

**A STUDY ON PATIENT DATA SHARING  
ACCEPTABILITY AMONG MALAYSIAN WITH  
DATA MINING TECHNIQUES**

**LEE XIN YING**

**UNIVERSITI TUNKU ABDUL RAHMAN**

**A STUDY ON PATIENT DATA SHARING ACCEPTABILITY AMONG  
MALAYSIAN WITH DATA MINING TECHNIQUES**

**LEE XIN YING**

**A project report submitted in partial fulfilment of the  
requirements for the award of Bachelor of Biomedical  
Engineering with Honours**

**Lee Kong Chian Faculty of Engineering and Science  
Universiti Tunku Abdul Rahman**

**May 2023**

**DECLARATION**

I hereby declare that this project report is based on my original work except for citations and quotations which have been duly acknowledged. I also declare that it has not been previously and concurrently submitted for any other degree or award at UTAR or other institutions.

Signature :



Name :

LEE XIN YING

ID No. :

1801629

Date :

28-04-2023

**APPROVAL FOR SUBMISSION**

I certify that this project report entitled “**A Study on Patient Data Sharing Acceptability among Malaysian with Data Mining Techniques**” was prepared by **Lee Xin Ying** has met the required standard for submission in partial fulfilment of the requirements for the award of Bachelor of Biomedical Engineering with Honours at Universiti Tunku Abdul Rahman.

Approved by,

Signature

:



Supervisor

:

Dr. Goh Choon Hian

Date

:

19 May 2023

The copyright of this report belongs to the author under the terms of the copyright Act 1987 as qualified by Intellectual Property Policy of Universiti Tunku Abdul Rahman. Due acknowledgment shall always be made of the use of any material contained in, or derived from, this report.

© 2023, Lee Xin Ying. All right reserved.

## **ACKNOWLEDGEMENTS**

I would like to express my immense gratitude to my supervisor, Dr. Goh Choon Hian, for his invaluable guidance, support and encouragement throughout the project. Additionally, I would like to extend warm thank you to the respondents who took time to participate in my study and everyone else who has supported me in this journey, particularly Ir. Dharmesh Doshi as an industrial expert who cleared my doubts in the initial stage of the project.

## ABSTRACT

The Coronavirus Disease (Covid-19) is a major stress test for the healthcare sector. Malaysia is one of the countries that quickly adopted telehealth in a large-scale. Sharing personal information and health records with healthcare providers are required in telehealth. During interactive video conference, patient data transmission happens at all sites. Growing interest and adoption in telemedicine bring cybersecurity pressure to the healthcare industry. Legal and ethical aspects are imperative in telemedicine. In Malaysia, Telemedicine Act 1997 has been promulgated and passed but is yet to be enforced. Patients must acknowledge the risks of private health information being exposed. Patient acceptability is essential to propel telehealth quality and adoption. This study aims to explore patient data sharing acceptability among Malaysian using data mining techniques. This study may lay a foundation for all authorities, highlighting current ambiguity and implementing effective interventions. Ultimately, state-of-the-art healthcare services could be accessed and delivered. Actual execution of questionnaire distribution yielded 162 respondents. Data analysis and data mining were conducted on survey results with SPSS and Python for closed-ended questions and open-ended questions respectively. Latent Dirichlet Allocation (LDA) topic modeling alongside term frequency-inverse document frequency (TF-IDF) score are used for analyzing unstructured texts collected from open-ended questions which allows true insights collection. More than 61% of respondents responded willing to disclose information on family contacts, phone numbers, email addresses, and geolocation. Additionally, more than 58% of respondents were willing to provide their medical prescription, medical history, and diagnostic results. Age, work or study field, income and marital status were found to be associated with data sharing attitude. Malaysians expressed privacy breach concerns and worried about the digital literacy of the elderly. Information and Technology (IT) industry received the lowest level of trust. Data privacy, transparent communication and education were suggested ways to enhance data sharing willingness.

## TABLE OF CONTENTS

<b>DECLARATION</b>		<b>i</b>
<b>APPROVAL FOR SUBMISSION</b>		<b>ii</b>
<b>ACKNOWLEDGEMENTS</b>		<b>iv</b>
<b>ABSTRACT</b>		<b>v</b>
<b>TABLE OF CONTENTS</b>		<b>vi</b>
<b>LIST OF TABLES</b>		<b>viii</b>
<b>LIST OF FIGURES</b>		<b>ix</b>
<b>LIST OF SYMBOLS / ABBREVIATIONS</b>		<b>x</b>
<b>LIST OF APPENDICES</b>		<b>xi</b>
 <b>CHAPTER</b>		
<b>1</b>	<b>INTRODUCTION</b>	<b>1</b>
	1.1 General Introduction	1
	1.2 Importance of the Study	3
	1.3 Problem Statement	3
	1.4 Aim and Objectives	5
	1.5 Scope and Limitation of the Study	5
	1.6 Contribution of the Study	6
	1.7 Outline of the Report	6
<b>2</b>	<b>LITERATURE REVIEW</b>	<b>7</b>
	2.1 Introduction	7
	2.2 Telemedicine services in Malaysia	7
	2.2.1 Telemedicine guidelines and privacy policy in Malaysia	8
	2.3 Related studies on telehealth implementation and acceptability	10



2.4	Survey platforms	16
2.5	Type of survey questions	17
2.6	Data analysis	17
	2.6.1 Statistical analysis	17
	2.6.2 Data mining	19
2.7	Summary	20
<b>3</b>	<b>METHODOLOGY AND WORKPLAN</b>	<b>21</b>
3.1	Introduction	21
3.2	Project planning	21
3.3	Study method	22
	3.3.1 Sample size	22
	3.3.2 Survey design	23
3.4	Pilot study	25
3.5	Data analysis	25
	3.5.1 Closed-ended responses analysis	25
	3.5.2 Open-ended responses analysis	26
3.6	Summary	28
<b>4</b>	<b>RESULTS AND DISCUSSION</b>	<b>29</b>
4.1	Introduction	29
4.2	Socio-demographic characteristic of study respondents	29
4.3	Closed-ended responses result analysis	31
	4.3.1 General perception towards telehealth	31
	4.3.2 Data sharing acceptability	34
4.4	Open-ended responses result analysis	40
4.5	Summary	58
<b>5</b>	<b>CONCLUSION AND RECOMMENDATIONS</b>	<b>59</b>
5.1	Conclusions	59
5.2	Recommendation for future work	59
	<b>REFERENCES</b>	<b>60</b>
	<b>APPENDICES</b>	<b>67</b>

