to deliver insights that can enable decision making in real-time. Various analytical notions such as artificial intelligence and data mining can be applied to analyze the data.

Big data analytics have shown to have the ability to support healthcare. Inside the huge amount of health data, trends and patterns are hidden, which with the right big data analytics tools has the potential to lower costs, save lives and improve overall health care services. Big data can be used to improve medical functions and healthcare services, including among others clinical decision support, personal and population health and disease surveillance (Raghupathi & Raghupathi, 2014). McKinsey is of the view that big data can improve quality and efficiency in the following areas: Public Health, Clinical operations, Research & Development (Manyika et al., 2011).

Big data analytics has altered the way we analyze, leverage and manage data in any sector. One of the most encouraging policy areas where it can be applied is the health policies. The heath sector analytics have a huge potential to avoid preventable diseases, predict outbreaks of epidemics, reduce cost of treatments and in general improving the quality of life for common citizens. The country's population is increasing along with the average human lifespan, which postures new challenges to the current treatment delivery procedures. Consequently, the researcher aims to incorporate healthcare sector analytics in public health policymaking process, and subsequently analyze its associated hurdles and challenges.

On the other hand, there is dire need to improve the health of individuals and the population and to shift from proactive or reactive towards predictive and prescriptive decision-making.

Therefore, there is urgent need of health policies which are backed by real-time data in order to meet the increasing healthcare demands of common citizens. One of the technical solutions the world is applying to address health issues is the use of big data analytics. Whereas in KPK, there is little to no adaptation of big data in healthcare and big data oriented public health policymaking.

1.2 Problem Statement

In the field of healthcare, several terms related to the collection, analysis and application of big data float around, comprising health information technology, digital health, health 2.0, e-health etc. (Forschungs, 2017). These varied terms are exceeded only with the help of heterogenous data sources that are joined to enhance health care services, for instance, electronic health records, medical test results and publications. Conversely, for instance, these data - stored in biobanks databases and registries are a highly valuable resource for current and future public health policies. Both individual and aggregate data can promote the analysis of patient health outcomes and disease paths as basis for improving personal health and treatment (Leyens et al., 2017). Contrariwise, the main aim of public health policymaking is to effectively predict and evaluate risk, foresee possible unintended outcomes and implicitly allocating scarce resources. Therefore, the key to effective and successful public health policymaking is to depend on timely and relevant data and evidence to assist in developing collaborative interventions.

The private sector in developing countries is on a fast track when it comes to the adoption of data strategies for decision-making purposes, whereas governmental organizations are more slowly embracing the newest instruments and technologies. Similarly, developed countries such as, US, UK and Singapore have worked extensively in managing healthcare big data and utilizing it for decision making (Song, 2012; Kayyali et al., 2013; Lawrence & Bradley, 2018). Although,

majority of the governmental organizations in developing countries are not ready to use big data analytics for their strategic advantage (Bhatt et al., 2017), in other words, to redefine their actions and role, to refine their working way, to understand their disruptive character of technological advancements and to examine the topic as a whole and to identify advantages for policy definition, development and evaluation. One of the main reasons for the sluggish pace of big data integration in healthcare is the nature of big data itself: due to regulations and concerns regarding data protection and privacy, several players, facilities and data integration issues. This study therefore looked at the challenges pegged in the way of inclusive big data analytics in healthcare and public health policy-making.

1.3 Research Purpose

The intent of the proposed study was to identify hurdles and challenges pegged in the way of inclusive big data analytics in healthcare and health policy-making. The study aimed to devise an extensive conceptual framework for its integration in public health policy formulation in Pakistan. The study examined the challenges, prospects and policy implications that big data possess for data-driven and evidence-based health policymaking in Pakistan.

1.4 Significance of the Study

The study examined the challenges, prospects and policy implications that big data possess for data-driven and evidence-based health policymaking. It offered a roadmap in order to help the government and private health providers identify the key factors that affect the adoption of big data analytics in healthcare and public health policymaking. The study, therefore, proposed a much-needed model for better healthcare and health policymaking using big data analytics.

1.4 Research Objectives

- Understand the current knowledge of big data and its application in healthcare and health policy formulation.

- To help the government and private health providers identify the key factors that affect the adoption of big data analytics in healthcare and public health policymaking.

- Evaluate the current health policy formulation process.

- Develop Conceptual Framework on the subject.

- Draw recommendations for data-driven and evidence-based public health policymaking in Pakistan.

1.5 Research Questions

The following research questions will be addressed in this study:

- What are the major hurdles and challenges pegged in the way of inclusive big data analytics and big data health policy making?

- What is the current status of big data and its application in the health sector of national and international levels?

- Understand the role of resistance towards change in adoption of BDA in healthcare and public health policy-making

- What are the implications for incorporating big data analytics in public health policy formulation in future?

1.6 Motivation for the Study

The key motivation behind this study was to find a practical solution to the growing healthcare issues in Pakistan. The researcher aimed that the relatively new concept of big data has a huge potential in health sector, with the help of which Pakistan can address the growing demand for health facilities with the effective use of data and evidence in the decision-making process. The researcher also desires to do a research which possess practical application and to be of service to the society.

1.7 Organization of the Study

The following are the chapters included in the study:

Chapter 1: Introduction chapter includes a brief background, research purpose, research objectives, research questions and motivation of the study.

Chapter 2: Literature Review chapter critically and extensively evaluate the literature to derive concepts for the conceptual framework.

Chapter 3: Conceptual Framework chapter briefly elucidate on the concepts that are derived from the literature.

Chapter 4: Research Methodology portion displays the foundation for the application of specific methods or processes used to recognize, select and securitize data and facts smeared to understand the research problem.

Chapter 5: Data Analysis and Findings chapter presents the research analysis and findings of the data obtained from library sources and face-to-face/telephonic interviews.

Chapter 6: Conclusion and Recommendations chapter summarizes the research findings, the outcomes and accomplishment of objectives and also the recommendations, research limitations and direction for future research.

CHAPTER 2

LITERATURE REVIEW

There is a lot of research work conducted on the concept of big data since it is introduced in this fast-changing technological world. Many researchers have worked for its application in different fields, such as business, education, transportation, banking, finance and marketing. Similarly, plenty of studies have been conducted to support big data uses in health sector. The studies show an enormous potential of big data inclusion in healthcare, and how that will revolutionize the healthcare system. The developed world and fast-growing economies like China and South Korea have made great progress with respect to big data uses in different fields. Therefore, this literature review aims to critically analyze and evaluate the concept of big data, analytics and its uses in the healthcare sector, and also to identify the opportunities and challenges associated with big data analytics for public health policies.

2.1 Definitions and Properties

There is no uniform and universally agreed definition of big data. Undoubtedly, the lack of one uniform definition indicates the interesting landscape of its meaning and operation. Many definitions have been widely used. According to Mckinsey the term "big data refers to datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze" (Mckinsey, 2011). One more widely used definition of big data is the "3V" model proposed by Gartner: "big data is high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making" (Beyer, 2012). The 3V Model is illustrated in detail below in Table 1:

Characteristics	Description	Attribute	Driver
Volume	The sheer amount of data generated or data intensity that must be ingested, analyzed, and managed to make decisions based on complete data analysis.	The digital universe is generating a high volume of data, which is expected to increase with exponential growth.	Increase in data sources, higher resolution sensors.
Velocity	How fast data is being produced, changed and the speed with which data must be received, understood and processed	Metrics used can be defined in the segments of accessibility, applicable and time value.	Increase in data sources, improve throughput connectivity and enhanced computing power of data generating devices.
Variety	The rise of information coming from new sources both inside and outside the walls of the enterprise or organization creates integration, management, governance, and architectural pressures in IT.	The data can be divided in the following segments: structured, unstructured, semi structured, and complexity.	Mobile, social media, videos, chat, genomics, and sensors.
Veracity	The quality and provenance of received data.	The quality of Bid Data may be good, bad, or undefined due to data inconsistency, incompleteness, ambiguities, and model approximations.	Data-based decisions require traceability and justification.

Table 1: Characteristics of big data (Wilson & Kerber, 2011)

2.2 Challenges in Big Data Integration

The review of the literature shows that there are few studies available that described that importance of big data Analytics in health policymaking (Groves & Knott, 2013), nevertheless, studies have inspected the required qualifications and technological aspects for big data in public health policymaking (Braunstein, 2015; Kim & Park, 2017). The previous researches mainly focused on policy and technological issues and not on the adoption issues, such as trust, fitness and security of technology for the duties needed to handle big data analytics in heath sector (Kim & Park, 2017). The available literature attempts to explain big data analytics adoption through the view of technology, for instance its usefulness and -ease of use (Petersen, 1997; Rehman, 2016). Nevertheless, stressing only on the end-user perception of technology may not be enough. People will approve and embrace the system only when the features of the new technology are suitable for the job requirements (Goodhue & Thompson, 1995). Acceptance will also occur when people

sees the new technology as advanced, useful and easy to use, nevertheless the new technology may not be approved and welcomed if an incongruity occurs with her mandatory tasks and the new technology cannot improve her job performance (Junglas & Watson, 2008).

Hence, it is necessary for the users to have a perception that the technology is easy and useful, and also the characteristics of technology must match with the defined job tasks.

2.2.1 Data Privacy and Security

The literature exhibited that perceived trust (Wang & Lin, 2017) and perceived security of information (Fife & Orjuela, 2012) were the main hurdles for users to adopt to innovative data driven systems. Security is one of the principal causes for the slow pace adoption of big data analytics (Jain, Gyanchandani & Khare, 2016). Data privacy and trust on the other hand is a main concern in the big data analytics acceptance, and therefore organizations must ensure data privacy and yield more trust in big data analytics adoption (Heart, 2010). Previous studies by Shin, Sivarajah and Malaka also emphasized on the perceived trust and perceived security to be the major hurdles and challenges for BDA acceptance in health sector.

2.2.2 Management and Governance

As the amount of global digital information and health-related data grows, so does the number of people using and accessing this information. There must be assurance given through relevant laws so that personal health data shall be used appropriately. There is still skepticism with respects to "where the data goes to", "by whom it is used" and "for what purpose" in the complex legal environment around the world (Europa, 2014). Therefore, for proper management and governance of the data there is a need of relevant data protection regulations which needs to be implemented in order to protect the privacy of individuals.

2.2.3 Data Reliability

On the other hand, one of the major concerns related to big data is the reliability of the data. As stated by Narayanan et al. (2013) "Traditional measures of data reliability are more difficult to calculate for such data. Also, patient data can contain significant proportions of missing values that further complicate the calculation of reliability." Therefore, for improved quality of healthcare services there is a need of robust feedback mechanism that can analyze patient records on frequent manner.

2.2.4 Expertise and Infrastructure

Big data proposes immense opportunities for new insights to understand human systems, and for predicting nonlinearities and interactions among variables. Nonetheless, insufficient infrastructure, traditional data analytics, shortage of data experts (IBM, 2013), lack of trust in databases (Davenport & Patil, 2012) and associated expertise impede the development of innovative data management solutions.

2.2.5 Interoperability

Another issue in the health sector is the lack of standardization to facilitate interoperability, as data is usually fragmented, or produced in IT systems having discordant formats (Roney, 2012). Hospital services, clinical activities, administrative services, education and research are segregated, and in several organizations, individual silo maintains its own distinct informational infrastructure and organizational data. The lack of technology integration and cross-border coordination espouses to facilitate interoperability amid the elements of the big data value chain. Global standardization can play an important role in addressing such challenges.

2.2.6 Resistance to Change

In developing countries, the resistance to change from staff is also one of the main reasons which affects the incorporation of innovative systems in the organizations (Sharma et al., 2020). Prior literature regarding 'electronic health records' (EHRs) system adoption, Resistance to Change (RTC) from doctors was frequently mentioned as a main hurdle for system adoption (Bates, 2005). RTC is also one of the main reasons that slows down the acceptance of information system in the healthcare (Groves et al., 2013).

Regardless of the recognition of big data Analytics, inadequate experimental research has examined aspects which can affect its adoption in health sector (Yan et al., 2017). Empirical studies conducted by a healthcare organization in Pakistan identified big gap in the literature from both sides i.e. the knowledge of BDA and the adoption of it (Tunio, 2018).

2.3 Prospects of Big Data in Healthcare

Big data Analytics holds problem-resolving and value creating-opportunities. Developed and catch-up economies, such as the UK, the USA, Australia, Singapore, and even the EU and OECD countries are actively formulating strategic plans for big data management, highlighting its potential, and utilizing it as an instrument to resolve persistent problems or to reveal new prospects of value (Song, 2012; Reimsbach, 2013). Manyika and Gartner have illustrated that how the utilization of big data can create economic value, such as looking for new projects and predicting market changes (Manyika, 2011; Gartner, 2012). The Economist on the other hand projected that if big data is properly utilized it might find solutions to problems related to food, energy, environment and healthcare (The Economist, 2014).

Big data analytics has a huge potential to boost the health and social safety of citizens in near future. It has the ability to support important and complex social issues affecting welfare, social safety and healthcare, for instance, the management and prevention of infectious diseases, individual-centered collaborative care, disaster, terrorism and global risks (Koch, 2013). For instance, it is now possible to thwart chronic diseases through monitoring individual lifestyle or consult data, and also to sense signs of social issues and the need for services provision by monitoring the collected data through various sources. Moreover, policy makers can shape policy agendas through analyzing and monitoring the big data of social media sites (Song, 2012).

Figure 4: Workflow of Big Data Analytics (Dash et al., 2019)
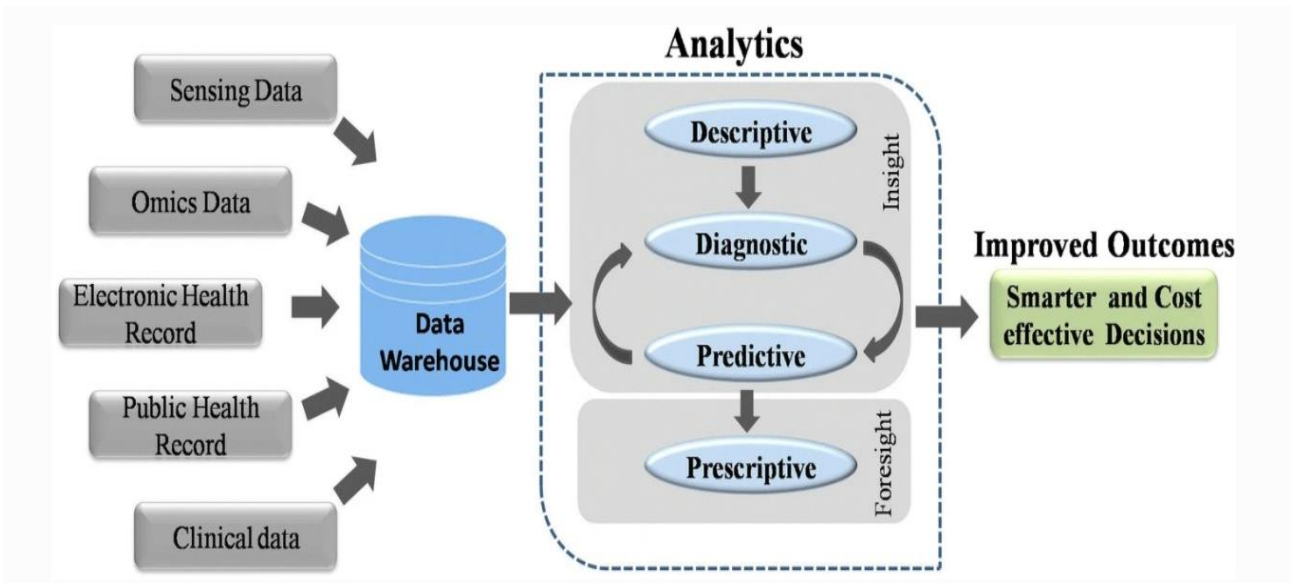
The workflow/process of big data Analytics. Data warehouses hold massive volumes of data that are produced from different sources. This data is analyzed with the help of computational/analytic pipelines for better and more affordable health care solutions.

There are plenty of examples from different countries who have successfully used big data analytics to solve their healthcare and social services problems. The Pillbox project in the United States resulting in reduction of $500 million annually in health provision cost with the help of Big

data analytics (Lee, 2012; Pillbox, 2013). The police department of San Francisco has designed a big data system which is actively used for crime prevention (Song, 2012).

The UK government has established a 'Foresight Horizon Scanning Centre' with the help of big data, which assists as an antidote to innumerable health and social issues, for instance epidemics, obesity and potential risk management (climate change, coastal erosion) (UK Government, 2012). The European Union has developed a project called iKnow (Interconnect Knowledge) that deals with uncertainty, and offer opportunities for research on terrorism, tsunamis, networking, earthquake and other global crises (Go & Jeoung, 2012). The 'Organization for Economic Cooperation and Development' (OECD) has adopted "Evaluating Economic Benefits of big data as an agenda for its 15th Working Party on Indicators for the Information Society (WPIIS)" by recommending big data use for increasing efficiency of different sectors (OECD, 2014).