## LIST OF TABLES

Table 2.1:	Healthcare services included in Malaysia's telemedicine application	7
Table 2.2:	Types of medications and information collected	10
Table 2.3:	Summary of related studies on patients	10
Table 2.4:	Summary of related studies on the public	14
Table 2.5:	Summary of online survey builder comparison	16
Table 2.6:	Types of closed-ended survey questions	17
Table 2.7:	Summary of commonly used statistical tests	18
Table 2.8:	Summary of common data mining techniques	19
Table 3.1:	Type of questions designed with references	24
Table 4.1:	Test of normality and statistics for age	29
Table 4.2:	Sociodemographic characteristics	30
Table 4.3:	Knowledge towards telemedicine	32
Table 4.4:	Perceived reliability and efficiency by region	33
Table 4.5:	Perceived efficiency by marital status	34
Table 4.6:	Financial data sharing willingness by working industry or field of study and monthly income	35
Table 4.7:	Phone number sharing willingness by age	36
Table 4.8:	Diagnostic data sharing willingness by age	37
Table 4.9:	Results on data sharing willingness by age group	39

**LIST OF FIGURES**

Figure 2.1:	General process of data analysis	18
Figure 3.1:	Project flowchart	21
Figure 3.2:	Gantt chart for part one of project	21
Figure 3.3:	Gantt chart for second part of project	22
Figure 3.4:	Phi and Cramer's V values interpretation	26
Figure 3.5:	Text mining process	26
Figure 3.6:	LDA topic modeling process	27
Figure 4.1:	Likert scale bar chart of general perception towards telemedicine	33
Figure 4.2:	Likert scale bar chart of different data sharing willingness	34
Figure 4.3:	Likert scale bar chart for medical data sharing willingness	37
Figure 4.4:	Likert scale chart of willingness to share PHI to organizations	38
Figure 4.5:	Likert scale chart of willingness to allow organizations to keep PHI	39

## LIST OF SYMBOLS / ABBREVIATIONS

ANOVA	Analysis of Variance
COREQ	Consolidated criteria for reporting qualitative research
COVID-19	Coronavirus Disease
CV	Coefficient of Variation
EFA	Exploratory Factor Analysis
EHR	Electronic Medical Records
EU	European Union
GDPR	General Data Protection Regulation
GDsP	Guide to Good Dispensing Practice
HIPAA	Health Insurance Portability and Accountability Act of 1996
KL	Kuala Lumpur
MDA	Medical Device Authority
MMC	Malaysian Medical Council
MoH	Ministry of Health
NPRA	National Pharmaceutical Regulatory Authority
PHI	Personal Health Information
R&D	Research and Development
SAS	Statistical Analysis Software
SEA	South East Asia
SPSS	Statistical Package for Social Sciences
THIS	Total Healthcare Information System
LDA	Latent Dirichlet Allocation
TF-IDF	Term Frequency-Inverse Document Frequency
NLP	Natural Language Processing
UK	United Kingdom
IT	Information and Technology
Q	Question

## **LIST OF APPENDICES**

Appendix A: UTAR's Ethical Approval for Research Project	67
--	----

## CHAPTER 1

### INTRODUCTION

#### 1.1 General Introduction

The Coronavirus Disease (Covid-19) is a major stress test for the healthcare sector to ignite healthcare transformation and solutions. Healthcare players are forced to commit high-quality clinical care with minimal contact to break the virus transmission chain. Telehealth is blooming and being utilized more than ever before in the pandemic world. Telemedicine and telerehabilitation are under the broad umbrella of telehealth. They allow patient care to be delivered remotely without the need for physical contact with the advancement of technology. According to Bestsenyy et al. (2021), digital health adoption had soared 38 folds from the pre-Covid-19 baseline, peaked in the first quarter of 2020 and stabilized since then. Malaysia is one of the South East Asia (SEA) countries that adopted telehealth quickly on a large-scale, chiefly in the private healthcare sector to sustain their revenue during the desperate time (Batumalai, 2022). In Malaysia, telemedicine targets different modalities such as store-and-forward approach for medical specialties like radiology and pathology based on documented medical history or images, remote monitoring for disease management and real-time consultation for diagnosis and medication prescription (Ministry of Health Malaysia, 1997).

Sharing personal information and health records to healthcare providers are required by patient to be benefited from telehealth. Personal information might also be gathered for upcoming studies, user experience enhancements, service upgrades, or even cross-organizational commercial partnerships. During interactive videoconference, patient data transmission happens at all sites. These data will be stored online through electronic medical records (EHR) system or Total Hospital Information System (THIS) of providers to support clinical decisions (Salleh, Abdullah and Zakaria, 2021). Nevertheless, growing interest and adoption in telemedicine bring cybersecurity pressure to the healthcare industry. Increasing in sharing of data from individuals to organizations means creating privacy issues and potential risks. Three main security services are being applied to telemedicine; availability in making sure data is readily

available even during disasters, confidentiality in prevention of data disclosure and integrity in the use of data by authorized personnel (Ali et al., 2013). Wireless patient data transfer makes these services susceptible and exposes the general public to dangers from cyber criminals. According to (Combs, 2021), both doctors and patients had data and privacy security concerns. It is no surprise that the public is concerned as health information breaches are prevalent today. Based on Healthcare Report 2021 of Kaspersky, 32% of respondents reported that their organizations have faced data leaks and ransomware attacks due to vulnerability in technologies (Kaspersky, 2021). In Malaysia, there was an event in which data records of 19992 patients were leaked online with personal information and medical scan images (Habibu, 2019).

Legal and ethical aspects are imperative in telemedicine. The law governs what someone may and cannot do, whereas ethics establish moral standards for what they should and shouldn't do. US Department of Health and Human Services (HHS) issued The Health Insurance Portability and Accountability Act of 1996 (HIPAA) as a benchmark in the protection of patient personal health information, covering entities from healthcare providers to business associates (Centers for Disease Control and Prevention, 2022). Besides, SEA countries such as Malaysia, Indonesia and Singapore have issued policy statements, advisories, laws and blueprints though the guidelines created were mainly to regulate healthcare providers rather than technologies, services and platforms (Intan Sabrina and Defi, 2020). Despite not having HIPAA, legal framework available to date in Malaysia does guide healthcare professionals to conduct telemedicine (Abdul Aziz, 2020).

In addition to infrastructure barriers, technical skills and technological challenges, it is equally important for organizations to address management issues including patient acceptance on the service adoption (Rouidi, Elouadi and Hamdoune, 2022). The current study investigates the acceptance of patient data sharing among Malaysian, willingness to share and adopt telemedicine and telerehabilitation as well as perceptions on these implementations. Statistical analysis and topic modeling were used as a tool to discover patterns, examine attitudes and recognize trends based on datasets obtained. Of note, the word the “telehealth” alludes to a broader scope of remote healthcare services, including telemedicine and telerehabilitation services.

## **1.2 Importance of the Study**

The study presents a significant impact on telehealth, telemedicine and telerehabilitation utilization in Malaysia. Quick adoption of telehealth has been noticed during the hit of unprecedented COVID-19 pandemic but patient acceptability and attitudes towards data sharing for the services are key considerations that are yet to be studied. This study may lay a foundation for all authorities and parties involved from the government, regulatory affairs to healthcare providers in understanding the current perspective of end-users regarding the topic, highlighting the ambiguity and implementing effective interventions. For instance, from the viewpoint of policy makers or healthcare institutions, this study can shed light on patients' privacy concerns thereafter revise the policy terms effectively. Furthermore, researchers can obtain knowledge regarding the viability of conducting studies that rely on patient data, leading to the development of more effective therapies. Besides, better healthcare results may result from this as it fosters patient-provider interaction and builds trust, forging closer bonds with their clients. Ultimately, state-of-the-art healthcare services could be accessed and delivered at instant without the need for physical and regular appointments in healthcare facilities.

## **1.3 Problem Statement**

There are papers being published about aspects to be addressed for telehealth to succeed such as tackling security risks and providing reimbursement (Hall and McGraw, 2014). Better acceptance of utilization of telemedicine technology is reckoned to uplift the quality of healthcare in overcoming geographical barriers (ECRI, 2021). Malaysia is expected to be an aged nation in 2044 with 14 percent of citizens aged 65 years old and over (David, 2022). The older population has privacy concerns about telerehabilitation and telemedicine services. With higher susceptibility to cognitive aging and lower technological literacy, they are at an increased risk (Jafni et al., 2019; Fadzil et al., 2022; Friedman et al., 2022).

To date, there are studies on acceptance of telemedicine without sufficient focus on data privacy considerations. Malaysian Medical Council (MMC) and Ministry of Health, Malaysia (MOH) had developed framework and guidelines pertaining different domains for telemedicine. These include clinical, ethical, legal and operational aspects (Intan Sabrina and Defi, 2020).



These guidelines are created in telemedicine to ensure healthcare providers obtain, store and process patient data lawfully. Telemedicine Act 1997 has been promulgated and passed by Malaysian Parliament yet to be enforced (Baker McKenzie, 2022). Since legislation available are not comprehensive enough to enforce telehealth technology, existence of privacy policies and guidelines cannot guarantee the confidentiality of personal information and hence this could be a stepping stone for telemedicine deployment (Bassan, 2020).

According to a study done by Thong et al. (2021) with respondents of 146 Malaysian doctors from KPJ private hospitals, it was reported that only 22% utilized telemedicine for consultations amid the Covid-19 pandemic. Plus, 43.80% of the healthcare practitioners were unaware of the existence of Telemedicine Act 1997 as a guideline. These unfamiliarities and ignorance of physicians on the advisories available affect the implementation of telemedicine in this society, considering they are the main telemedicine service providers to patients. There are papers published regarding knowledge and opinions of practitioners on telemedicine but end-users perspectives are scarce in Malaysia. For that reason, it is essential to explore end-users perceptions of telemedicine services (Woo and Dowding, 2018).

As telehealth requires constant data collection from patients and remains significant barrier to its adoption, it is essential to comprehend patient data sharing willingness and the factors involved. Data exploitation happens when a party benefits unfairly from the other and causes a series of consequences including harm or discrimination (Cassel and Bindman, 2019). In fact, patients could be the victims of long-term identity theft or leverage details for extortion (Steger, 2019). Therefore, patients ought to acknowledge the risks of private health information (PHI) being exposed, whether there will be fair distribution of benefits since data shared might not be solely for research and development (R&D). Nevertheless, despite having privacy policy consisting of terms and conditions to data sharing, patients are hardly given opportunities to question the exact purpose of sharing personal data even for future research (McCoy, Joffe and Emanuel, 2020). The ultimate purpose of approving the sharing of PHI with healthcare organization is to seek healthcare advice and services, which results in patients' lack of control.

#### **1.4 Aim and Objectives**

This study aims to examine patient data sharing acceptability among Malaysian and analyze it using data mining techniques.

The specific objectives of the study are:

- (i.) To develop a survey questionnaire on patient perception and willingness to share personal data and healthcare information for telemedicine services use in the Malaysian context.
- (ii.) To perform data analysis on closed-ended survey results using statistical and inferential analysis with SPSS
- (iii.) To perform data analysis on open-ended survey results using topic modeling with Python.

#### **1.5 Scope and Limitation of the Study**

The study focuses on patients who reside in Malaysia, as a citizen of Malaysia. The data collection for pilot study was conducted based on convenience sampling. Each participant was given the same questionnaire to answer. The prime limitation of this study is the smaller sample size for older age groups (55 and above) and divorced or separated groups, it has limited representation from older aged respondents and this could impact the findings. Besides, there is lack of representation from certain states, primarily from East Malaysia. Another limitation of this study is the lack of participants from other races other than Chinese. Hence, this study may not fully reflect and report on the perspectives of the Malaysian population.