Similarly, in order to prepare for future uncertainty with respect to diseases outbreaks and terrorism the government of Singapore developed a 'Risk Assessment and Horizon Scanning (RAHS) system (RAHS, 2012). In the same way the 'Australian Government Information Management Office' has saved resources and time by establishing an automated and computerized system that can search, analyze and reuse huge amount of data through government 2.0 (NISCK, 2012).

In South Korea, the 'National Information Society Agency' (NIA) has shown potential for the prevention of suicide with the help of analyzing online buzzwords. Proceeding this, a study conducted in 2012 discovered a relationship among a suicide rate and the quantity of searching 'suicide' through evaluating Google search trends (Song, 2012). The above examples identify and

concludes that there is a huge potential of big data in the healthcare and public health policymaking.

2.4 Public Health Policymaking

Policymaking is defined as, "a process that defines and pursues the right course of action in a given context, at a particular time, for a certain group of people, with a particular allocation of resources" (Greenhalgh & Russell, 2006). In a hypothetical scenario, to see how public health policy and its procedures can establish paths for its results: If pharmaceutical companies or patients are accessed for HIV policy in a country, the policy will be remedial in nature. On the other hand, if only policy economists are referred, prevention will be stressed, because preventive measures are often cost-efficient. The progression an HIV outbreak takes in that country will mainly rely on who took part and facilitated in shaping the HIV policy (Buse et al., 2005).

Public health policy development is an intricate phenomenon (Bosch et al. 2012) in which the inclusion of data or evidence is a complex and complicated tread (Gilson & Raphaely, 2008; Greenhalgh & Sietse, 2011). The idea that policies can or should be formulated or backed entirely by evidence has been widely debated in the academic realm. Scholars have argued that the basic notion 'evidence-based policy making' is blemished as it proposes that there are only technical solutions available to what are fundamentally political problems (Greenhalg & Russel, 2006). Policies primarily tackles questions like 'what or how a society should look like'— that only an evidence cannot answer. In practice, a policy is comprised of various social customs along with diverse evidence grounds pertinent to each of those customs (Parkhurst, 2016).

There is a general agreement among many policy makers and in the research arena that a policy could be 'evidence-informed' rather than 'evidence-based', and simply implying the

incorporation of knowledge emerged from research is not enough for health policy making as an evidence, but in also incorporate information from several other sources (Smith & Joyce, 2012; Hawkin & Parkhurst, 2015). The concept of 'policy' is very broad in nature and includes, regulations, laws, judicial decrees, budget priorities, along with agency guidelines (Brownson et all., 2009). The engagement of scholars with all policy making developments hence is essential for including evidence into policies (El-Jardali et al., 2012). The advocates of public health can play better role in policymaking if they grasp the processes of policy formulation and its implementation, along with several other factors which effects policy design, such as, institutions, individuals, power, politics, interest groups and ideas (Walt & Gilson, 1994; Buse & Dickinson, 2007).

According to World Bank (2016), Pakistan is lower middle-income country. Parallel to many other countries in this group, her policy settings are also different from the ones in high-income countries (Walt et al., 2008). The lower middle-income states usually share similar attributes such as, weak regulatory capacity, weaker regulations, favoritism in their political system, lack of purchasing power and more dependence on donor's support as compare to high-income states. Similar to many other developing states, the health budget in Pakistan is largely spent on hospitals which provides treatment services to the urban elite of the country, while the poor and downtrodden segments on the other hand die from diseases which can easily be obstructed for a few rupees (Filmer et al., 2000). The state controls the health sector and is both the purchases as well as provider of health facilities in the country, and also do public private partnerships while being in charge of its regulations simultaneously (Nishtar et al., 2013).

In Pakistan, very little emphasis has been given to policy development by research. The review of the literature displays a paucity of information overall. A small number of studies exist

on policy content (Siddiqi et al., 2004) or policy context (Khan, 1996; Collins et al., 2002; Khan & Heuvel, 2007). The absence of information backed by research is not new to Pakistan. Studies conducted in other places have also revealed that at large, the research on health policies is more focused on better health policies (Innvaer et al., 2002; Gilson & Raphaely, 2008) than on how public health policies should be drafted, and systems progress. A current report shows that out of 238 research scholarships approved by a donor agency in the past 15-16 years, only 11 percent cared for 'health policy analysis' signifying that such analysis in developing countries are still in their early stages and requires more attention (Ghaffar et al., 2016).

2.5 Prospects of Big Data in Pakistan

Big data analytics has altered the way we analyze, leverage and manage data in any sector around the world. Similarly, acknowledging the importance of Big-Data analytics in different sectors, especially for effective public policy making and growth and development, the government of Pakistan has initiated policy actions to boost the Big-Data industry. The Pakistan's Ministry of Information & Technology is currently working on a blueprint for Big-Data, which will synchronize and link different government ministries through ICT that will digitize the government data (Dodhy, 2017; Mahmood et al., 2013). Pakistan is very adaptive and welcoming when it comes to the inclusion and use of technology in government functions. She was the first country in SAARC to establish the e-government system back in 2002 (Warf, 2016; Younas et al., 2016). On the other hand, the telecom industry which is the largest donor of the ICT industry that lured massive Foreign Direct Investment (FDI) into the country from last few decades. They are the richest sources of Big-Data currently in the Country. Similarly, the ambitious China Pakistan Economic Corridor (CPEC) project, which is initiated recently by Pakistan and China also enabled Pakistan to develop its economy along with technology. Pakistan and China started the idea of

'digital CPEC' and initiated the laying-down of 'high capacity optical fiber' that will help in transforming the ICT in the country (Zahid et al., 2017).

The government of Pakistan proposed various policies with an eye to establish Big-Data industry in the country, and to make it one of the key national industry. The Ministry of Information Technology and Telecommunication (MoITT) is a leading ministry in the country for the development and promotion of ICT in the country. Back in 2015, the then Ministry of Information Technology (MOIT) initiated different policy actions for the promotion of internet through re-organization. As a result, six Pakistan's government ministries, which includes The Ministry of Information Technology and Telecommunication (MoITT), Ministry of Finance, Revenue and Economic Affairs, National Database Registration Authority (NADRA), Ministry of Planning, Development and Reforms, Ministry of Interior and Pakistan Telecommunication Authority proposed a policy draft for the initiation and operationalization of Big-Data industry back in August 2016 (Latif et al., 2018). The previous government also introduced "National Center of Big Data & Cloud Computing" in different universities such as, LUMS, IBA and NUST at the end of their tenure in 2018. The present government is more serious towards the development of ICT and Big-Data industries. It can be seen from their efforts when the federal cabinet approved the first ever Digital Pakistan Policy on 22 May 2018, it was a momentous day and a landmark development in the IT sector in which they have talked extensively about development of ICT and Big-Data industry in the country. (Pakistan. MoITT, Digital Pakistan Policy, 2018).

Consequently, in order to investigate the adoption of big data analytics in public health care the researcher has employed the Data Readiness Concept' and 'Digital-era Governance' (DEG) framework in this study.

2.6 Data-based policymaking frameworks

New ideas and concepts such as the "Digital-era Governance" (DEG) (Dunleavy et al., 2005) and 'Data Readiness Concept' (Klievink et al., 2016) seek to understand the way public sector operates with novel technologies like big data. these are focused on the premise that technology- and-data-driven innovations in public realm require an infrastructure to generate value from data. this notion is closely linked to the idea of e-government technologies that transform government policies into being more accountable and responsive (Jetzek, 2016). The concept of DEG is an heir to the concept of "new public management" and is described as a novel macro-theory in the public sector development (Margetts & Dunleavy, 2013).

In the line of DEG research, it is identified that digitization and new technologies in public services are widely presented as effortless and 'positive notion,' nonetheless in practice the public sector is far short in development as compared to the private sector leading to low rates of literacy linked to technological innovation and sometimes even computer in general (Dunleavy et al., 2006). That means that government employees have to learn a range of new skills, increasing personnel and training costs or outsource expertise. In addition to workers skills, public sector organizations require the aptitude to process knowledge and produce desired results by hiring personnel, implementing SOPs and establishing agencies (Dunleavy et al., 2006). In broad terms, the growing usage of data and technology itself does have an effect on the definition and quality of information, as there is not only excess of information, but it is also linked with the diversity of kinds of knowledge, along with the skill and capacity to view, comprehend and use it (Rose, 1999; Dunleavy et al., 2006).

The concept of data readiness evaluates the government capacities by observing the data readiness of the public organization and gives complementary point to digital era governance (DEG). In conjunction with big data, the concept of data readiness emphasizes on organizational capabilities, alignment and maturity. The capabilities reflect the use of big data by organization connected to IT and data governance, legal compliance and data science expertise. Alignment applies to how the use of big data a good match with the structure and core operations of the organization and finally the organizational maturity is the integration of e-government initiatives inside the organization. These characteristics are connected to a value chain by the DR concept that involves data collection, combination, analysis and use of data in the big data process (Klievink et al., 2016).

The organizational maturity criterion in particular is embedded in the e-government practice of examining the growth stages of e-government: joined-up government, integrated organizations, nationwide portal, stovepipe organizations, inter-organizational integration and demand-driven government (Klievink & Janssen, 2009). In a nutshell, when it comes to digital government infrastructures, the more advanced the public organization with the respect to responding to environmental changes, the better its performance and efficiency will be (Klievink & Janssen, 2009).

Therefore, this investigator finds these two concepts (Data Readiness Concept' and 'Digital-era Governance' (DEG)) most appropriate to the study as it supports the inclusion of big data and new technology to the realm of public policymaking. On the other hand, Digital-era Governance' (DEG) which itself is a macro-theory for public sector development will help in assessing the organizational capacity whether it is capable of adapting to new technologies.

This theory will also allow the researcher to identify the hurdles in the inclusion of new technologies in the public health policymaking process.

Consequently, the above literature review summarizes studies concerning big data application in healthcare sector around the world. It also reveals the current developments of big data supported policies and the slow and steady process of adoption of BDA. The gap between the slow and low-geared adoption and potential pros of big data signifies a greater prospect for researchers to comprehend how big data analytics can be adopted and incorporated in the healthcare in order to move towards data-driven decision making. In Pakistan big data analytics is in the early developing and adoption stage, and the government must design a clear policy and machinery for the acceptance of BDA in government as well as in the private health sectors. The gap focus for this study would be to analysis and identify any and all hurdles, challenges and develop a model for better healthcare and health policy making using big data analytics. Hence, to reduce the gap in the literature, one of the key focus of this research would be to provide a comprehensive and wide-ranging research insight for the integration of BDA in healthcare and public health policymaking.

CHAPTER 3

CONCEPTUAL FRAMEWORK

3.1 Introduction

The Conceptual Framework consists of different constructs and concepts which support the model of the study. It describes how different concepts and constructs may affect the research area. It presents, describes and explains why the research problem under consideration exists. The Conceptual Framework has been designed from various concepts, lines and labels which combinedly creates meaningful statements. The researcher first attempts to identify the hurdles and challenges pegged in the way of big data analytics and big data health policymaking. The key hurdles and challenges identified after rigorous review of the literature include: data privacy, data security/safety, data reliability, interoperability, data management and governance and expertise and infrastructure challenges. Consequently, this research will try to propose a data-driven model that will help in overcoming the above-mentioned challenges including all other social, political and technological challenges (if identified further by the respondents), that will result not only in efficient healthcare system but will also lead to data-driven and evidence-based health policies. Therefore, with the successful application of data-driven model in the healthcare policymaking process, it will ultimately result in policy empowerment. These health policies will be effective on grounds as it will be backed by real-time data.

3.2 Conceptual Framework

Figure 5: Conceptual Framework of "Investigating the Adoption of Big Data in Healthcare and Health Policy-Making".

Concepts		**Key Sources**
Data Privacy & Security	←	(Heart 2010; Fife & Orjuela 2012)
Management & Governance	←	(Europa, 2014)
Data Reliability	←	(Narayanan et al. 2013)
Data Interoperability	←	(Roney, 2012)
Expertise & Infrastructure	←	(IBM, 2013; Davenport & Patil 2012)
Resistance to Change	←	(Sharma et al., 2017; Bates, 2005)
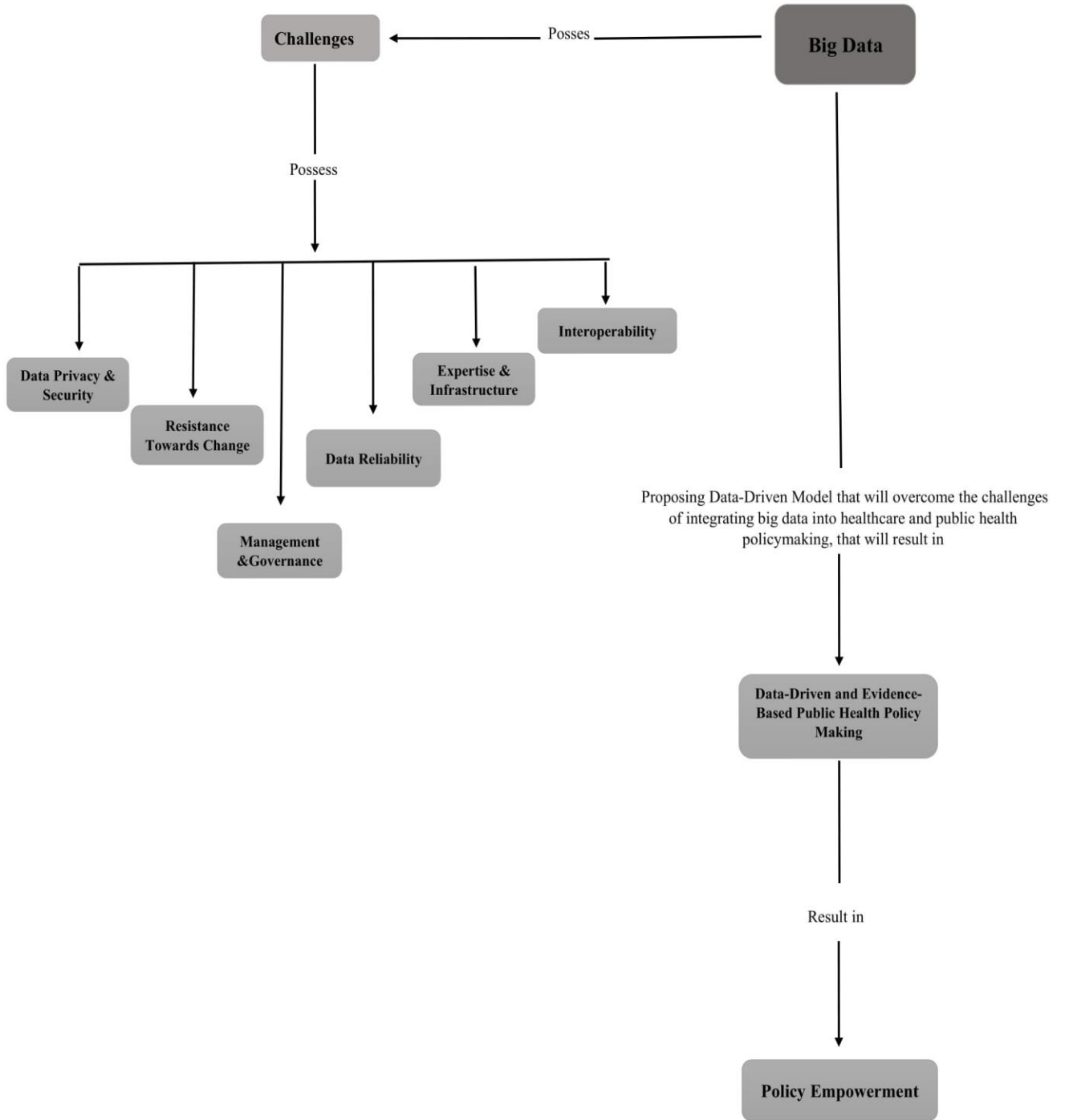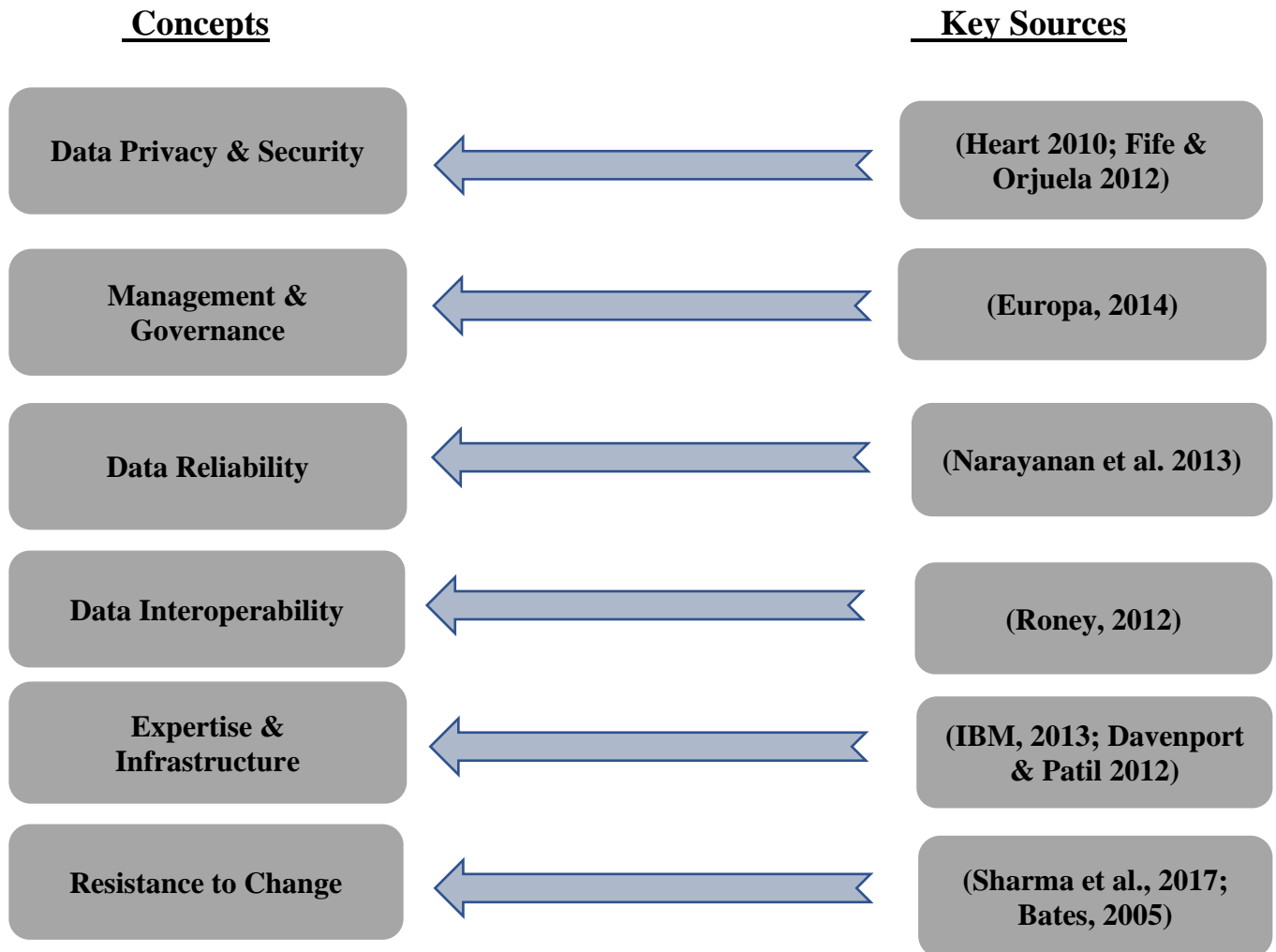
Figure 6: Constructs and Key Sources