Besides, lack of funding is another existing limitation as no honorarium will be granted to participants and hence voluntary response rate will be lowered and this might affect the quality of responses received. Since this study involves self-administered questionnaire, the response rate might be low or people may be reluctant in responding. At the same time, there is always a risk that participants were not honest when answering the questions, different interpretations or incorrect feedback could be received.

## **1.6 Contribution of the Study**

This study identifies patient data willingness, factors and their behaviors towards telehealth and healthcare in general. It can facilitate healthcare providers in understanding, advancing patient-centered healthcare and improving patient outcomes in the future. Besides, the study fills the gaps in previously published papers, specifically in the context of the Malaysian population.

## **1.7 Outline of the Report**

This report is being divided into four main parts. The second section (Chapter 2) provides a review of pertinent literature on the current study, analysis, and results, along with a summary of the main points. Part three (Chapter 3) of the report outlines the methodology used in the current study holistically, from questionnaire development, pilot study analysis, sampling method, and responses collection to analysis for open and closed-ended questions. Part four (Chapter 4) presents the findings of analysis of this study in detail. The study's future course is concluded in the last part (Chapter 5).

## CHAPTER 2

### LITERATURE REVIEW

#### 2.1 Introduction

This chapter presents an in-depth review on two major parts, including telemedicine services background in the Malaysian context and survey form framework. Sources are mainly selected with publication years after 2014. Types of telehealth and telemedicine services offered in Malaysia along with the privacy policy, guidelines and statutes issued will be discussed. The topic also covered previous works that have been published in various nations.

#### 2.2 Telemedicine services in Malaysia

In Malaysia, telemedicine refers to provision of healthcare services that links consumers, public institutions, healthcare providers, agencies and suppliers through multimedia networks in achieving health service goals (Ministry of Health Malaysia, 1997). Table 2.1 shows the healthcare function and services included in this nation.

Table 2.1: Healthcare services included in Malaysia's telemedicine application

Services	Details
Information and education	Education and training services on healthcare issues
Consultation, diagnostic and treatment	Healthcare clinical services including rehabilitation, consultation, diagnostic and treatments provided by or between healthcare professionals
Support	Pharmaceutical, medical supply, insurance, electronic government services and payment systems
Governance	Policies, law, accreditation, standards, guidelines and related services in assisting high-quality healthcare delivery

There are multitude of forms of telemedicine platforms for clinical and support services available in Malaysia, these include clinics, hospitals, pharmacies and companies. Ng et al. (2022) had conducted a study on the availability and scale of teleconsultation use among Malaysian public clinics. It was reported that less than half of total responded clinics provided teleconsultation and majority of the consultation services were being conducted through phone calls rather than video calls. In addition, the consultation services were primarily offered to individuals with diabetes and hypertension rather than for follow-up needs.

From the perspective of private healthcare sectors, hospitals such as Pantai Hospitals, Sunway Medical Centre, Columbia Asia Hospitals and Ara Damansara Medical Centre provide telehealth services and virtual care. Virtual consultation with practitioners (e.g., doctors, psychologists, therapists and rehabilitation team), in-home phlebotomy, prescription and delivery of medication can be accomplished without having to be present physically. These services are appropriate for individuals, companies and nursing homes who require second opinion or follow-up on their medical condition concerning the medical specialties. Apart from private hospitals, companies such as Doc2Us, Speedoc and Naluri offer a range of healthcare services and support to cater the public's quality healthcare needs. For instance, aside from teleconsultation, Naluri provides interactive, personalized information and education support for overall physical and mental health with information storage function for progress-tracking and monitoring (Medtech Innovator, 2021).

### **2.2.1 Telemedicine guidelines and privacy policy in Malaysia**

Privacy policy refers to statement that details the way of an organization in handling one's personal information (Office of the Australian Information Commissioner, 2022). These important documents are developed by healthcare providers and should be viewed by consumers prior supplying consensus in giving and receiving services. Generally, there are existing statutes and guidelines that are applied in most forms of trades, these include Personal Data Protection Act 2010 that governs processing of personal data in commercial transactions and Guide to Good Dispensing Practice (GDsP) which was developed based on Poisons Act 1952 in ensuring quality medicine

dispensing(Personal Data Protection Act 2010; Pharmaceutical Services Division Ministry of Health Malaysia, 2016).

The key healthcare regulatory agencies in Malaysia are Ministry of Health, Malaysia (MOH), Medical Device Authority (MDA) and Malaysian National Pharmaceutical Regulatory Authority (NPRA). In the context of telemedicine use, the Malaysian government including Malaysian Medical Council (MMC), MOH have issued different legislations, guidelines, advisories and blueprints to guide and regulate the practices. The MMC published advisory on virtual consultation during the unprecedented COVID-19 pandemic to guide effective, safe and cost-effective primacy virtual care with online platforms and mobile applications. Other initiatives to promote telehealth employment include Malaysia's Telemedicine Plan and Garis Panduan Pelaksanaan Klinik Virtual di Hospital, 2020. The Telemedicine Act of 1997 was passed, but since it has not yet taken effect, it has no legal effect (Raja Eileen Soraya et al., 2021).

Private hospitals and entities such as Sunway Medical Centre and Gleneagles Hospitals have established privacy policy on the use of telemedicine and teleconsultation services. The policies spell out the terms and conditions pertaining multiple aspects and liabilities. These hospitals promised to keep all medical reports and calls private and confidential under patient-doctor confidentiality. Sunway Medical Centre protects patients' medical information private by complying with HIPAA standards alongside General Data Protection Regulation (GDPR) of Malaysia and those who reside in European Union (EU).

In terms of e-pharmacies, online telemedicine services platforms such as DoctorOnCall and Doc2Us abide to legislations and guidelines in their dispensation policies. For instance, Group B (prescription only) medicine requires valid prescription before purchase. This means Group B medicines are strictly dispensed upon prescription by doctor through online consultation with electronic signature. Table 2.2 shows the information collected for each category of medicines. All dispensed medication is done according to the regulations spelled out in Laws of Malaysia Poison Act 1952. However, drugs can be delivered without appointment of a licensed delivery company (Baker McKenzie, 2022).

Table 2.2: Types of medications and information collected

Medications	Information
Group B (Prescription only)	<ul style="list-style-type: none"> <li>- Signature, address, prescription date of prescriber</li> <li>- Name, ID number and address of patient</li> <li>- Dosage, instruction of taking medication and quantity</li> </ul>
Group C	<ul style="list-style-type: none"> <li>- Name, ID number and address of patient</li> <li>- Drug dosage, name, quantity and date of prescription</li> </ul>
Over-the-counter	<ul style="list-style-type: none"> <li>- Sold over the counter without a prescription</li> </ul>

### 2.3 Related studies on telehealth implementation and acceptability

There are different studies and papers published regarding perception towards telehealth and its acceptability around the world. However, resources were scarce in the Malaysian context. Table 2.3 shows the summary of related work on patients while Table 2.4 shows the summary of related work on the public including medical practitioners and elderlies. According to the tables, it can be deduced that telemedicine possessed high acceptance but data security is primary concern.

Table 2.3: Summary of related studies on patients

Author/ Year	Study / Resource Type	Study Methods/ Type	Key Findings/ Conclusion
(O'Brien et al., 2019)	Examine the United States patients in perspective on the pros and cons of sharing data to healthcare practitioner	Questionnaire - Frequency calculation for categorical variables, standard deviation for continuous variables - Analysis done by Statistical Analysis Software (SAS)	More than half of agreed to share PHI without disclosure of name and social security number (SSN). Better communication and clear intentions increase patient comfort level.

Table 2.3 (continued)

(Karampe la, Ouhbi and Isomursu, 2019)	Explore European patients' willingness and conditions on sharing PHI	Online Questionnaire -Analysis done by Statistical Package for Social Sciences (SPSS) Statistics with data visualizations	Majority users were more willing to share PHI, 22.63% were willing to share PHI for scientific research. Trust establishment is crucial to develop positive attitudes. Age, education and jobs were important factors.
(Tan et al., 2022)	Explore Kuala Lumpur, Malaysian older patients and caregivers on acceptance of virtual consultation	Questionnaire - Utilized Unified Theory of Acceptance and Use of Technology Scale model	High level of acceptance for virtual consultations as a vulnerable group.
(Jansen-Kosterink, Dekker-van Weering and van Velsen, 2019)	Explore understanding and acceptance of Netherlands 14outpatient rehabilitation program patients on telerehabilitation.	Study/ Interview -Technology Acceptance Model and Unified Theory of Acceptance and Use of Technology Scale model -Involves portal for rehabilitation care - Reported by means of Consolidated criteria for reporting qualitative research (COREQ) -Analysed with three coders, thematic analysis	Some participants claimed the need to have therapist as motivation is vital. Participants felt telerehabilitation provide sense of autonomy.



Table 2.3 (continued)

(Aggarwal et al., 2021)	Explore United Kingdom patients and public on data sharing awareness on Artificial Intelligence for healthcare and research use.	Questionnaire -Responds manually entered into Excel -Analysis included descriptive analysis, logistic regression - Analysis with SPSS.	Patients have low knowledge level about the field but understands how PHI are used for research than the commercial organizations and researchers in university. Anonymization of data was strongly supported (83.8%). Trust is important factor when sharing sensitive information, patients were not comfortable sharing data to for-profit entities due to privacy fears.
(Seltzer et al., 2019)	Explore willingness of United States academic Emergency Department patients to share digital health and non-health data for research.	Questionnaire -Descriptive statistics to characterize components - Analysis done using Exploratory factor analysis (EFA) in identifying clusters of data sources based on groups	65% patients willing to share at least 1 data type, prefer to share all types after death. Willing to share financial and location data the least. Majority have concerns about privacy breaches, in ways not stated in policies.

Table 2.3 (continued)

(Kato-Lin and Thelen, 2022)	Investigate how virtual patients with acute conditions in United States weigh illness and pandemic against privacy concerns for telemedicine utilization.	Questionnaire -Analysis done by confirmatory factor analysis and construct reliability, bivariate analysis (experience/characteristics) linear regression models in finding relationship between respondent's characteristics and privacy/intention outcomes, Structural equation modeling.	Patients did not think privacy concerns will affect continued use intention during environmental risk factors. Satisfactory and privacy concerns are vital in telemedicine adoption. New telemedicine patients were more concerned in privacy issues than repeat users.
(Haque et al., 2022)	Study perception towards telemedicine of Bangladesh residents with chronic diseases.	Study/ Interview -Analysis done with SPSS, logistic regression.	Most patients were willing to learn but those with higher education were less willing to receive telemedicine services due to privacy concerns.
(Fadzil et al., 2022)	Scoping review on telerehabilitation implemented among older adults with cognitive frailty and mild cognitive impairment.	Scoping review	Older adults have problems exploring technology and require assistance with device. Data privacy, cost and familiarity are concerns.

Table 2.4: Summary of related studies on the public

<b>Author/Year</b>	<b>Study/ Resource Type</b>	<b>Study Methods/ Type</b>	<b>Key Findings/ Conclusion</b>
(Thong et al., 2021)	Investigate perception and barriers of Malaysian healthcare practitioners towards implementation of telemedicine.	Online Questionnaire -Analysis done by Statistical Package for Social Sciences (SPSS) Statistics	Only 22% reported using telemedicine. 67.1% agreed that telemedicine is ideal for follow-up basis. Barriers mainly target legal consent, billing charges and reimbursements.
(Bell, Ohno-Machado and Grando, 2014)	Explore United States healthy individuals on data sharing preferences	Questionnaire - Built Graphical User Interface for study	Respondents prefer to have data sharing control. Large proportion have concern on use of data by for-profit organizations. Transparency and trust are biggest factor to sharing willingness.
(Kim and Choi, 2019)	Explore factors that influence elderly's PHI sharing willingness in South Korea	Questionnaire -Logistic regression	Most old adults were less willing to share PHI to government agencies, devices corporation and insurance companies. Education and age were factors influencing willingness to share PHI to entities.
(Lee, Wong and Lee, 2020)	Study mobile-Health perception and views of people living in Selangor, Malaysia	Questionnaire - Analysis with multivariable logistic regression models	Most respondents were not familiar but had positive attitude. Education and training are important. High acceptance in sharing PHI with professionals and family members. Baby boomers and younger generations were concerned about data privacy but younger generation were inclined to mobile-Health adoption.