CHAPTER 4

RESEARCH METHODOLOGY

4.1 Introduction

In the earlier the chapter researcher identified concepts that made the conceptual framework for this study. It assisted the investigator in the identification of different kinds of concepts which might someway affect the thesis problem under discussion.

The research methodology or design deals with the compressive process that an investigator choose to syndicate the distinct foundations of the study in a rational and comprehensible manner. Consequently, it reassures that the investigator will lucidly discourse the research problem under discussion. "It comprises the strategy for the collection, measurement, and analysis of the data" (Labaree, 2013).

The methodology portion displays the foundation for the application of specific methods or processes used to recognize, select, and securitize data and facts smeared to comprehend the research problem. This part answers to two rudimentary questions: "how the data was generated? and how the data was analyzed" (Labaree, 2013)?

4.2 Philosophy

Exploratory research design was adopted by the researcher to assess the problem, because "the researcher aimed to investigate the hurdles and challenges pegged in the way of inclusive big data analytics and big data health policy making. The philosophy adopted for this study is based on *Social Constructivism and Unobtrusive data assessment techniques,* and "the notion behind this is, individuals try to make sense of things they experience and world they live in. constructionist

focus on the procedure by which meanings are generated, sustained, negotiated and modified (Andrews, 2012) whereas unobtrusive measures will observe and analysis the existing phenomenon in human interactions and statistics without any contact or interference with subject under study." This helps assist in tackling known biases such as selection bias and experimenter's bias and create a triangulation (Triangulation means using more than one method to collect data on the same topic. This is a way of assuring the validity of research through the use of a variety of methods to collect data on the same topic) of the results by employing two different methodologies. The investigator therefore, used inductive and deductive methods, and the meaning and explanation was produced from interaction and observation with the respondents.

4.3 Research Strategy

The research was conducted in two steps, with both quantitative and qualitative reasoning to acquire a triangulated and unbiased results.

Firstly, the researcher employed unobtrusive quantitative strategy such as content analysis (identifying patterns of data relating to the study), that includes; Thematic analysis of text (similar theme or topics), Quantitative descriptive analysis (frequency of data) and Secondary Analysis of quantitative data (statistical data from secondary sources like Governmental surveys) to investigate the progress of big data in developing and developed world. Here, the research collected data from secondary sources to establish the concepts and triangulation of the study and to, later, verify the primary source data obtained through interviews.

The second strategy adopted by the investigator is '*phenomenology',* in which investigator designates the understandings of persons regarding a phenomenon as demonstrated by respondents (Creswell, 2003) in Pakistan due to researcher's limited access. It facilitated the investigator to

acquire maximum information from the respondents with the help of in-depth interviews based on open-ended questions. During this study the investigator preserved its concentration on learning the meaning which the respondents hold about the problem under discussion.

4.4 Locale of the Study

The second part of the study was conducted in Pakistan. The four provinces of Pakistan with a population of 212.2 million people. Specifically, Khyber Pakhtunkhwa (KPK) of the four provinces was selected based on the resources available to the researcher. This place was selected for knowing the provincial government vision of transformation. The provincial health budget is also gradually increased from past few years. On the other hand, the health ministry of the said province has also introduced different health initiatives for the common citizens such as, the Sehat Sahulat Programme and Universal Health Insurance Programme, which is first of a kind in Pakistan. Similarly, the place was selected because of the availability of secondary data in the form of policy document, health sector review document and annual reports of District Health Information System

4.5 Research Method

The investigator adopted both qualitative method (Primary source data) and quantitative method (Secondary source data) for data collection, interpretation, and report writing. "The investigator considers that it is the finest and exclusive method to validate knowledge claims grounded on productive and constructive viewpoint (Creswell, 2013)."

4.5.1 Data collection

The data collection for this study has been done in two parts; Secondary data was collected from existing database like articles, journals, governmental databases, statistics bureaus/surveys and hospital archives. Whereas, the primary data was collected through the process of face to face interviews (Observational data).

4.5.1.1 Secondary Data Collection

Presently, researchers around the world are gathering and archiving a lot of secondary data for investigation that is becoming more common (Andrews et al., 2012). Someone else collects secondary data for his primary research purposes to provide the basic principles of study. Researchers with limited resources and time may use the secondary data for their investigations. The researcher has used both published and unpublished data for the collection of secondary data. Published data are collection from: 1) several provincial and federal government publications and policy documents (e.g., health policy documents, census reports and statistical statement), 2) review reports and publications of foreign governments and international organizations and their subsidiaries (e.g., WHO, ADB and World Bank etc.), 3) numerous publications and research reports in different fields prepared by economists, universities and research scholars, etc., 5) varied sources from university libraries, 6) scientific and technical journals 7) blogs and websites and other sources of published documents.

The unpublished data is gathered from various sources. For this research the unpublished data was collected primarily from two sources: data collected from public/private individuals and organizations relevant to the study.

Three documents namely KP Health Policy 2019, KP health sector review 2019 by ADB and District Health Information System (2019 Annual Report) were used for secondary data collection and subsequent analysis.

4.5.2 Data Collection Tools

Secondary Data was collected by Google search engine, Boolean search tools and Governmental Ministry's archive or Library.

Primary Data was collected with the help of semi-structured interviews from the respondents for in-depth analysis. The respondents separately participated in face-to-face and telephonic interviews with the help of semi-structured and open-ended questions.

4.5.3 Interviews

According to McNamara (2009) "interviews are basically verbal conversation between two people with the objective of collecting relevant information for the purpose of research". In-depth interviews are nondirective and unstructured in nature, which provides interviewees liberty to answer related to the topic under discussion.

4.5.4 Population and Sampling

The participants of the study involve a preliminary descriptive examination of the perceptions of public policy experts, public health experts, big data experts and international partners working in collaboration with Pakistan's health Ministry. Because of the time constraints involved in interviewing and subsequent data analysis it was limited to fourteen subjects.

For Primary Data, purposive sampling (also known as judgment, subjective or selective sampling) was used. "It is a sampling technique in which researcher relies on his or her own

judgment when choosing members of population to participate in the study". The sampling was conducted on the basis of reaching a saturation point in terms of repetition of data respondents provide. Purposive sampling is a non-probability sampling technique and it happens when "elements selected for the sample are chosen by the judgment of the researcher. Researchers often believe that they can obtain a representative sample by using a sound judgment, which will result in saving time and money (Black, 2010)."

4.6 Data Analysis

The investigator adopted qualitative technique for data analysis and interpretation, the purpose was to uncover and understand the big depiction by using the data to define assured certain phenomenon in both developed and developing countries. Qualitative methods contain the gathering of information in bulks with the help of focus group discussions, interviews, observations, and documentary analysis, which needs to be described and summarized further (Lacey & Luff, 2001).

Qualitative data analysis includes some or all the subsequent phases, however the order might differ: reading through all data by hand or with the help of computer, organizing and preparing data for analysis, , themes and description of coded data, interpreting the meaning of those themes/description, transcription of raw data, interrelating themes or description, , all these stages leads to accuracy and validation of data (Creswell, 2013).

4.7 Conclusion

In this chapter, an introduction was presented that explains the research design adopted in this thesis. The discussion on population and sample study were provided before explaining the research methodology used to perform this work. Library and filed research are used for the

collection of data. For library research, KP Health Policy 2019 and KP health sector review 2019 by ADB was reviewed and analyzed. A qualitative approach was used for field study, in which a number of interviewees are included in the research process. The collected data was then analyzed, discussed, and the findings were presented accordingly.

CHAPTER 5

DATA ANALYSIS AND FINDINGS

5.1 Introduction

This chapter presents the research analysis and findings of the data obtained from library sources and face-to-face and telephonic interviews. The purpose of this study was to identify hurdles to the inclusion of big data in health policy formulation process in Pakistan. The both primary source and secondary source of the data were employed through quantitative and qualitative methods such as content analysis, secondary data analysis and interviews respectively. The data collected through interviews was then transcribed, analyzed and interpreted by the researcher. Whereas, the library and archives were combed through for secondary data. The findings were presented in align with the research questions and objectives. The method used for analyzing the data is discussed already in the methodology chapter. The foremost attention of the analysis was to evaluate the theoretical framework in order to further develop the framework or minimize the current concepts, if the concepts are not associated with the subject under consideration. The objectives and questions played a vital part in deriving findings of the research. The extensive review of the documents including KP Health Policy 2019 and KP health sector review 2019 by Asian Development Bank further strengthened the analysis section.

It also offered a detailed picture of the current problems persists in the health sector. These documents supplemented with interviews played a crucial role in offering a detailed picture of the current health care sector and how big data can play its role in reforming the said. Consequently, the analysis is distributed among two sections. The first section presents the evaluation and examination of the current health sector of Pakistan using content analysis where both quantitative

and qualitative data are being analyzed. The second section covers the subsequent analysis and interpretation of the data collected through face-to-face and telephonic interviews. This section mainly focused on identifying the hurdles to the inclusion of big data in health policy formulation process in Pakistan. It also further helped the researcher in drawing the discussion and recommendation of this study.

5.2 Demographics of the Respondents

5.2.1 Categories of the Respondents

A total of 14 interviews were conducted to analyze the research questions. The respondents were categorized on the basis of their association and expertise in the fields such as public policy, healthcare, big data and international partners who assists the government in the policy formulation process. Out of 14 participants 4 participants were public policy experts, 3 were part of healthcare, 4 big data expert and 3 participants were from international agencies working in collaboration with Pakistan government. One of the main problems which caused difficulties while conducting interviews was the Covid-19 pandemic. It was due to the pandemic which caused safety issues for the respondents. The interviews took longer time than estimated to receive responses from the respondents. Resultantly, majority of respondents agreed to write their responses and some interviews were conducted via phone. The researcher therefore collected written responses as a main source for the collection of primary data.

Table 2: Classification of the Respondents

Informants	Number
Public Policy Experts	4
Public Health Experts	3
Big Data Experts	4
International Partners	3

Table 2. includes detail on the types of informants interviewed based on their association and expertise in the fields such as public policy, healthcare, big data and international partners who assists the government in the health policy formulation process. These individuals have vast knowledge and expertise in their respective fields.

It was discovered by the investigator that the data obtained from 14 participants is saturated. It implies that the investigator recognizes a common trend/similar pattern in the responses taken from interviewees. The collection of new data would not necessarily result in a new understanding of the problem. As stated by Ritchie et al. (2003) it is ideal to collect data till theoretical saturation is achieved. In qualitative sample there comes a point of diminishing return as the analysis of additional data does not automatically result in more details or information. The above-mentioned individuals were the individuals who were part of the field-work interview process and their views were consequently obtained, analyzed and interpreted.

5.3 Findings

This section of the chapter discusses the findings and analysis of the data collected from library sources and interview with the respondent on the hurdles to integrate big data in public

health policy formulation process in Pakistan. The data were obtained and analyzed in accordance with study questions and objectives.

5.4 Dynamics of Evidence-Based Health Policy Formulation

5.4.1 Health Policy Formulation in the Post 18th Amendment

Under the 18th Constitutional Amendment of Pakistan which took place in 2010 transferred health governance and management to the provinces. The provinces were granted the legislative and executive authorities of the health sector that were previously under the control of federal government. The respondents were asked in the beginning of each interview about the health policy formulation process in the country, to which they responded with the following answers. One respondent had a detail response to the question and elaborated the whole scenario of the 18th Constitutional Amendment. The respondent stated that:

"The 18th Amendment in the 1973 Constitution has devolved health planning, service delivery and programmed implementation to the provincial level. This constitution has two lists of responsibilities. List-1 which includes subjects like defense, foreign policy etc. is the responsibility of federal government. List -2 which include subject like rural development, local government, health and education etc. is the responsibility of provincial government. A 3rd list called the concurrent list of responsibilities was also kept in the constitution for a limited period. The concurrent list included subjects which became the joint responsibility of both federal and provincial government. Health was also included in the concurrent list and a federal ministry of health was looking after this subject until 2010 when 18th Amendment in the constitution was implemented and federal ministry of health was dissolved along with 21 other ministries. After the 18th amendment the

provinces in addition to service delivery and program implementation also became

responsible for strategic planning." /R7

Another respondent stated that:

"Post 18ᵗʰ Amendment, health policy is primarily driven by provincial health ministry's

instead of the federal government. Ministry of Health, of the particular province and its

staff are responsible of devising a health policy in the province." /R2

The role of World Health Organization (WHO) cannot be ignored in the health policy formulation for development and underdeveloped countries. In the case of Pakistan, WHO vigorously provides technical assistance to federal and provincial governments in the policy formulation process. Here is how one participant responded:

"To meet the new responsibility of strategic planning all the provinces established health

policy units and developed provincial health policies; WHO provided technical support in

this process of devolution. However; some federal function in health became fragmented

and distributed between different ministries and divisions." /R10

One the other hand, one of the respondents further stated that:

…All the provincial health strategies used WHO health system framework of six building

blocks to design their strategies. These strategies still await costed and financed

implementation plan. However; these strategies provide good guidelines for planning at

the provincial level…"? /R11

It can be observed that WHO plays a crucial role in providing technical and financial assistance to provinces in drafting their health policies and programs.

One of the respondents stated that the devolving the legislative and executive authorities to provinces under the 18th constitutional amendment of Pakistan was not ideal but initially worthy, as it had several problems in the later stage:

"After the devolution we have policies made on provincial level, it was not ideal but that was something which was initially worthy. But after the devolution, the federal devolved health sector to the provinces, there was a lot of bounce because the health ministry, NIH and some other stakeholders came up with National Health Services to implement some mechanism and projects at the federal level, but if you talk about the health policy it is very fragmented, it is not the best, it has issues and they are ineffective…" R/1

On the other hand, one respondent stated that:

"Although, after 18th constitutional amendment the health is a provincial subject but in practice the health is controlled by the center with political influence and context. The aim of controlling health from center is controlling the financial resources rather than the better public service delivery system. Thus, from agenda setting to policy evaluation, there should be a decentralized approach with the effective collaboration and consultation of all stakeholders." /R12

The health sector was devolved to the provinces as a result of 18th constitutional amendment. This has resulted in the transfer of legislative and executive responsibilities to the provinces. But it has been identified that the all the four provinces lags capacity to formulate health policies in the province. That is why the role of international actors such as, the World Health

Organization, the World Bank, German Development Agency Deutsche Gesellschaft für Internationale Zusammenarbeit GmbH (GIZ, German development bank Kreditanstalt für Wiederaufbau (KFW) and several other donor agencies comes into place.

5.5 The Use of Big Data in Developed World

The health sector data has increased from last few years. It is also expected that this data will grow dramatically in years ahead (Cottle et al., 2013). Especially in developed world like the US and the UK, Germany, Japan, Canada and several other countries have already developed their infrastructure when it comes to handling healthcare sector big data and utilizing it for better and informed decision making. The developed world is consistently working on bringing more and more efficiency when it comes to handling vast amount of health data. They have developed systems which manages these large data sets on a daily basis. It also helps the decision makers by providing clean data after rigorous analysis so that evidence-based decisions are made. The following is the consequent analysis of healthcare sector of the UK and Japan who have made great progress when it comes to healthcare data handling.

5.5.1 United Kingdom

The UK has one of the best healthcare systems. Big data analytics and Artificial Intelligence of healthcare data is at the forefront of UK government's long-term strategy. With the National Health System (NHS), the UK, a global leader, has a strong competitive advantage, possessing massive amount of healthcare data. Although research have shown that the UK is still faced with several hurdles concerning governance frameworks, establishing procedures to gain access to data and patient agreement to trade privacy risk for better health services provisions

patient (Lawrence & Bradley, 2018). Similarly, linking two or more discrete data sets is another big challenge and a laborious process. Nevertheless, the country is constantly working on bringing new laws and regulations in order to operationalize the process of data management and its use for informed decision making.

The following are some of the developments that took place during the past few years in the UK:

> *"The UK has established a national strategy that has resulted in the likes of the UK Biobank and 100,000 Genomes projects. These projects dovetail with a national strategy for the implementation of genomic medicine with the opening of multiple genome-sequencing sites, and the introduction of genome sequencing as a standard part of care for the NHS." – (Topol, 2019)*

Similarly, the UK has also introduced a Biobank for individuals between 40-60. As one of the studies reported:

> *"The UK Biobank is a prospective cohort initiative that is composed of individuals between the ages of 40 and 69 before disease onset. The project has collected rich data on 500,000 individuals, collating together biological samples, physical measures of patient health, and sociological information such as lifestyle and demographics. In addition to its size, the UK Biobank offers an unparalleled link to outcomes through integration with the NHS." – (Agrawal & Prabakaran, 2020)*

It is a unified healthcare system which enables investigators to link initial baseline measures to disease outcomes and multiple sources of medical information from clinical admissions to hospital admissions. This makes it possible for researchers to be better able to eliminate mistakes in the classification and diagnosis of diseases.

In addition to the UK Biobank, the 100,000 Genomes project was also launched by Public health England. As stated by a study:

> *"Public Health England launched the 100,000 Genomes project with the intent to understand the genetic origins behind common cancers. The massive effort consists of NHS patients consenting to have their genome sequenced and linked to their health records. Without the significant phenotypic information collected in the UK Biobank." – (Turnbull et al., 2018)*

The project is an addition to identify disease that result in single-nucleotide polymorphisms (SNPs). This project is also important in stimulating private sector growth.

Although there are some complexities when it comes to the privacy, integration and use of these systems. For instance, the NHS in the UK is an integrated platform with a variety of data sets, nevertheless this system is full of vulnerabilities. As identified by one of the studies:

> *"The UK provides an informative contrast in its NHS, a single government-run enterprise that provides free healthcare at the point of service. Currently, the NHS is able to*

successfully integrate a variety of health records—a step ahead of the US—but relies on outdated technology with security vulnerabilities such as fax machines." – (Macaulay, 2016).

Another report stated that the EHRs are rather misguided and often cause hindrance:

"Another issue often identified with the modern incarnation of EHRs is that they are often not helpful for doctors in diagnosis—and have been identified by leading clinicians as a hindrance to patient care." – (Gawande, 2018)

Another major issue is the data ownership and data sharing. There are complications when it comes to the compliance of data with different laws and regulations. Therefore, the government of UK has to adopt a clear-cut policy on this.

5.5.2 Japan

Japan is world's number 1 in longevity. It is one of the countries that has made remarkable progress when it comes to big data handling in the healthcare sector. Japan is determined and earnest in order to bring continuous improvement in healthcare services delivery. The country is keen to prolong life expectancy and make healthcare sector a growth industry. For that reason, the country has also introduced a law concerning the use of patient data to facilitate medical research. As one of the report states:

"The law, commonly called Jisedai Iryo-kiban Ho — roughly translated as the next-generation medical infrastructure law — allows medical big data to be pooled anonymously so it can be utilized for research into diseases and the development of new drugs." – (Otake, 2017)

Medical data is largely untapped due to privacy concerns and lack of relevant laws. That is why the new law introduced by the government of Japan will ensure that the data is used for better use without compromising patient privacy. As the report stated:

"Medical records of people who receive care at hospitals and clinics have remained largely untapped due to privacy concerns and the difficulty of handling sensitive data. But the government believes they can be put to better use if they are made anonymous, with the goal of spurring drug development and other medical innovations." – (Otake, 2017)

The government ensured that the data is now available for analysis purposes, so that every patient receive the best possible care they need. Furthermore, when it comes to privacy the law states that:

"The Law on the Protection of Personal Information that came into effect this month have, in fact, enabled the use of clinical records for research without asking for a patient's consent. But to use such data, each hospital and clinic is responsible for making its data anonymous by deleting patients' names and other privacy information, which is seen as cumbersome when their primary job is providing health care." – (Otake, 2017)

According to this law:

"All clinical records, including a summary of conditions, prescriptions written by doctors and visual data such as MRI scans, will be included." – (Otake, 2017

When it comes to any danger of data breach and leaks that law ensures that:

"Privacy protection is key, especially with medical data, which are highly sensitive. The government will allow only a small number of trustworthy operators under a small project so that privacy protection is ensured. Confidentiality will be ensured through agreements, penalties, recording entry into data rooms and separating data basis from the internet." – (Otake, 2017)

Japan has also developed cloud-computing system in order to support sequencing analysis: As one of the studies explained:

"The Read Annotation Pipeline®36 by the DNA Data Bank of Japan (DDBJ) is a cloud-based pipeline for high-throughput analysis of next-generation sequencing data. DDBJ initiated this cloud-computing system to support sequencing analysis. It offers a user-friendly interface to process sequencing datasets, which supports two levels of analysis: (1) the basic-level tools accept FASTQ format data and preprocess them to trim low-quality bases and (2) during the second analysis, the data are mapped to genome references or assembled on supercomputers." – (Luo et al., 2016)

On the other hand, there are several other Acts through which the government of Japan is collecting data such as The Acts on Assurance of Medical Care and the National Cancer Registry. These laws help the government to provide personalized services to citizens. On the other hand, the country also has Hospital Information System (HIS) which records physicians' orders that is then used for examination and decision-making purposes (Ishikawa, 2016). It can be concluded that Japan has one of the best healthcare systems in the world.

5.6 The Use of Big Data in Developing World

The developing world lags behind on many grounds when it comes to management of big data. The developing world is currently faced with several issues. There are issues concerning rules and regulations, privacy, security, lack of resources, lack of innovation, patient safety, government will and integration of data. With some exceptions these issues are prevalent in majority of the developing countries. As big data is in early adoption stages, many countries in the developing world are trying to formulate strategies in order to use healthcare big data for personalized health services provisions and making informed decisions. Studies shows that, in developing countries there is still a need for technological infrastructure and resources that can collect, store, evaluate and visualize vast quantities of varied structured and unstructured data. The following section shows the progresses and issues concerning big data integration in healthcare sector of developing countries including Bangladesh and Iran. After reviewing the studies and reports, the above countries possess similar issues and prospects concerning big data integration in the healthcare sector.

5.6.1 Bangladesh

Bangladesh is among the fast-growing economies of South Asian region. The country has made a lot of progress when it comes to bringing new innovation and technological development. One of the areas which they have targeted is the public procurement. The government has introduced a Public Procurement System, which is an e-procurement system to make the procurement process transparent and efficient at the same time (Roy, 2018). Studies shows that it is a promising system which can bring improvements in the healthcare sector with the use of BDA. As one of the studies stated:

"The new notion 'Procurement 4.0' comes into the forefront. Procurement 4.0 integrates technologies such as the Internet of Things, Blockchain, Big Data Analytics (BDA) in procurement framework to overcome the limitation of e-GP." – (Murad, 2020)

This study further stated that with the adoption of this system which is currently in the development stage, will result in transparency:

"For example, if a media report that some medical equipment is now dust-ridden, badly damaged, uncared, BDA will consider those purchases were done without prior need assessment and will advise keeping the particular organization on constant watch. Suppose a media reports that a sub-district level hospital purchased some generators, out of which none has been found in operation, costing a massive loss to the public, depriving the citizens of sufficient healthcare. BDA will analyze this unstructured dataset with its natural language processing intelligence to identify the types of equipment that need to oversee

with more vigilance. If any pricey medical devices and instruments have been found either unused or lying idle for years, some of them never unboxed ever, BDA can detect those from RFID signals. Accordingly, if equipment found wasted such as anesthetic equipment, hydraulic OT Table, delivery table, incubator, BDA will detect which other hospital settings need those." – (Murad, 2020)

This is indeed remarkable to bring transparency in the healthcare sector of Bangladesh if this system is implemented in its true shape. On the other hand, there are several issues concerning the use of healthcare data in Bangladesh. For instance, a study pointed out:

"One of the significant difficulties in utilizing health care's huge information to its full degree is policies that protect the privacy of patient's information in Bangladesh which is yet not strict till date. Numerous laws ensure the patient's data and not uncover the patient's identity that makes the big data analytics troublesome. Actually, now and then healthcare providers are themselves reluctant to share patient's data because of huge competition in Bangladesh. A doctor may not want their competitors to know precisely what number and which sorts of techniques they performed and where." – (Kabir, 2018)

Studies and reports show that there is a great prospect of healthcare data utilization in Bangladesh. Although, the government will need to ensure that there are laws in place before using citizens data for decision making purposes.

5.6.2 Iran

Iran is one of those developing countries which has an efficient healthcare sector. The efficiency came with the introduction of Electronic Health Records (EHR), also called SEPAS, in Iranian hospitals. According to a study:

"SEPAS is a large-scale project which aims to build a nationally integrated system of EHR for Iranian citizens. Over the last decade, Iran has progressed from having no EHR to 82% EHR coverage for its citizens. EHR is one of the most widespread applications of medical big data in healthcare. In effect, SEPAS is built with the aim to harness data and extract value from it and to make real-time and patient-centered information available to authorized users." – (Financial Tribune, 2017)

Although it has been pointed out that the data generated through EHR is not utilized to its full potential. The harnessed data should inform decision-making process and drive actionable outcomes, if big data system is to be successful. Similarly, a study stated that:

"Currently, data is gathered effectively in Iranian public hospitals, meaning that the raw and unstructured data is mined and classified to create a clean set of data ready for analysis. This data is also transferred into summarized and digestible information and reports, confirming that real potential value can be extracted from the data." – (Tabrizi, 2018

The study further stated that:

> *"In spite of this, the benefit of big data is not yet realized in guiding clinical decisions and actions in Iranian healthcare. SEPAS is only being used in hospitals by IT staff and health information managers who work with data and see the reports from the system. However, the reports and insights are not often sent to clinicians and little effort is made by management to extract lessons from some potentially important streams of big data." – (Tabrizi, 2018)*

The studies show that Iran has an efficient healthcare data system in place, although the government needs to bring systems that will analyze the data for decision making purposes. This will not only result in better health services provision but will also ensure sustainable health practices.

5.7 The Use of Big Data in Under Developed World

The healthcare sector in under developed world is even bleak as compare to developing world. There are several issues in underdeveloped world. Some of the major issues in these countries include, lack of resources, population growth, lack of human and resource capital, infectious diseases, mortality and so on. These issues can only be addressed if international organizations join hands with governments of the concerned countries. On the other hand, there are few countries who are making progress when it comes to healthcare services provision. Similarly, there is very little research being carried out on the topic of big data in healthcare sector of underdeveloped countries. Nevertheless, on the African continent, governments and commercial organizations are starting to take interest in using big data-related technologies to evaluate the

enormous amount of data they currently produce and wish to do so in real-time. This uptake is fueled by the developments being made in big data technologies. Ghana is one of the countries who is working on integration of innovative approaches in their respective healthcare sectors.

5.7.1 Ghana

Ghana is a country of western African, located on the coast of the Gulf of Guinea. Until recently, Ghanaian government and businesses were unaware of the usefulness of big data analytics, primarily, because of the lack of awareness and minimal diffusion of these innovations. Nevertheless, businesses are now rapidly recognizing the difference in value that big data analytics can make in their decision-making process and designing strategies that will give them a competitive advantage. For many of these corporate organizations, it has become apparent that they have vast quantities of data that if properly examined, can provide then with a wealth of expertise to operate their companies more effectively and productively. One of the studies conducted in this regard states that:

"In recent times, Ghana has seen a major shift from paper based to electronic record keeping in most of the agencies and ministries. For example, recently the National Health Insurance Authority (NHIA) introduced biometric data collection for all clients on their scheme. Other agencies such as the National Identification Authority (NIA), the Electoral Commission (EC), the Ghana Education Service (GES), the Social Security and National Insurance Trust (SSNIT) and the Ghana Health Service (GHS) are all transitioning from the traditional data collection and progressing to electronic data processing and collection." – (Abdulai, 2018)

The study further stated that:

> *"The health industry in Ghana generates millions of data records, but most of these are stored in hard copy form, whereas the current trend is toward rapid digitization of these large amounts of data. A number of health facilities are now moving towards digital records, but currently all these efforts are segmented and disjointed. To derive benefit from the digitization process, these efforts by individual facilities have to be coordinated and centralized (Asangansi & Braa, 2010). Effort must be made to implement an architectural platform onto which individual agencies can simply 'plug in and play'." – (Abdulai, 2018)*

The study also identified the hurdles pegged in the way of implementing big data in Ghana:

> *"There are several challenges when it comes to the inclusion of big data in Ghana. One of which is heterogeneity and incompleteness. Another challenge is with the volume of data to be worked on within an organization. Managing large and rapidly increasing volumes of data can be challenging and requires that faster processing components and storage systems be designed and built. Also, with large data sets to be processed, speed could be an issue to deal with. Another challenge is privacy. For instance, there are strict laws governing what can and cannot be done with electronic health records. Big data raises concerns and fears regarding the inappropriate use of personal data, particularly through linking of data from multiple sources.*

"The implementation of Big Data in Ghana comes with its own challenges apart from the ones discussed above. The lack of infrastructural support and the right technology is also a challenge to the implementation of big data in Ghana. Another challenge is the unavailability of the skilled personnel with the knowledge and skills related to big data analytics." – (Abdulai, 2018)

It can be concluded that, Ghana is currently going through a transformational phase where traditional data systems are being gradually replaced with automated systems. Although there is a need for infrastructural support and skilled human resource to back this transformation.

5.8 The Use of Data and Evidence in the Health Policy Formulation Process

Health policy making is a multifaceted process and data and evidence has greatly influenced and affected by the institutional aspect of policy formulation process. Evidence and data are fundamentals in efficient policy making. Quality data and credible evidence may act as a bridge in health policy formulation. However, when it comes to health policymaking in Pakistan very little attention is given to data and evidence. Especially, when it comes to the KPK health policy document it has been stated that the process of developing a health policy in Khyber Pakhtunkhwa was based on the issues and priorities of the current government. Nowhere it is mentioned that the policy is drafted on the basis of the data and evidence which was reviewed through a consultative process.