Table 2.4 (continued)

(Grande et al., 2022)	Study on United States adults on willingness to share personal digital information for health-related uses	Study -Conjoint analysis -Latent class analysis in R - Stata version 16 (StataCorp)	Participants were less willing to share financial information. Many have strong underlying opinions on privacy issues. New protections created could increase confidence on sharing personal data.
(Lin, 2017)	Study comparing overall telemedicine service perception and use intention among Chinese frequent travellers and non-frequent travellers.	Study -Quantitative analysis -Descriptive analysis -Factor analysis (T-test and ANOVA) -Multigroup structural equation modeling	Reliability, ease of use and needs of interaction resulted in higher perceptions. Financial cost and usefulness play lesser role. Frequent travellers have stronger intention to use telemedicine.
(Lam et al., 2020)	Study United States older adults on their unreadiness for telemedicine during Covid-19 pandemic	Study -Multivariable logistic regression	Unreadiness increases with age, lower income and lower education. Policies should be recognized. Older adults are left behind due to inexperience with technology.
(Jafni et al., 2019)	Study on Malaysian clinical professionals' barriers to telerehabilitation.	Study/ Interview -Thematic analysis	Barriers for telerehabilitation includes demotivated, no urgency to change, less awareness, less software and low connectivity as well as lack if exposure on e-Healthcare knowledge and resistance to change.

Table 2.4 (continued)

(Friedman et al., 2022)	Perspective on ethical considerations and recommendation in addressing online health privacy risks for older adults.	Article	Online health data seeking erodes personal privacy. Older adults should understand privacy risks to protect personal information. People should ensure that existing and emerging legislative proposals are informed by the needs of older population.
-------------------------	--	---------	--

## 2.4 Survey platforms

Online survey builders come in a variety of forms. To be a fantastic platform for survey form distribution, the builder needs to be affordable, able to collect a large number of responses, simple to use, and equipped with powerful and eye-catching design capabilities (Matthew, 2021). Table 2.5 presents a summary of comparison between SurveyMonkey, Google Forms and GorillaSurvey. SurveyMonkey is also a prominent platform as it can automatically build a survey form according to pre-written questions. Gorilla is an experiment builder that assist scientists in online experiments development, it enables behavioural study with design tools like randomization and delays (Anwyl-Irvine et al., 2020). When compared to other programmes, Matthew (2021) determined that Google Forms is the fastest and most cost-effective. To make the process easier, questions can also be reused and copied.

Table 2.5: Summary of Online Survey Builder Comparison

Features	Gorilla Survey	Google Forms	SurveyMonkey
Ease of Use	Intermediate	Beginner	Intermediate
Price	Paid after free trial	Free	Freemium
Response Limit	Unlimited	Unlimited	Free for certain numbers
Customization	High	Medium	High
Mobile Support	Yes	Yes	Yes

## 2.5 Type of survey questions

To create a successful survey questionnaire, it is crucial to decide on the type of questions intended to be used. The two main categories of survey questions are closed-ended and open-ended format, selecting the correct type directly influence the collection of feedbacks for analysis. With a robust collection of questions, insightful data and quality engagement can be achieved to fit the purpose of project. Open-ended questions allow respondents to answer based on elaborations rather than static options while closed-ended questions provide predefined selections. Closed-ended questions come in myriad of forms, including multiple choices, Likert scale, rating scale choices and checklist type choice questions (QuestionPro, 2022). Table 2.6 shows the types of closed-ended survey questions.

Table 2.6: Types of closed-ended survey questions

Question	Details
Multiple choice/Dropdown	Allow single answer from multiple options
Likert scale	Rank range of items from ordered set
Checklist type	Choose choices from list of options
Rating scale	Evaluate weight of answer choice

## 2.6 Data analysis

Data analysis is vital in research as it allows raw information to be distilled into an accurate and relevant form utilizing various methods such as descriptive analysis, statistical analysis and data mining (Karin, 2023). There are different tools that could be used to perform data analysis, which include SPSS, Tableau, RapidMiner, Excel, R Programming and Python.

### 2.6.1 Statistical analysis

SPSS is an excellent software tool to analyze scientific data easily, it offers plethora of statistical tests (Bevans, 2020). For instance, Chi-square test, Fischer's exact test, T-tests, correlation, regression and ANOVA. Figure 2.1 presents the general process of data analysis.

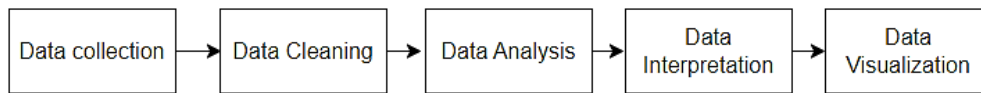


Figure 2.1: General process of data analysis

There are two categories of data in data analysis which are numerical and categorical (Donges, 2023). Numerical data is represented by numbers and is further classified into two types, discrete and continuous. Discrete numerical data are values that can only be represented by integers. In contrast, continuous numerical data can have decimal values. On the other hand, categorical data is expressed through categories, which are also referred to as components or groupings. Categorical data is further subdivided into two types, nominal and ordinal. There is no logical order between categories in nominal categorical data. Ordinal categorical data have categories that may be arranged in a meaningful order. Table 2.7 shows a summary on types of commonly used statistical tests.

Table 2.7: Summary on commonly used statistical tests

<b>Statistical Test</b>	<b>Purpose</b>	<b>Data Type</b>
Chi-Square	association between two categorical variables	Nominal or Ordinal
Fischer Exact	association between two categorical variables	Nominal or Ordinal
t-test	difference between two group means	Continuous
ANOVA	difference between means of three or more groups	Continuous
Mann-Whitney	difference between two independent group medians	Ordinal or Continuous
Kruskal-Wallis	difference between three or more independent group medians	Ordinal or Continuous
Linear Regression	linear relationship between two continuous variables	Continuous

### 2.6.2 Data mining

Data mining is the process of searching for patterns and extracting valuable data from large amounts of data. It is widely used in the "big data" environment because it enables data analysis and profit through predictive and descriptive data mining. There are several commonly used data mining techniques (Taylor, 2023). Table 2.8 outlines the techniques and use.