"The health policy was prepared in view of the key issues/problems related to health status of the population, health system operations, intersectoral coordination for population health improvement, priorities of current government in health sector achieving health

related SDGs and Universal Health Coverage, national priorities in health sector (as envisaged in National Health Vision 2016- 25.), and international/regional commitments. etc." /KPK Health Policy Document

The use of data and evidence in the health policy process is a far cry. One of the participants elaborated the current condition with an example stating that:

"Sadly, in Khyber Pakhtunkhwa, the policy making process is directed externally from WHO, World Bank and SDGs. Very little policy making is done on provincial level taking in account the problems and evidence of that area…

…For example, if you look at Polio, we will see that the majority of cases are coming from Khyber Pakhtunkhwa, the government, instead of targeted policy creation from that data, has interventions of three-days and one-week anti-polio march of healthcare workers going door to door. This is a waste of manpower and resource as one month later the polio count is back to being high…

…The idea is to collect data, Numbers of polio cases and then from that numbers identify patterns why people do not vaccinate against polio and then solve the problem." /R7

It is true with every disease outbreak in the province, the government use a reactive approach without consulting the data and evidence to address the problem from the roots. Another participant has given an example of HIV in the country stating that:

When it comes to HIV spread and patients. Where we do not know the number of cases or the cause/pattern behind that numbers like why people are contracting this disease or what the rate of spread is in patients, but solution is derived from international standards and no attention is paid to our own reasoning." /R12

Upon asking a question whether data and evidence is used in the health policy formulation process, almost all the participant stated that the use of data and evidence is not there. The policymakers give very little attention to data and evidence while drafting health policies. One respondent provided a broad view of the prevailing health policymaking process:

In KPK, there exists a vide gap and health policies had made and implemented without quality data and credible evidence." /R13

One of the respondents was of the view that:

"...some broad data points might be used, but generally the use of data to formulate a broad-based health policy is not available." /R7

Similarly, another respondent stated that:

"The use of data and research is not there in the health policymaking process, although there are different projects and programs where data is used." /R10

This was supported by another respondent in the following statement:

"There is very minimal use of data in the policy process." /R3

Similarly, another participant had similar views:

"There are practices around the world where policies are backed by data, and they have given good results. But here unfortunately it is not there, and we don't consider it, and we are not really going into that direction. You can see in the recent outbreak of Covid-19, there was no awareness nor any data there. These surveillance units were collecting data, the policies should have been based on data, that was a policy document which should

have been supported by data so that it could have predicted future prospects related to the current pandemic and any other disease outbreak." /R1

One of the participants was of the view that the data is collected on the diseases and the number of cases, but that data is not being utilized to identify the policy problem itself:

"No, the government has collected statistical data on diseases and causes and number of patients but that is not being using to identify the policy problems in itself. Like we know 5 people are sick, but we do not know why they are sick and how they got sick, what sort of treatment could they have and what is they medical, family, social history." /R18

On the other hand, one of the respondents was of the opinion that the role of data and evidence is not significant when it comes to health policy formulation in the country, as stated:

"In my opinion, the role of data and evidence is not significant. Decisions are based on broad indicators, such as areas where there are in general lack of facilities and disease, but data does not dig in to the details. I don't think the ministry of health uses that data for policy formulation." /R2

One of the participants had a different view when it comes to the use of data and evidence in the health policy formulation process:

"I think the trend has started to change and now the practices have indeed started to shift towards data and evidence backed policy making. A recent example of this is recent pandemic where data was used to locate the COVID-19 hotspots and then being sealed afterwards to reduce the spread of the virus." /R9

It is true on one side, but these are reactive approaches when it comes to any disease outbreak and epidemic. In the case of Covid-19 pandemic there was still no sign of preparation in the beginning on how the government will deal with the situation because there was no timely data which could help the government to formulate a better strategy in order to stop the outbreak in the country.

5.9 Awareness of Big Data and Big Data Analytics Concept Among Respondents

In this specific question the respondents were asked whether they have any knowledge and understanding about the concepts such as big data and big data analytics, their understanding of the mentioned concepts is quoted and discussed below.

These are the following responses after asking about the understanding and knowledge regarding the concept of big data and big data analytics. One of the respondents elaborated in detail stating that:

> *"Yes, Big data refers to a large amount of information being digitized, consolidated, standardized, analyzed and modeled/patterned. In healthcare, big data uses specific statistics from a population or an individual to research new advancements, reduce costs, and even cure or prevent the onset of diseases. In recent years, healthcare data collection has moved into the digital realm, making analysis faster and more accurate. Health professionals, just like business entrepreneurs, are capable of collecting massive amounts of data and look for best strategies to use these numbers." /R7*

Another participant provided a comprehensive definition of big data analytics by stating that:

"Big data analytics is a process used to extract meaningful insights, such as hidden patterns, unknown correlations, market trends, and customer preferences. Big Data analytics provides various advantages it can be used for better decision making, preventing fraudulent activities, among other things." /R11

One respondent stated that:

Big data is a term that describes the large volume of data – both structured and unstructured. But it's not the amount of data that's important. It's what organizations do with the data that matters. Big data can be analyzed for insights that lead to better decisions and strategic business moves. /R14

Similarly, other participants have defined the concepts in the following statements:

"It is the amount of data just beyond technology's capability to store, manage and process efficiently." /R8

"It is data whose scale, diversity and complexity require new architecture, techniques, algorithms, and analytics to manage it and extract value and hidden knowledge from it." /R10

"It is high-volume and high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight, decision-making, and process automation." /R5

"It refers to technologies and initiatives that involve data that is too diverse, fast-changing or massive for conventional technologies, skills and infrastructure to address efficiently. Said differently, the volume, velocity or variety of data is too great." /R4

"As the name suggests big data possess large data sets that are analyzed through computers in order to reveal associations, trends and patterns for better decision making." /R3

"Yes. Large data sets, and their analysis for macroeconomic decisions." /R2

"Yes. It is the way of filing large data sets using advanced systems where there is less human interaction." /R1

"It is about handling large data sets with the help of computers." /R6

One of the respondents explained the concept with respect to health data stating that:

"...I believe big data refers to the large datasets collected in national surveys and through other health registries; analyzing big data is indeed an effective practice as it can expand our empirical knowledge towards issues of importance." /R9

As the definitions and explanations suggests the majority of the respondents have understanding and knowledge of big data and BDA concepts. They were also aware of their importance and uses in the health care.

5.10 Challenges to Big Data Integration

In order to achieve the goal of maximizing the potential of current and rapidly evolving data sources needs to overcome a number of hurdles and challenges that defines today's

increasingly heterogenous data environment along with a number of complex policies, governance, normative and regulatory constraints. Consequently, it is important to tackle not only infrastructural hurdles/challenges, but also societal, legal and ethical challenges need to be addressed.

On the other hand, one of the main approaches for development and growth is the use of data and empirical evidence in the policy formulation process. The role of data and evidence is therefore inevitable to formulate effective policies. Similarly, in the health policymaking process, data and evidence plays a very crucial role. In the case of health policy formulation in Pakistan, there are several major issues which has resulted in ineffective public health policies.

5.10. 1 Issues Concerning Interoperability, Data Integration, and Data Sharing

Health data is the area where there are a lot of issues and sensitivity involved when it comes health data collection in the country in general and Khyber Pakhtunkhwa in particular. It is because the reason that there are several organizations involved in the data collection process. Most of the times this data is collected for or under specific targeted programs. This data is available in silos and not integrated at one place to make policy decisions. As stated by one of the respondents:

> *"When health data comes in, there is a lot of sensitivity and issues with the data. data is in area whether it is health or any other sector or organization, they are very particular about data…. talking about data, unfortunately we don't have statistics or data. although there are donor driven projects, we gathered some data, which is available in understandable form, but if you talk about broader health data, I don't think we have the kind of data to make decisions or policies and programs…" /R1*

CHAPTER NO 5: DATA ANALYSIS AND FINDINGS   69

The respondent further elaborated that:

"…Whether this data is available. It has multiple angles, how much we are collecting, from where we are collecting, do we need that, are we sharing it with the right people. if you say policies are based on the data, then I would say no the current data doesn't match, nor the current condition, we don't have that capacity. The provinces have District Health Information System (DHIS) which collects routine health data, but that data is also not used for decision making purposes." /R1

When it comes to the use of data and research in health policymaking, one respondent stated that data is defiantly generated but there are several issues involved with it starting from the collection of data to the use of data. Similarly, these organization are not integrated with each other therefore the data collected is not utilized in a proper manner. As stated below:

"Data is there but it is not integrated in the policy making process. Questions remains, who is collecting the data, who make policies, are these organizations integrated with each other?" /R5

Similarly, one respondent stated that:

"…one the one hand, there is lack of data and research, and, on the other hand, there is incapability to utilize data and research for health policymaking in the country…" /R4

Another participant stated that:

"Data integration is one of the main issues in health sector because data is not there is one centralized system which can be utilized for policies". /R6

One respondent was of the view that there is a need for synchronization of data across data sources. It is in fact very crucial because if data lives in silos and separate systems it will be difficult to use it for any decision-making purpose:

"There is a need for data synchronization across data sources." /R12

Similarly, one participant mentioned a few other challenges that are linked with the issue of data integration:

"There are many challenges in implementing big data in healthcare especially in relation to integration of data, data accommodation, data classification, incorporation of technology etc." /R14

From the above statements it can be observed that there is a lack of integration in the departments who are responsible for collecting, dissemination and utilization of data. The data is generally not utilized in the health policy realm. The above statement is therefore further supported by another respondent, who stated that:

"There are electronic systems developed for the collection, reporting and dissemination of routine data on provincial level such as the District Health Information System (DHIS) of Khyber Pakhtunkhwa. It is a very helpful system for collecting and aggregating data related to routine facilities and community services. On the basis of this data there are different programs designed but I don't think this data is being utilized for health policy formulation."/R5

Another respondent rightly found out that there is a need of a centralized information system where the data is gathered, analyzed and shared in order to make informed decisions:

> *"The biggest hurdle is understanding the importance of a centralized information flow or system where everyone can use that information for their benefit. For example, Hospitals share their data on patient's medication to Pharmaceutical companies so that they are create exact amount of medicine and do not waste drugs, ingredients of medicine and cost of production and import of drugs in creating useless product." /R7*

Another participant was of the view that due to the segregated systems and silos working nature of public and private health sectors the data is not shared and eventually it is not utilized by the government for policymaking process:

> *"I think with a large number of healthcare being privatized and depends on profit generation; the biggest challenge is data sharing. Private hospitals in Pakistan are recording patient's information but for their use only, they are not sharing this with the patients, let alone with government or other hospitals. This data collection and information is being blocked from reaching policy makers and important figure heads…*
> *…It is like they have the means to collect information, but they are not using or utilizing it completely or they do not know how to use it completely." /R8*

The above responses show that the government does collect a large amount of data but there is the lack of capacity and will to integrate the data in one centralized information system so that it can be used for better decision making.

5.10.2 Resistance Towards Change

Resistance towards change is one of the key highlights of this study. In developing countries like Pakistan, the resistance towards change from doctors and hospital staff is one of the main reasons which affects the inclusion of data-driven systems in the healthcare sector. Previously, when the system of 'electronic health records' (EHRs) were brought in the KPK healthcare sector, the doctors, physicians and hospital staff were seen as a main hurdle for the new system adoption. Resistance towards change (RTC) is therefore one of the main hurdles that slows down the acceptance of new data-driven and information system in the healthcare and eventually in public health policy-making. In order to highlight the issue of RTC, the respondents were asked about the main hurdles in the adoption of big data analytics in healthcare. Their responses show that it is one of the major hurdles in the process.

One participant stated that:

"Whenever there is a new change coming in the organization, the employees always resist it, it is due to the fact that they don't want to leave their comfort zone, nor they want their comfort zone to be disturbed by any new change. Similar is the case with health sector, whenever the government tries to initiate new programs the medical practitioners resist it harshly." /R3

Another participant was of the view that when the provincial government tried to bring reforms in the health sector it was rejected sternly by the majority:

"The current PTI government faced severe backlash and their reforms were opposed due to various reasons. There was a continuous battle in the KPK healthcare sector where changes are being implemented on the model of international health standards and

practices such as practiced in Shaukat Khanum Hospital. These changes were resisted fiercely by many who were contended by the earlier system with no responsibility and accountability." /R5

Similarly, one respondent stated that:

"Yes, it is true, whenever there is a new system comes in place the doctors and healthcare staff always resist it. It is no doubt they don't want to be accountable." /R1

On the other hand, one respondent mentioned the new provincial health act which was fiercely resisted mainly by the doctors:

"When the "Regional and District Health Authorities Act, 2019" in KPK were coming into force which will promote accountability and transparency in health services provision. The act was resisted by doctors to tooth and nail because they did not want to be accountable and transparent to the system while serving people." /R8

Another respondent was of the opinion that:

"Resistance to change by the staff is always there in every organization. If we want to integrate and promote data-driven system in healthcare sector, we will need to consult every stakeholder in order to avoid resistance in the first place." /R11

It is true that there is always a strong lobby and opposition to reforms. In this scenario, one respondent talked about the MTI system:

"Yes, that is true, when the Medical Teaching Institute (MTI) Act 2015 was coming into effect which is a new and modern system of corporate management and autonomy that will ensure the accountability of all employees be it a doctor or hospital staff (paramedics and

nurses), ensuring their performance measurement and attendance based on international

standards on the basis of waiting lists, mortality, morbidity etc. It was indeed ideal for a

common citizen to get better health services, but it was resisted by a strong opposition who

wanted to return to the outdated administrating system." /R14

Stakeholders involvement is key to a successful change in any organization. It was rightly highlighted by one of the respondents:

"In order to avoid strong opposition, the government will need to take all the

stakeholders in the decision-making process." /R2

The above responses show that resistance to change is always there, and it is even stronger in the public sector organizations. There were also several examples of opposition mentioned by the respondents whenever the provincial government had tried to bring reforms in the healthcare sector. The respondents also mentioned that stakeholder's consultation is necessary for bringing data-driven systems in the provincial healthcare sector that will eventually support the big data-driven health policy-making.

5.10.3 Expertise and Infrastructure

Big data provides enormous potential for novel approaches to comprehend human systems, and to predict interactions and nonlinearities between variables. Nevertheless, the development of innovative/creative data management and processing systems is hampered by shortage of data experts, insufficient infrastructure, conventional data analytics, lack of trust and confidence in databases and related expertise. On the other hand, the health sector infrastructure is also inappropriate when it comes to buildings and installation. A comprehensive report on the health sector review of Khyber Pakhtunkhwa by ADB stated that:

Figure 7: Health Infrastructure

Inappropriate infrastructure (buildings and installations). The infrastructure was inappropriate in almost all facilities visited. It included often outdated facilities in terms of (i) standard requirements for adequate spaces (consultation rooms with less than 16 square meters and 2–3 doctors and patients inside, no equipment for clinical examination let alone of privacy); (ii) issues related to infection control (e.g., a tuberculosis lab at the dead-end of a corridor); (iii) long distances between services that should be interlinked and close to each other

Source: Khyber Pakhtunkhwa Health Sector Review: Hospital Care 2019

Similarly, the report identified that despite the thousands of doctors being produced every year, the doctor-patient ratio remains 1 for every 1200 patients. Similarly, there is lack of hospital facilities.

Figure 8: Number of Hospitals, Doctors, Patients & Beds in KPK

Numbers of Hospitals, Doctors, Patients and Beds in KPK

	Public Hospital	Dispensaries	Basic Health Units	Rural Health Centers	MCH Centers	TB Centers	First Aid Points
Numbers of Hospitals	924	4,916	5336	595	1,138	371	1080
Numbers of Patients	11,657	32,897	23,998	1,764	43,900	25,000	20,849
Numbers of Doctors	768	5,638	2,654	675	5,973	1,000	1,200
Numbers of Beds	2,879	348	10,836	1,500	20,600	1,643	65
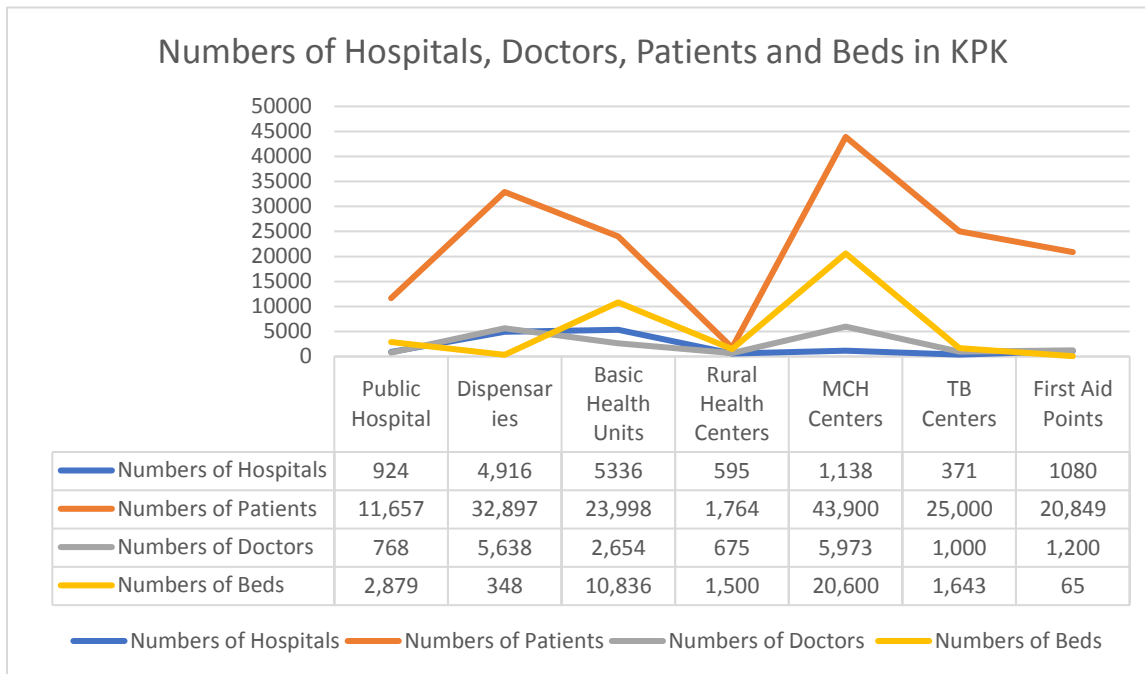
Source: Khyber Pakhtunkhwa Health Sector Review: Hospital Care 2019

On the other hand, there is already a stress on existing public health infrastructure, because majority of the citizens are underprivileged who visits government health facilities to get

treatment. The below chart shows the number of populations covered by each facility where 69 percent of population is covered by public hospitals.

Figure 9: % of Population Covered by Public & Private Hospitals
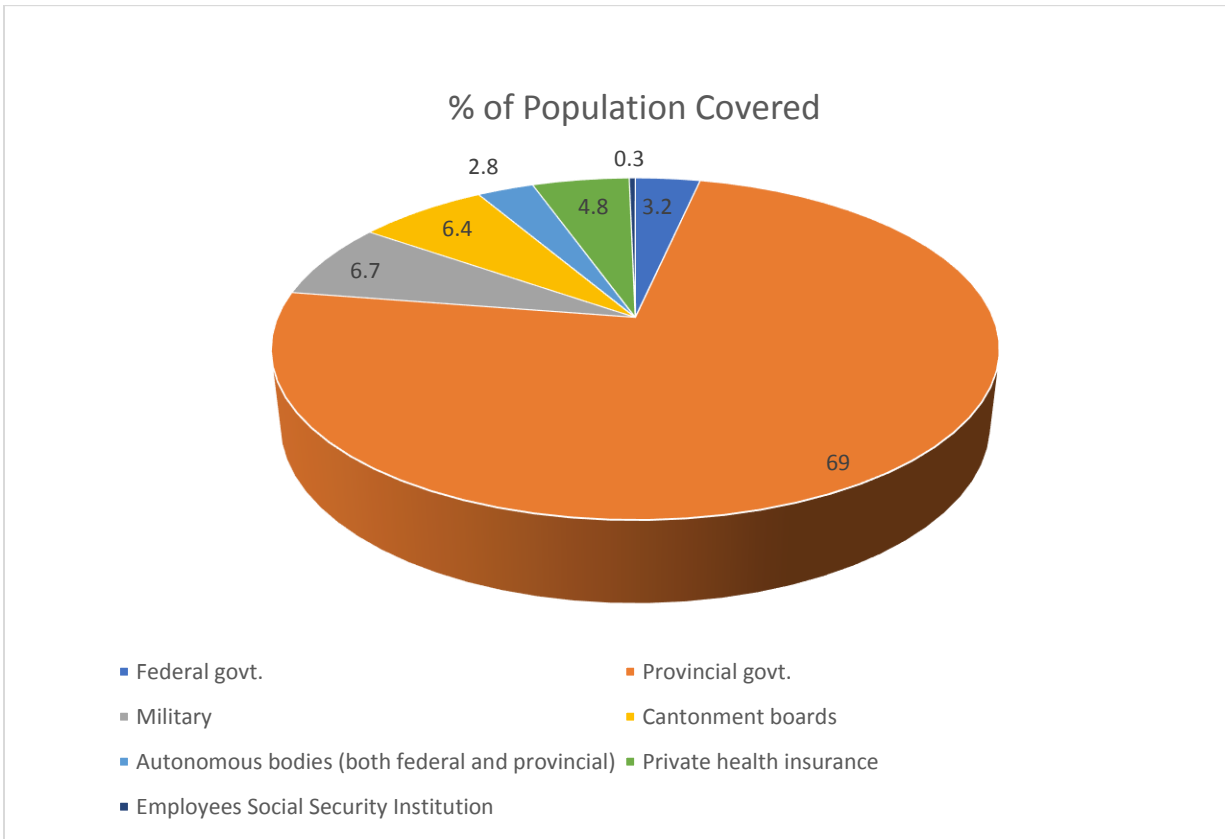
Source: Khyber Pakhtunkhwa Health Sector Review: Hospital Care 2019

The following are the views of respondents when it comes to expertise and infrastructure issues concerning big data and healthcare.

One of the participants stated that:

"Challenges include IT infrastructure, budget allocations by the government and in general comfortableness of doctors and hospital staff in using computers for

documentation and data collection. It also includes the government's facility of providing

relevant software that collects that data in a coherent manner." /R2

On the other hand, one respondent stated that:

"There are some key issues due to which over the past years public health policies are

ineffective in the province such as financial constraints, bureaucratic norms, ineffective

monitoring system and lack of skilled human capital." /R1

Another participant stated that:

"The biggest challenge in my opinion is the insufficient IT infrastructure." /R3

Similarly, one participant was of the view that training of hospital staff is more important:

"In my opinion the main challenge is the untrained hospital staff. In order to bring new

systems, we need to properly train our hospitals staff first." /R5

This statement was supported by another respondent who stated that:

"One of the challenges to big data integration is the lack of appropriate trained staff." /R8

Another participant stated that:

"According to me, trained personnel and required infrastructure are two of the major

issues." /R10

Likewise, one respondent stated that:

> *"There is big data talent gap for which there is a need of proper training and development of skilled personnel." /R12*

One participant posed some thought-provoking questions when it comes to expertise and systems concerning the management of big data, he stated that:

> *The main issue I think would be the data collection and implementation of efficient data collection mechanisms. The problem is that data is out there every interaction, every event generated data but are we recording that data? Or do we have trained people in place to record this data? Or is there any data collection system in place? /R9*

Consequently, it is identified that there is a lack of expertise and infrastructure when it comes to incorporating data-driven technologies in the healthcare sector.

5.10.4 Data Quality and Reliability

Similarly, the quality and reliability of big data is also one of the key concerns when it comes to big data. It is even more tough when we use traditional measure to ensure data reliability of such data. Patient data can also hold large amounts of missing values that will eventually obscure to calculate reliability. Despite the fact that data reliability is extremely important when it comes to analyzing big data for decision making purposes there were only few participants who mentioned the data reliability issue. Their views are quoted below:

"There are various challenges like storage, cleaning, accuracy, security, sharing, and querying and proper data handling." /R11

"There will be difficulty in management of large data sets and ensuring the data is true." /R10

"One of the major challenges in big data is that data analyzing which data is better data and which data is considered bad data, so that it can be sent further for analysis eventually decision making." /R3

"HER or Electronic Health Records is a good source of data, but clinicians are usually busy due to which they miss values, click on incorrect menu items or paste information in wrong patient record…" /R7

"One issue would be to decide which data sets should be discarded, archived and analyzed further." /R13

"Moreover, data from the field of health are often significantly fragmented." /R6

It is therefore true that large amount data does not inevitably means better data. Because the data is not usually collected for analysis purposes. The data is often collected for different purposes such as for billing which may not be utilized for decision making purposes. Thus, a comprehensive monitoring and feedback mechanism is required to enhance the quality of

healthcare services which can analyze/review patient health records on regular basis and thus therefore provide better data for decision making.

5.10.5 Management and Governance

When the amount of digital information and data relating to health rises, so it does it will rise the number of people who access and use this information. Assurance must be provided by applicable legislation such that the data related to personal health must be used properly. Throughout the dynamic legal environment, there is always confusion about "where the data goes", "by whom it is used" and "for what reason". Consequently, there is a need for appropriate data security laws to be enforced to safeguard individual's privacy for the proper management and governance of the data. In the light of the above topic the individuals were asked about the issues concerning data management and governance. Their responses are recorded and quoted below.

One participant was of the view that:

> *"There are several challenges when it comes to the implementation of big data in healthcare particularly with respect to data management and governance. There is a need for relevant laws so that the data could be protected." /R6*

Another participant supported the same while stating that:

> *"The management and governance of data is very crucial especially healthcare data, because healthcare data is very sensitive and currently in our country there is no known regulation to healthcare data management and governance." /R8*

Similarly, one respondent was of the view that:

"I believe one of the main challenges is the governance of health sector data. When data is properly managed under laws and regulations, this protected and secured data can be used for decision making purposes" /R3

Another participant also supported the above statements by stating that:

"In my opinion data ownership and/or governance issues are very much there if we talk about healthcare data." /R4

One participant was of the opinion that along with data governance there are other pertaining issues:

"There are many challenges in implementing big data in healthcare especially in relation to privacy, security, standards, governance."/R14

Another respondent stated that there are no regulations when it comes to private health sector regulation and that is why the data is not shared and eventually not integrated:

"Private sector is not regulated. When we don't have any regulation, therefore they are not bound to provide us any data, nor they are providing…

…They did the disease outbreak investigation in Sindh, but they did not provide the data to the provincial government. You can see how much is the burden. I don't think they are regulated. They are not bound. Government would like to keep them in lope in most of the things, but why they would do, as they don't see any benefit. As by law they are not bound then why they would provide data or be integrated in the process." /R1

Management and governance of healthcare data is the first step towards safety and security of data. For this purpose, there is a need of relevant laws on the provincial level so that individual's privacy should be safeguarded when it comes to proper usage of healthcare data.

5.10.6 Data Privacy and Security

Healthcare data is extremely sensitive as it mainly involves patient's health record. It has been identified that perceived trust and perceived security of information as being the key barriers for users adopting to innovative data-driven systems. Privacy, security and trust were some of the key causes for the big data analytics slow pace adoption in the health sector. Consequently, it is crucial for the health care sector organizations to guarantee data privacy and security and ensure greater trust in adoption of big data analytics. The respondents were asked about data privacy and security issues concerning big data and their responses are quoted below:

"Yes, data privacy and security are some of the major issues in adopting data-driven systems. The government will need to ensure that the data is protected in each step from collection of data to utilizing it for decision making. This will require strict regulations. Perhaps new laws related to data protection could help." /R3

"Whenever we talk about data especially big data, there is always a question of data privacy and data protection. Before we introduce new systems in place, we will have to ensure that patients data is protected." /R5

"There are many implications when it comes to big data. First of all, government is not prepared. It would not be easy for the government because there are data privacy and data sensitivity issues." /R1

"...also, data privacy issues are there because without proper regulations and systems in place there will be data breaches." /R14

"Yes. We might have data breaches which will in return pose risk to patient's privacy." /R9

"...protection of patients data is another issue which is largely debated…" /R2

"There are various challenges like storage, cleaning, accuracy, security, sharing, and querying and proper data handling." /R11

"I think when we are to bring such system in place, we will have to take citizens into account, because after all they are the one who will be most affected by it." /R13

"I believe there is uncertainty of data management, which will eventually result in data breaches. To make sure data is safe and protected, the organizations must ensure data security." /R12

The above responses show that data privacy, security and trust are some of the major issues when it comes to big data analytics in healthcare sector. The respondents stressed on bringing new laws and regulations that will ensure the privacy and protection of user's data.

5.10.7 Financial Constraints

There is always a need of huge finances when it comes to bringing new changes and updating the existing systems in the public sector organizations. Financial constraints are therefore one of the key reasons which may hinder the adoption of big data technologies in the healthcare sector. Consequently, the issue of financial constraints was highlighted by the majority of the participants.

One respondent was of the view that:

"When it comes to spending on health, I believe fiscal space is one of the main issues for any government because they will need to ensure fiscal sustainability while at the same time increasing the spending on health sector." /R4

This was supported by another respondent who stated that:

"...Budget allocation by the government is also one of the main issues for the adoption of big data analytics in the healthcare sector..." /R3

In the same way another participant stated that:

"...lack of funding is another issue concerning big data integration..." /R14

On the other hand, one respondent was of the view that:

"I believe cost related to setting up the data collection mechanism and infrastructure would

be one of the main challenges…"/R10

One participant had expounded on the issue of financial constraints in the health sector. The respondent was of the opinion that the main reason of fragile health system is the budgetary gap between the health issues and the required money for health services:

"Our health system is very fragile; we can't afford to have pandemic like this. We have

epidemics like on daily and monthly basis. Out budget level with health issues are not

matching. We have not really improved on that point. From the last few years our budget

has minimized. The politicians are very smart, they give us different excuses. The revenue

collection is weak, if you can collect well, you will spend well. Money is not there. You are

taking that chunk out of health. So, our budget is not matching with the health issues. We

are not prepared for any epidemic. And this a pandemic. We are very match lacking

capacity, because we don't have enough money to spend. For quality services we need

money to pay for them and keep that role which we don't have. There are capacity issues.

Provinces are responsible for their own, they need to play proactive role. We need reforms.

Now we have four health systems. Their analysis would be good as well, still there are

budgetary issues." /R1

It is true when it comes to the budget allocation for health. Although in the epic time where the province was faced with the Covid-19 pandemic, the government allocated 124 billion rupees for health sector. On the other hand, the overall health expenditure has been increased to 4 percent of the total provincial budget. It is a good sign as the provincial government is considering health sector a top priority.

5.10.8 Government Priorities, Power Dynamics, Lack of Capacity and Political Will

There were few other challenges / hurdles that were highlighted by the respondents when it comes to incorporating big data in the health policy formulation process. These challenges were relevant in the context of this region.

One respondent provided a broad view of the current health policymaking practices by stating that:

"There are multiple hurdles, there are certain people who involve in the policymaking process are the ones who develop it, do the initial drafts and they don't share it with other relevant institutions who have stake in it. but the main issue is, there is one institution who is drafting the policy without considering any relevant stakeholders in the policymaking process. This is one of the limitations that the process is not consultative." /R1

Another participant was of the opinion that:

"It is usually the dynamics of the power, top-down approach is very much there, also we do have capacity issues, we lag that part, we don't have will, for instance if we bring the evidence in the policy, the policy will look different. Our priorities are not based on the facts." /R4

Similarly, one of the respondents stated that:

"There is lack of political will that does not allow for innovation and transformation in the healthcare sector." /R5

Another participant also stated the same:

"There is lack of political will from the government side". /R1

In the same way, another participant was of the view that:

"A major hurdle is the nature of decisions and data itself. Health care decisions require timely information and actions. Other barriers may include data conventions, and institutional practices." /R11

Nevertheless, one participant was of the view that despite capacity issues the country is currently going through a transformative phase and innovation is taking place when it comes to the collection and use of data:

"We do have capacity issues, but a government vision is there, we do have the manpower at the federal level, and also at the provincial level, and there are innovation taking place in the health sector, they continuously collect data under certain programs for specific issues. Now this platform (big data) will provide an opportunity to integrate the existing data and to collect the data and manage the data. in terms of capacity, despite we have challenges, still we have a lot of workforce involve in the collection of data." /R3

The above responses show that government lags behind several issues such as government priorities, power dynamics, lack of capacity and political will when it comes to bringing new systems in place which will support healthcare and evidence-based decision making.

5.11 Health Data Cooperative Model (HDC) for Policy Empowerment
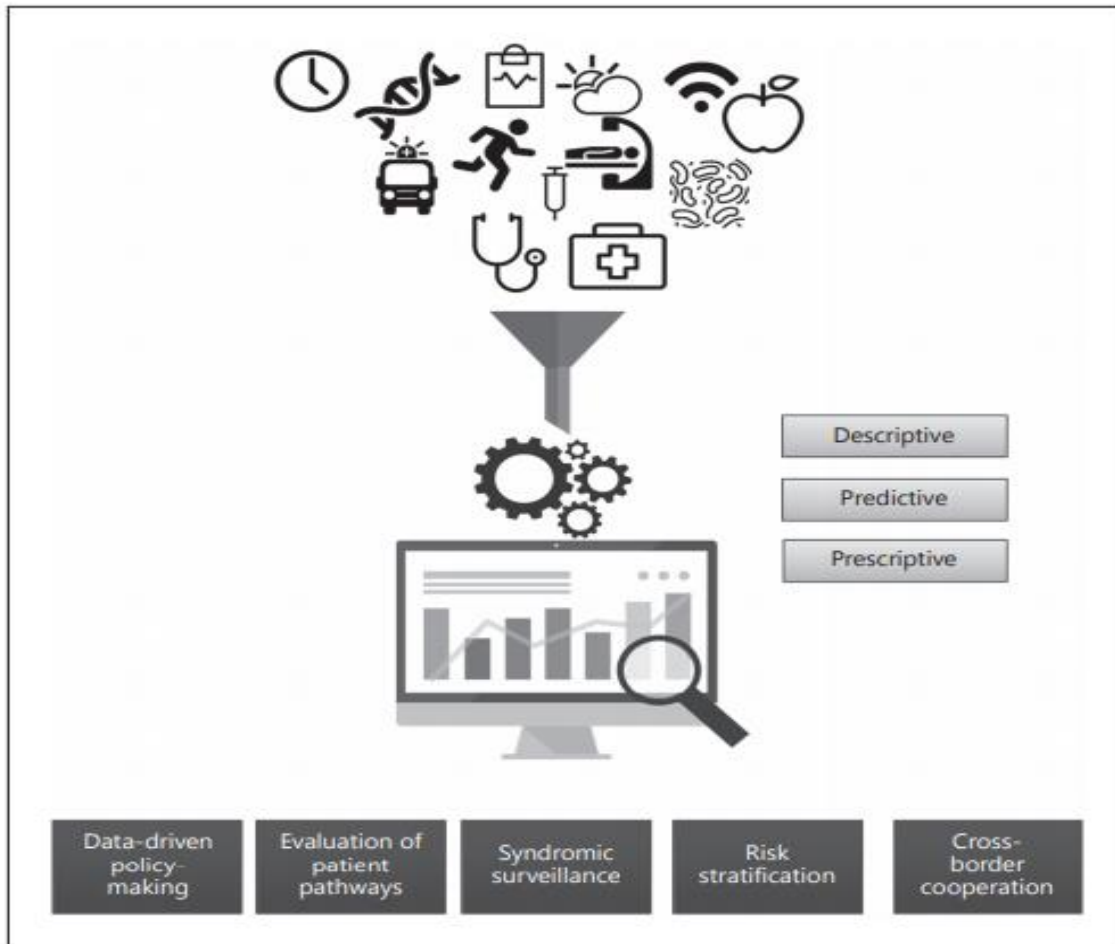
Figure 10: Health Data Cooperative Model(Mählmann et al., 2017)

This study proposes the Health Data Cooperative (HDC) as a needed model to overcome the hurdles and challenges of integrating big data into public health policymaking. The HDC model is recommended by several authors and it is the considered the most suitable model for policy empowerment. The rapid advancements in technology coupled with the declining costs of computer power and data storage and at the same time the improved security of cloud-based storage and analysis systems, make an ideal environment to develop HDC (Mählmann et al., 2017).