Table 2.8: Summary of Common Data Mining Techniques

<b>Techniques</b>	<b>Description</b>
Clustering	Group data into clusters according to similarities and attributes. Use to discover collections of data objects.
Classification	Categorize data into predefined classes. Data analyzed based on algorithms.
Regression	Identify and analyse connections between variables. Used for prediction and forecasting numeric values.
Text mining	Analyze and extract unstructured text data to get insights by categorizing main topics. Use Natural Language Processing (NLP) or linguistic principles to evaluate documents.
Association rules	Search for association between items while discover hidden patterns.

In text mining, topic modeling is a popular approach. Topic modeling is an unsupervised machine learning method similar to clustering to uncover hidden semantic structure of texts. The key approach to address such technique was by utilizing the Natural Language Processing (NLP) Latent Dirichlet Allocation (LDA) algorithm. LDA is a mathematical generative probabilistic model that treats each document text as a mixture of topics with a mixture of words (Li, 2018). LDA determines the likelihood each topic is in each document. Besides, it presumes that topics are constant and unchanged over time. In LDA, each document consists of a mix of topics and each topic consists of a mix of words.



## **2.7 Summary**

Healthcare organisations have adopted telehealth services, including e-pharmacy, but different regulations, rules, and criteria must be followed in order to deliver these services legally and ethically. Telemedicine Act of 1977 was passed but has yet to be enforced. While telemedicine services are widely accepted, privacy violations remain a serious worry. Trust, education, and improved policies all seemed to be major contributors in increasing public satisfaction and acceptance. Google Forms has deemed to be the best survey generator. Data analysis extracts relevant insights and information. Statistical analysis is the process of analysing data and drawing conclusions using statistical methods while data mining is the process of identifying patterns and insights.

## CHAPTER 3

### METHODOLOGY AND WORK PLAN

#### 3.1 Introduction

This chapter will be divided into subsections, detailing the methods that are being conducted in development of questionnaire and pilot study. Figure 3.1 presents the project flowchart to illustrate workplan.

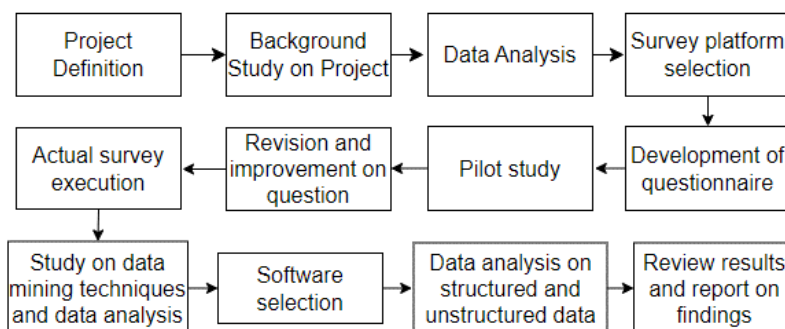


Figure 3.1: Project flowchart

#### 3.2 Project planning

Gantt charts as shown in Figure 3.2 and Figure 3.3 to illustrate the workplan timeline for the first and second part of the project respectively. The charts act as guide to ensure progress is on track and achieve outcomes.

Phase No.	Task Description	W1	W2	W3	W4	W5	W6	W7	W8	W9	W10	W11	W12	W13	W14
		13/6 19/6	20/6 26/6	27/6 3/7	4/7 10/7	11/7 17/7	18/7 24/7	25/7 31/7	1/8 7/8	8/8 14/8	15/8 21/8	22/8 28/8	29/8 4/9	5/9 11/9	12/9 18/9
Phase 1- Project Planning	Brief background study on scope of project														
	Project timeline planning														
Phase 2- Literature Review	Study on papers and evaluation														
	Problem identification														
Phase 3- Project Development	Study on survey patterns and platform selection														
	Develop survey questionnaire														
	Pilot study and study validation														
	Preliminary evaluation and review data														
	Questionnaire modification														
Phase 4- Report Findings and Presentation	Actual execution and survey form distribution														
	Logbook updates														
	Progress report writing														
	Project presentation														

Figure 3.2: Gantt chart for part one of project

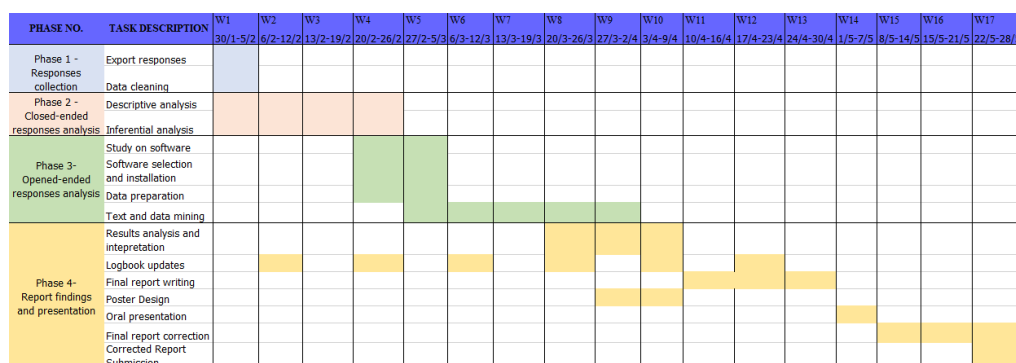


Figure 3.3: Gantt chart for second part of project

### 3.3 Study method

Study was conducted among Malaysians who are aged 18 and above. Survey questionnaire was selected as study design as the purpose of the study is to discover and present the perception and chiefly the acceptability of patient data sharing in the Malaysian context. Convenience sampling method was used for the pilot study as this is the easiest and most suitable method for preliminary phase of a study (Cint, 2021). The official study was done in voluntary basis without honorarium. Participation information sheet have been created to allow potential respondents to understand the study, their rights as a participant and obtain informed consent. All results collected were pooled so that no responses can be attributed to individual persons. Ethical clearance for this study was obtained as shown in Appendix A and the study was conducted in accordance with ethical principles outlined.