HDC are comprised of centralized data structures that facilitate data linkage of and data access to heterogeneous data from diverse sources inside and beyond the health field. HDC offers a roadmap to address technological, political and social hurdles/challenges and empower policymaking by incorporating and implementing big data. HDC model is largely focused on metadata and has an innately patient/citizen-centered trust model design for the governance of health data. In this manner, HDC facilitate the inclusion of heterogenous health data of citizens, along with the linking to information published in databases, guidelines and articles, as a result allowing value and insights to be generated via an analytics system that can be placed on top of any form of data governance model (Hafen, 2014). Therefore, the Health Data Cooperative model tackles all the above-mentioned hurdles and proposes and holistic solution/answer within a single framework.

As shown in Figure 5, Health Data Cooperative (HDC) is an ecosystem where data, materials, knowledge and services of stakeholders are incorporated in a cloud-based platform. Firstly, their heterogenous elements may contain data sources for example electronic medical record (EMR), nonelectronic systems data, citizen and patient m-health data, proteomic and genomic data, research data, biological material repositories, biostatistics, sensor data, social media data, environmental data and more. These varied data sources are comprehensively processed, examined and mapped in order to develop a web of knowledge and actionable insights. The analytical models that are developed will be continuously re-defined, new requisites shall be identified and new data sources of data acquirement will be captured. "The aims of analysis are to obtain descriptive analytics to study causal mechanisms for the current state, predictive analytics to improve prognoses on the spatial and temporal development of ill-health and disease, and, if possible, prescriptive analytics to inform policy-making on how to decide on shaping the future

(Hafen, 2014)." The dispersed character of big data needs technical solutions and considerable investments.

Consequently, in order to accomplish the challenge of evidence-based health policymaking, large and varied data sets must be converted into actionable knowledge to policymakers for planning and implementing healthcare programs. Through HDC, health policymaking will not only be evidence-based but also data-driven. A big change awaits the onset of health policymaking.

5.12 Conclusion

The analysis and findings chapter confirmed the current concepts developed earlier in the conceptual framework concerning hurdles and challenges pegged in the way of inclusive big data analytics and big data health policy-making such as, Interoperability, Data Integration and Data Sharing, Resistance Towards Change, Expertise and Infrastructure, Data Quality and Reliability, Management and Governance, Data Privacy and Data Security. New concepts were also emerged during the collection of data and responses from the participants such as, Financial Constraints, Government Priorities, Power Dynamics, Lack of Capacity and Political Will. The preliminary concepts were confirmed from the respondents through their input. Consequently, health data cooperative (HDC) model was proposed in order to accomplish the hurdles and challenges that big data possess for data-driven and evidence-based health policymaking.

CHAPTER 6

CONCLUSION & RECOMMENDATIONS

6.1 Introduction

In the preceding chapter the data and responses collected through interviews and secondary sources confirmed the already developed conceptual framework. The findings were thoroughly examined and analyzed, and with the help of participant responses and secondary sources the conceptual framework was further developed. This chapter summarizes the study on the research outcomes, accomplishment of objectives and recommendations. Research limitations were provided, and future research directions were recommended.

6.2 Research Outcomes

After the thorough analysis and discussion on the already established concepts concerning hurdles and challenges pegged in the way of inclusive big data analytics and big data health policy making. It has been identified that there are many hurdles and challenges which hinders the inclusion of big data in the policy formulation process. The study also offered recommendations on how to overcome these hurdles in the healthcare sector so that data and evidence-based health policies could be encouraged.

The available literature on the hurdles and challenges concerning big data inclusion in the health policymaking was based on Interoperability, Data Integration and Data Sharing, Resistance Towards Change, Expertise and Infrastructure, Data Quality and Reliability, Management and Governance, Data Privacy and Data Security. This research offered a conclusive conceptual

framework on hurdles and challenges pegged in the way big data analytics and big data health policy making. The new concepts which were identified during the data collection process, such as Financial Constraints, Government Priorities, Power Dynamics, Lack of Capacity and Political Will. These new concepts were not mentioned or studied before. Following a thorough review of the literature seven concepts were identified concerning hurdles and challenges to big data integration in health policymaking which are Interoperability, Data Integration and Data Sharing, Resistance Towards Change, Expertise and Infrastructure, Data Quality and Reliability, Management and Governance, Data Privacy and Data Security.

After the collection of data, new concepts were emerged such as, Financial Constraints, Government Priorities, Power Dynamics, Lack of Capacity and Political Will. The investigator claimed that it may be due to the contextual discrepancies among various studies that contributed to the concepted being identified.

The first overarching challenge and hurdle that is being identified and widely mentioned by the respondents is the big data integration into the health policymaking to take advantage of the data and generate knowledge and insight for public health. This also involves the issues concerning data interoperability and data sharing among different organization. When the organizations work in silos and there are no mechanisms and regulations for data integration among organizations, it makes it difficult for experts to make informed decisions without having complete data. Another major hurdle that was discussed is the expertise and infrastructure. The health sector is currently in the transformative phase but when it comes to big data, we lack the required expertise and infrastructure.

On the other hand, there is a huge challenge of data quality and reliability. Large amount data does not inevitably mean better data. Because the data is not usually collected for analysis purposes. The data is often collected for different purposes such as for billing which may not be utilized for decision making purposes. Thus, a comprehensive monitoring and feedback mechanism is required to enhance the quality of healthcare services which can analyze/review patient health records on regular basis and thus therefore provide better data for decision making. Similarly, management and governance of data is another major obstacle when it comes to big data in healthcare. When the amount of digital information and data relating to health rises, so it does it will rise the number of people who access and use this information.

Similarly, data privacy and data security are another major challenge. Healthcare data is extremely sensitive as it mainly involves patient's health record. It has been identified that perceived trust and perceived security of information as being the key barriers for users adopting to innovative data-driven systems. On the other hand, there is always a need of huge finances when it comes to bringing new changes and updating the existing systems in the public sector organizations. Financial constraints are therefore one of the key reasons which may hinder the adoption of big data technologies in the healthcare sector. There are several other obstacles such as, government priorities, power dynamics, lack of capacity and political will. Government lags behind several issues such as government priorities, power dynamics, lack of capacity and political will when it comes to bringing new systems in place which will support evidence-based decision making.

6.3 Achievements of Objectives

The objectives of the research presented in chapter one was accomplished in the following way:

- In order to formulate the conceptual framework on the hurdles and challenges pegged in the way of inclusive big data analytics and big data health policy making, literature applicable to the research area was comprehensively and critically reviewed. The gap in knowledge was identified, which was explored further in the study.

- The conceptual framework was developed with the assistance of concepts drawn from a thorough review of the relevant literature. The methodology was designed in a way that would further develop the conceptual framework.

- Relevant data was collected through using both secondary sources existing database like articles, journals, governmental databases, statistics bureaus/surveys and hospital archives. Whereas, the primary data was collected through the process of face to face and telephonic interviews (Observational data). The primary data was then reduced, transcribed and interpreted. The conceptual framework which was already developed was examined via findings.

- With the help of data analysis and new findings the revised conceptual framework was developed.

6.4 Recommendations

These recommendations are drawn from the aforementioned discussions, analysis and conclusion. This will help the decision makers to take advantage of the new data paradigm in the best way possible. It will also help the decision makers to overcome these challenges and hurdles

which hinders the inclusion of data and evidence in the decision-making process and become part of the transformative process.

- In order to address the issue of data integration and data linkage a robust platform should be developed that will collect heterogenous cross-border data sets, along with a strong interface for analysis so that it can support near/real-time data analysis and to provide accessible and usable data and evidence to policy makers. To realize this, the public and private health service providers must work together to ensure smooth cross-border flow of data.

- This study proposes the application of Hadoop and HDFs (Hadoop File System) tools for implementation which is widely mentioned by authors to perform the analytics on the big data that is generated. Hadoop is an open-source and open access software designed to store data and run application on commodity hardware clusters. In addition, there are other tools which can be used such as, cloud databases of large organizations (such as IBM and Intel) that use SaaS-software as service facility and the OpenRefine service, a robust tool/software that is deigned to clean up and transform large data sets for informed decision making.

- In order to find the proper balance amid the aims of providing data for public health policymaking and concerns related to individual privacy. New ways to connect and incorporate separate data sets pose challenges to personal privacy, and thus involve highly reliable organizations and approaches to good governance. Therefore, legislation concerning data management and governance is highly important on provincial and national level so that data is carefully managed and protected.

- The government can agree on international technical standards in this context, while taking into account transparency and openness which is a deemed as a prerequisite framework for digital transformation. Furthermore, innovative methods need to be developed to turn the large volume of data into useful knowledge that can be accessed by healthcare practitioners. Additional work will be carried out to ensure the information is more understandable and accessible for the doctors and health practitioners.

- To accomplish this, current education and training programs for health practitioners must incorporate data management/handling issues into the curriculum to make sure the requisite competencies and skills are developed. The training of new generation of healthcare practitioners should be made a top priority.

- The government of Khyber Pakhtunkhwa is already on a road to transformation with its healthcare initiatives such as, the Sehat Sahulat Programme and the Sehat Insaf Programme which is providing universal healthcare to every family in the province. It is therefore a wonderful opportunity for the provincial government to bring transformation in the health infrastructure and expertise so that it can handle big data. This will help the government to make smart and targeted health policies without wasting a lot of resources.

6.5 Research Limitations

- For the collection of primary data, the identification and access to relevant participants were either denied or limited. The author also faced difficulties while convincing participants for interview.

- Similarly, due to Covid-19 epidemic it was difficult to carry out interviews in person, therefore, majority of respondents agreed to write their responses and some interviews were conducted via phone.

CHAPTER NO 6: CONCLUSION & RECOMMENDATIONS 97

- Correspondingly, the author also faced uncertain environment due to Covid-19 due to which the study took more time to complete then it planned. There were also resource constraints.

- Conversely, the prescribed model is only limited to health sector and can only be applied to healthcare policy issues. Also, the data was collected from prospects of health policymaking. It contains the opinions of public policy experts, public health experts and big data experts. It does not include the views and opinions of medical professionals and consumers of healthcare system.

6.6 Future Research

This study proposed a model for the implementation and capacity assessment for the adoption of big data driven health policy making process. The author suggests conducting a survey based on the model proposed in the study at national and provincial level concerning the topics of big data and its application in healthcare and health policymaking process.

6.7 Summary

This chapter concludes the study by briefly summarizing the research findings, the outcomes and accomplishment of objectives. The recommendations, research limitations and direction for future research were also presented.

BIBLIOGRAPHY

Almeida, F., & Santos, M. (2014). A conceptual framework for big data analysis. In Organizational, legal, and technological dimensions of information system administration (pp. 199-223). IGI Global.

Andrews, L.; Higgins, A.; Andrews, M. W. and Lalor, J.G. (2012), Classic Grounded Theory Beyer, M.A. and Laney, D., 2012. The importance of 'Big data': a definition. *Stamford, CT: Gartner*, pp.2014-2018.

Agrawal, R., & Prabakaran, S. (2020). Big data in digital healthcare: lessons learnt and recommendations for general practice. Heredity, 124(4), 525-534.

Abdulai, J. (2018). Redefining the Concept of Big Data: A Ghanaian Perspective.

Bosch-Capblanch, X., Lavis, J. N., Lewin, S., Atun, R., Røttingen, J. A., Dröschel, D., ... & Haines, A. (2012). Guidance for evidence-informed policies about health systems: rationale for and challenges of guidance development. PLoS Med, 9(3), e1001185.

Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. Information, Communication & Society, (2012), 662–679. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/1369118X.2012.678878

Braunstein, M. L. (2015). Practitioner's Guide to Health Informatics (pp. 1-162). Springer.

Bates, D. W. (2005). Physicians and ambulatory electronic health records. Health affairs, 24(5), 1180-1189.

Brownson, R. C., Chriqui, J. F., & Stamatakis, K. A. (2009). Understanding evidence-based public health policy. American journal of public health, 99(9), 1576-1583.

Buse, K., Dickinson, C., Gilson, L., & Murray, S. F. (2009). How can the analysis of power and process in policy-making improve health outcomes. The Official Journal of the International Hospital Federation, 45(1), 14.

Beaver, D., Kumar, S., Li, H. C., Sobel, J., & Vajgel, P. (2010, October). Finding a Needle in Haystack: Facebook's Photo Storage. In OSDI (Vol. 10, No. 2010, pp. 1-8).

Buse, K., Mays, N., & Walt, G. (2005). Making health policy (understanding public health). UK: Bell & Brain Ltd.

Beyer, M. A., & Laney, D. (2012). The importance of 'big data': a definition. Stamford, CT: Gartner, 2014-2018

Collins, C. D., Omar, M., & Tarin, E. (2002). Decentralization, health care and policy process in the Punjab, Pakistan in the 1990s. The International journal of health planning and management, 17(2), 123-146.

Cottle, M., Hoover, W., Kanwal, S., Kohn, M., Strome, T., & Treister, N. (2013). Transforming Health Care Through Big Data Strategies for leveraging big data in the health care industry. Institute for Health Technology Transformation, http://ihealthtran. com/big-data-in-healthcare.

Dey, N., Ashour, A. S., & Bhatt, C. (2017). Internet of things driven connected healthcare. In Internet of things and big data technologies for next generation healthcare (pp. 3-12). Springer, Cham.

Dash, S., Shakyawar, S. K., Sharma, M., & Kaushik, S. (2019). Big data in healthcare: management, analysis and future prospects. Journal of Big Data, 6(1), 1-25.

Dodhy, M. (2019, May 1) *Pakistan ministries are upgrading to the E-Governance model.* Retrieved from: https://www.techjuice.pk/13-pakistan-ministries-are-upgrading-to-the-egovernance-model/.

Dunleavy, P., Margetts, H., Bastow, S., & Tinkler, J. (2005). New public management is dead. Long live digital-era governance. Journal of Public Administration Research and Theory, 16, 467–494.

Dunleavy, P., Margetts, H., Bastow, S., & Tinkler, J. (2006). Digital Era Governance: IT Corporations, the State, and E-Government. New York: Oxford University Press.

El-Jardali, F., Lavis, J. N., Ataya, N., & Jamal, D. (2012). Use of health systems and policy research evidence in the health policymaking in eastern Mediterranean countries: views and practices of researchers. Implementation science, 7(1), 1-16.

Fife, E., & Orjuela, J. (2012). The privacy calculus: Mobile apps and user perceptions of privacy and security. International Journal of Engineering Business Management, 4(Godište 2012), 4-11.

Financial Tribune. (2020, December 24). E-Health File for 66m Iranians. Available at: https://financialtribune.com/articles/people/64502/e-health-files-for-66m-iranians.

Filmer, D., Hammer, J. S., & Pritchett, L. H. (2000). Weak links in the chain: a diagnosis of health policy in poor countries. The World Bank Research Observer, 15(2), 199-224.

Forschungs, G. Ö., (2016). Study on Big Data in, Brussels: EUROPEAN COMMISSION.

Gartner. (2019, May 11) *Hype cycle for emerging technologies identifies "tipping point" technologies that will unlock long-awaited technology scenarios.* Stamford (CT): Gartner Inc.; c2012. Available from: http://www.gartner.com/newsroom/id/2124315.

Ghaffar, A., Gilson, L., Tomson, G., Viergever, R., & Røttingen, J. A. (2016). Where is the policy in health policy and systems research agenda?. Bulletin of the World Health Organization, 94(4), 306.

Gilson, L., & Raphaely, N. (2008). The terrain of health policy analysis in low and middle income countries: a review of published literature 1994–2007. Health policy and planning, 23(5), 294-307.

Go, S. J., & Jeoung, Y. H. (2012). Health risk prediction using big health data. In Health Welf Policy Forum (Vol. 193, pp. 43-52).

Gawande, A., (2018) Why doctors hate their computers. The New Yorker, 12 https://www.newyorker.com/magazine/2018/11/12/why-doctors-hate-their-computers

Goodhue, D. L., & Thompson, R. L. (1995). Task-technology fit and individual performance. MIS quarterly, 213-236.

Greenhalgh, T., & Russell, J. (2006). Reframing evidence synthesis as rhetorical action in the policy making drama. Healthcare Policy, 1(2), 34.

Greenhalgh, T., & Russell, J. (2006). Reframing evidence synthesis as rhetorical action in the policy making drama. Healthcare Policy, 1(2), 34.

Groves, P., Kayyali, B., Knott, D., & Kuiken, S. V. (2016). The'big data'revolution in healthcare: Accelerating value and innovation.
Yan, X., Song, T., & Wu, Q. (2017). An improved cultural algorithm and its application in image matching. Multimedia Tools and Applications, 76(13), 14951-14968.

Hafen, E., Kossmann, D., & Brand, A. (2014). Health data cooperatives-citizen empowerment. Methods Inf Med, 53(2), 82-86.

Hawkins, B., & Parkhurst, J. (2016). The'good governance'of evidence in health policy. Evidence & policy: a journal of research, debate and practice, 12(4), 575-592.

Heart, T. (2010). Who is out there? Exploring the effects of trust and perceived risk on SaaS adoption intentions. ACM SIGMIS Database: the DATABASE for Advances in Information Systems, 41(3), 49-68.

Ishikawa, K. B. (2016). Medical big data for research use: current status and related issues. Japan Medical Association Journal: JMAJ, 59(2-3), 110.

Innvaer, S., Vist, G., Trommald, M., & Oxman, A. (2002). Health policy-makers' perceptions of their use of evidence: a systematic review. Journal of health services research & policy, 7(4), 239-244.

Junglas, I. A., & Watson, R. T. (2008). Location-based services. Communications of the ACM, 51(3), 65-69.

Jain, P., Gyanchandani, M., & Khare, N. (2016). Big data privacy: a technological perspective and review. Journal of Big Data, 3(1), 1-25.

Jetzek, T. (2016). Managing complexity across multiple dimensions of liquid open data: The case of the Danish basic data program. Government Information Quarterly, 33(1), 89–104.

Kabir, A. I., Karim, R., Hossain, M. I., & Newaz, S. (2018). Prospects and challenges of using big data in healthcare sector of Bangladesh: focus on the reformation of the healthcare system. Journal of Asian and African Social Science and Humanities, 4(1), 1-16.

Kayyali, B., Knott, D., & Van Kuiken, S. (2013). The big-data revolution in US health care: Accelerating value and innovation. Mc Kinsey & Company, 2(8), 1-13.
Khan, A. (1996). Policy-making in Pakistan's population programme. Health Policy and Planning, 11(1), 30-51.

Khan, M. M., & Van den Heuvel, W. (2007). The impact of political context upon the health policy process in Pakistan. Public Health, 121(4), 278-286.

Kim, M. K., & Park, J. H. (2017). Identifying and prioritizing critical factors for promoting the implementation and usage of big data in healthcare. Information Development, 33(3), 257-269.

Klievink, B., & Janssen, M. (2009). Realizing joined-up government: Dynamic capabilities and stage models for transformation. Government Information Quarterly, 26(2), 275–284.

Klievink, B., Romijn, B., Cunningham, S., & de Bruijn, H. (2016). Big data in the public sector: Uncertainties and readiness. Information Systems Frontiers, 19(2), 267–283.

Koch, S. (2013). Achieving holistic health for the individual through person-centered collaborative care supported by informatics. Healthcare informatics research, 19(1), 3.

Kwon, T. H., Kwak, J. H., & Kim, K. (2015). A study on the establishment of policies for the activation of a big data industry and prioritization of policies: Lessons from Korea. Technological Forecasting and Social Change, 96, 144-152.

Latif, Z., Jianqiu, Z., Ullah, R., Pathan, Z. H., & Latif, S. (2017). Application of optical frequency comb in high-capacity long distance optical communication for China-Pakistan economic corridor. Journal of Optical Communications, 38(3), 331-340.

Latif, Z., Tunio, M. Z., Pathan, Z. H., Jianqiu, Z., Ximei, L., & Sadozai, S. K. (2018, March). A review of policies concerning development of big data industry in Pakistan: Subtitle: Development of big data industry in Pakistan. In 2018 international conference on computing, mathematics and engineering technologies (iCoMET) (pp. 1-5). IEEE.

Lawrence, N. R., & Bradley, S. H. (2018). Big data and the NHS–we have the technology, but we need patient and professional engagement. Future healthcare journal, 5(3), 229.

Lavis, J. N., Lomas, J., Hamid, M., & Sewankambo, N. K. (2006). Assessing country-level efforts to link research to action. Bulletin of the World health Organization, 84, 620-628.

Luo, J., Wu, M., Gopukumar, D., & Zhao, Y. (2016). Big data application in biomedical research and health care: a literature review. Biomedical informatics insights, 8, BII-S31559.

Lawrence, N. R., & Bradley, S. H. (2018). Big data and the NHS–we have the technology, but we need patient and professional engagement. Future healthcare journal, 5(3), 229.

Mählmann, L., Reumann, M., Evangelatos, N., & Brand, A. (2017). Big data for public health policy-making: Policy empowerment. Public health genomics, 20(6), 312-320.

McKinsey, G. I. (2011). Big data: The next frontier for innovation, competition, and productivity. McKinsey Global Institute, 1-6.

Macaulay, T. A. (2016). Progress towards a paperless NHS. BMJ: British Medical Journal, 355.

Murad, M. A. U., (2020, December 18). Big Data Analytics in Bangladesh's Healthcare. [Online] Available at: https://www.daily-sun.com/post/498105/Big-Data-Analytics-in-Bangladesh%E2%80%99s-Healthcare

Mahmood, T., Iqbal, T., Amin, F., Lohanna, W., & Mustafa, A. (2013, December). Mining Twitter big data to predict 2013 Pakistan election winner. In INMIC (pp. 49-54). IEEE.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Hung Byers, A. (2011). Big data: The next frontier for innovation, competition, and productivity. McKinsey Global Institute.

Margetts, M., & Dunleavy, P. (2013). The second wave of digital-era governance: a quasi-paradigm for government on the Web. Philosophical Transactions of the Royal Society A, 371, 1–17.

National Information Strategy Committee of Korea. Implementation of smart government using Big data [Internet] Seoul: National Information Strategy Committee; 2012. [cited at 2019 May 5]. Available from:http://mmlab.snu.ac.kr/courses/2011_information-universe/presentation/1121.pdf.

Nishtar, S., Boerma, T., Amjad, S., Alam, A. Y., Khalid, F., ul Haq, I., & Mirza, Y. A. (2013). Pakistan's health system: performance and prospects after the 18th Constitutional Amendment. The Lancet, 381(9884), 2193-2206.

Otake, T., (2017). Medical big data to be pooled for disease research and drug development in Japan.[Online] Available at: https://www.japantimes.co.jp/news/2017/05/15/reference/medical-big-data-pooled-disease-research-drug-development-japan/

Pakistan. Ministry of Information Technology and Telecommunication. (2018, June 14). Digital Pakistan Policy-Gateway to Pakistan. Available at: http://moib.gov.pk/Downloads/Policy/DIGITAL_PAKISTAN_POLICY(22-05-2018).pdf

Parkhurst, J. (2017). The politics of evidence: from evidence-based policy to the good governance of evidence (p. 182). Taylor & Francis.

Patrizio, A. (2018, December 3). Network World. Retrieved from Network World Web site: https://www.networkworld.com/article/3325397/idc-expect-175-zettabytes-of-data-worldwide-by-2025.html

Petersen, G. S. (1997). High-impact sales force automation: A strategic perspective. St. Lucie Press.

Roy, T. (2018). Implementing electronic government procurement (e-GP) in Bangladesh: a study on 10 district LGED offices (Doctoral dissertation, BRAC University).

Rahman, N. (2016). Factors affecting big data technology adoption.

Reimsbach-Kounatze, C., & Van Alsenoy, B. (2013). Exploring data-driven innovation as a new source of growth. Organization for Economic Co-operation and Development.

Richterich, A. (2016). Using transactional big data for epidemiological surveillance: google flu trends and ethical implications of 'infodemiology'. In The Ethics of Biomedical Big Data (pp. 41-72). Springer, Cham.

Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: promise and potential. Health information science and systems, 2(1), 1-10.

Rose, N. (1999). Powers of Freedom: Reframing Political Thought. Cambridge: Cambridge University Press.

Ritchie, J., Lewis, J. and Elam, G., (2003). *Designing and selecting samples* (pp. 77-108). London: Sage.

Singapore: National Security Coordination Secretariat. (2019, May 8). *Risk Assessment and Horizon Scanning (RAHS)* [Internet]. Available from: http://app.nscs.gov.sg/public/home.aspx.

Sharma, M., Gupta, R., & Acharya, P. (2020). Prioritizing the critical factors of cloud computing adoption using multi-criteria decision-making techniques. Global Business Review, 21(1), 142-161.

Song YJ. (2012). Big data world! Evolution of SNS and the public policy. IT Future Strategy. (13):1–26.

Shiffman, J., & Smith, S. (2007). Generation of political priority for global health initiatives: a framework and case study of maternal mortality. The lancet, 370(9595), 1370-1379.

Siddiqi, S., Haq, I. U., Ghaffar, A., Akhtar, T., & Mahaini, R. (2004). Pakistan's maternal and child health policy: analysis, lessons and the way forward. Health policy, 69(1), 117-130.

Sippe, R. (2015). The relation between big data and informational privacy in the context of the healthcare.

Smith, K. E., & Joyce, K. E. (2012). Capturing complex realities: understanding efforts to achieve evidence-based policy and practice in public health. Evidence & Policy: A Journal of Research, Debate and Practice, 8(1), 57-78.

Song, T. M., & Ryu, S. (2015). Big data analysis framework for healthcare and social sectors in Korea. Healthcare informatics research, 21(1), 3.

Song, T. M. (2012). Efficient utilization of big data on healthcare and welfare area. In Healthc Welf Forum (Vol. 193, pp. 68-76).

Turnbull, C., Scott, R. H., Thomas, E., Jones, L., Murugaesu, N., Pretty, F. B., ... & Caulfield, M. J. (2018). The 100 000 Genomes Project: bringing whole genome sequencing to the NHS. Bmj, 361.

Tabrizi, N., (2020, December, 19). Big Data and Healthcare in the Global South. [Online] Available at: https://ict4dblog.wordpress.com/2018/08/31/big-data-and-healthcare-in-the-global-south/

Topol, E., (2019) The topol review: preparing the healthcare workforce to deliver the digital future. Health Education England https://topol.hee.nhs.uk

The Economist. (2019, May 12). *Data, data everywhere* [Internet] London: The Economist; c2010. Available from: http://www.economist.com/node/15557443/ .

The UK Government. (2019, May 10). *The Horizon Scanning Programme Team* [Internet] London; 2014. Available from: https://www.gov.uk/government/groups/horizon-scanning-centre.

Vayena, E., Haeusermann, T., Adjekum, A., & Blasimme, A. (2018). Digital health: meeting the ethical and policy challenges. Swiss medical weekly, 148, w14571.

Walt, G., & Gilson, L. (1994). Reforming the health sector in developing countries: the central role of policy analysis. Health policy and planning, 9(4), 353-370.

BIBLIOGRAPHY   105

Walt, G., Shiffman, J., Schneider, H., Murray, S. F., Brugha, R., & Gilson, L. (2008). 'Doing'health policy analysis: methodological and conceptual reflections and challenges. Health policy and planning, 23(5), 308-317.

Wang, E. S. T., & Lin, R. L. (2017). Perceived quality factors of location-based apps on trust, perceived privacy risk, and continuous usage intention. Behaviour & Information Technology, 36(1), 2-10.

Warf, B. (2016). e-Government in Asia: Origins, politics, impacts, geographies. Chandos Publishing.

World Bank. (2016). *The World by Region: World Development Indicators.* Washington DC: The World Bank.

Younas, N., Asghar, Z., Qayyum, M., & Khan, F. (2016, September). Education and socio-economic factors impact on earning for Pakistan-A bigdata analysis. In International Conference on Future Intelligent Vehicular Technologies (pp. 215-223). Springer, Cham.

APPENDICES

Appendix A: Letter of Introduction to Interviewees

<div align="right">

Shahid Umar
School of Public Policy,
Pakistan Instituted of Development
Economics (PIDE), Islamabad

</div>

TO WHOM IT MAY CONCERN

Dear Sir/ Madam,

RE: Invitation to Participate in a Research Study

I am a student enrolled in the Department of Public Policy at Pakistan Institute of Development Economics (PIDE) Islamabad. My supervisor has recommended your name, and I would be most grateful if you could provide some time from your busy schedule.

We are conducting interviews as part of a research study to increase our understanding of the topic "Investigating the Adoption of Big Data Analytics in Healthcare: The Moderating Role of Behavior Towards Change". As your expertise and vast knowledge in the above field makes you an ideal position to give us valuable firsthand information from your own perspective.

The interview takes around 30 to 40 minutes and is very informal. We are simply trying to capture your thoughts and perspectives on the above subject. Your responses to the questions will be kept confidential. Each interview will be assigned a number code to help ensure that personal identifiers are not revealed during the analysis and write up of findings. You will receive a copy of the final paper upon request.

There is no compensation for participating in this study. However, your participation will be a valuable addition to our research and findings could lead to a greater understanding of the challenges to big data analytics in healthcare and health policy formulation process which may help policy makers in the future.

If you are willing to participate please suggest a day and time that suits you and I'll do my best to be available. If you have any questions please do not hesitate to ask.

Thanks!

Sincerely,

Shahid Umar
PIDE School of Public Policy,
Islamabad.

Appendix B: Informed Consent Form

**Pakistan Institute of Development Economics (PIDE),
Department of Public Policy**

Consent to Participate in a Research Study

Title of the Study: "Investigating the Adoption of Big Data Analytics in Healthcare: The Moderating Role of Behavior Towards Change"

Faculty Advisors: Dr. Attiya Javid & Dr. Fazli Hakim Khattak

Investigator/Researcher: Shahid Umar (School of Public Policy)

Introduction:

You are cordially invited to participate in a research study conducted by Shahid Umar (Researcher) and Dr. Attiya Javid & Dr. Fazli Hakim Khattak (Faculty Advisors) at the Pakistan Institute of Development Economics (PIDE) Islamabad, because you are a part of this sector. We believe that your experience and insight about the topic will enhance the quality of this research. This research is only for academic purpose, and your participation is completely voluntary.

Purpose of the Study:

The intent of the proposed study is to identify hurdles and challenges pegged in the way of inclusive big data analytics in healthcare and health policy-making. The study aims to devise an extensive conceptual framework for its integration in public health policy formulation in Khyber Pakhtunkhwa. The study will examine the challenges, prospects and policy implications that big data possess for data-driven and evidence-based health policymaking in Khyber Pakhtunkhwa province of Pakistan.

Study Procedure:

The research requires your participation through interview with the help of open-ended questions, which will take 25 to 30 minutes at any comfortable place. If you do not wish to answer any of the questions during the interview, you may say so and the interviewer will move on to the next

question. No one else but the interviewer will be present unless you would like someone else to be there.

Confidentiality

The records of this study will be kept strictly confidential, and no one else except the researcher will access to the information documented during your interview. The entire interview will be tape-recorded, but no-one will be identified by name on the tape. Tape records will be kept in a locked file and all electronic information will be coded and secured using a password protected file. The information recorded is confidential, and no one else except the researcher will have access to the tapes. The tapes will be destroyed after two months of the study.

Consent

Your signature below indicates that you have decided to volunteer as a research participant for this study, and that you have read and understood the information provided above. You will be given a signed and dated copy of this form to keep, along with any other printed materials deemed necessary by the study investigators.

I agree to participate in the study.

Participant's Name & Signature _____

Date_____

SIGNATURE OF INVESTIGATOR

I have explained the research to the participant and answered all of his/her questions. I believe that he/she understands the information described in this document and freely consents to participate.

Investigator Name & Signature: _____

Investigator's Signature_____

Faculty Advisors: Dr. Attiya Javid & Dr. Fazli Hakim Khattak

Signature_____

Appendix C: Interview Questions

1. Could you please briefly elaborate on the health policymaking process in Pakistan?

2. What is the role of data and evidence in health policy formulation process?

3. Are health policies in KPK backed by data and evidence?

4. Are you aware of the terms such as big data and big data analytics and its uses?

5. Are you aware of any positives/opportunities associated with big data analytics for healthcare?

6. Are you aware of any challenges associated with big data analytics for healthcare?

7. What do you think what are the hurdles to incorporate big data analytics in the healthcare and healthcare policymaking?

8. Do you think big data analytics can empower the health policy formulation process?

9. Do you think our province is ready to embrace data-driven technologies in healthcare?

10. Do you think big data analytics possess any implications for health policymaking?

11. Do you think KPK is improving when it comes to health services provision and health policymaking process?

12. Any further comments you would like to add apart from these questions? (Optional)