#### 3.3.1 Sample size

To determine the sample size needed, Slovin's Formula was used. Calculation of sample size necessary to obtain targeted confidence interval in population sampling can be done with the following Equation 3.1 as below:

$$n = \left( \frac{N}{1+Ne^2} \right) \quad (3.1)$$

Where,

n = sample size

N = population size

e = margin of error

The population size was set at 3000 with a confidence level of 95% and a margin of error of 5% since participants would be solicited using a non-probability sampling method known snowball sampling. 3000 population size is assumed so that the survey can be feasible within available time frame and resources. The calculated sample size is 138 persons based on a 90% response distribution.

### **3.3.2 Survey design**

The survey questionnaire is constructed with Google Forms as builder. Red was selected as theme color of the study as it is perceived as energizing. Questionnaire developed for pilot study encompassed 3 sections. Upon completion of pilot study, finalized questionnaire consist of Section A (10 demographic questions and health status), Section B (9 general perceptions towards telehealth), Section C (3 data sharing acceptability questions) and Section D (9 opinions, experience and thoughts). Both open and closed-ended questions were used to obtain quantitative and qualitative overview on the current landscape. For closed-ended questions, Likert scale, multiple choice and rating scale question techniques were applied according to predicted feedback.

Demographic questions including socioeconomic information were included to understand respondents better or assist in analysis and describe findings. In the following section, questions were designed to obtain general perception of respondents towards telehealth services in multiple aspects. This is to study respondents' perceived value towards telehealth and their awareness onto related privacy policies. To determine patients' data sharing willingness, a series of questions pertaining personal data, medical information and confidence level were included in the form of rating scale in sections. The last section included open-ended questions designed to elicit comments about telemedicine, telerehabilitation, and e-pharmacy in terms of data protection, concerns, and ideas with given descriptions and situations. Table 3.1 displays all finalized designed questions with their respective references. The questions were designed with minimum word limits to maximize the potential possibility of receiving quality responses.

Table 3.1: Type of questions designed with references

Question	Type of question	References
<b>Section A - Demographic data and general health status</b>		
1	Ratio scale	
2	Multiple choice	
3	Multiple choice	
4	Multiple choice	
5	Multiple choice	
6	Multiple choice	
7	Multiple choice	
8	Multiple choice	(Indeed, 2021)
9	Ratio scale	(CompareHero, 2022)
10	Likert Scale	
<b>Section B - General Perception towards Telehealth (Telemedicine and Telerehabilitation)</b>		
1	Multiple choice	(Ministry of Health Malaysia, 1997)
2	Multiple choice	
3	Likert Scale	(Lin, 2017)
4	Multiple choice	
5	Multiple choice	
6	Open-ended	
7	Multiple choice	(Lin, 2017)
8	Multiple choice	
9	Open-ended	(Lee, Wong and Lee, 2020; Department of Statistics Malaysia, 2021)
<b>Section C- Data sharing acceptability</b>		
1	Rating scale	(Grande et al., 2022)
2	Rating scale	(Grande et al., 2022)
3	Rating scale	(Grande et al., 2022)
<b>Section D- Opinions, Experience and Thoughts</b>		
1	Open-ended	(Ministry of Health Malaysia, 1997; Baker McKenzie, 2022; Chin, 2022)
2	Open-ended	(Baker McKenzie, 2022)
3	Open-ended	
4	Open-ended	
5	Open-ended	
6	Open-ended	
7	Open-ended	(Majeed, 2019; O'Brien et al., 2019; Aggarwal et al., 2021)
8	Open-ended	
9	Open-ended	(Ozeran, Solomonides and Schreiber, 2021)

### **3.4 Pilot study**

Pilot study was conducted to pre-test the survey questionnaire before using it to collect data. Targeted respondents' number was 10% of the total 138 respondents needed. Survey form was sent out to respondents via convenience method for easier feedback collection while completion time was predicted to be within an hour. In addition, response time was gathered to determine the average expected time required for study completion. This pre-test was important and helpful to identify and eliminate confusion and potential issues in constructed questions. Amendments were made to finalize the form for actual study execution.

### **3.5 Data analysis**

Data analysis was a significant component during the study endeavor. This part details the data analysis procedure to acquire thorough knowledge from the data and make meaningful conclusions from the study results.

#### **3.5.1 Closed-ended responses analysis**

Descriptive and statistical analysis were performed using the SPSS program version 26. Descriptive analysis was done for calculating percentages and counts for categorical data to assess patterns or directions of variables such as baseline data. Meanwhile, statistical analysis was carried out with bivariate analysis, Fisher's exact test and Chi-square test. The purpose of analysis is to discover potential association and assess independence between each pair of categorical variables. By reviewing prior studies and finding relevant linked factors of interest, such as sociodemographic characteristics with perceived attitudes and perceptions, the variables for statistical and descriptive tests were meticulously chosen.

When testing larger sample sizes, the chi-square test is frequently more effective than Fisher's exact test. It presumes that there are at least 5 (20%) predicted frequencies in each cell of the contingency table. Nevertheless, if the anticipated frequencies are less than 5, the Fisher test is applied, thereby being more suitable for lower sample sizes (Kim, 2017). Prior to conducting the statistical tests, null hypothesis ( $H_0$ ) was formulated by assuming there is no

significant association between the variables, the factors are not reliant on each other.

Confidence interval of 95% was used and with every statistically significant result ( $p < 0.05$ ) obtained, null hypothesis was rejected. Measure of association with Phi or Cramer's V was then performed to association strength between the nominal groups. In contrast to Cramer's V, Phi is appropriate for a 2x2 contingency table or two dichotomous variables. Figure 3.4 presents the Phi and Cramer's V values interpretation (Akoglu, 2018).

Phi and Cramer's V	Interpretation
> 0.25	Very strong
> 0.15	Strong
> 0.10	Moderate
> 0.05	Weak
> 0	No or very weak

Figure 3.4: Phi and Cramer's V values interpretation

### 3.5.2 Open-ended responses analysis

Python was utilized for conducting open-ended survey responses analysis with the Scikit-learn library as it processes much faster than Gensim especially when the number of topics is relatively low (Greer, 2018). Topic modeling was applied with LDA to comprehend all 162 responses to each question at once. Figure 3.5 presents a flowchart of text mining process applied.

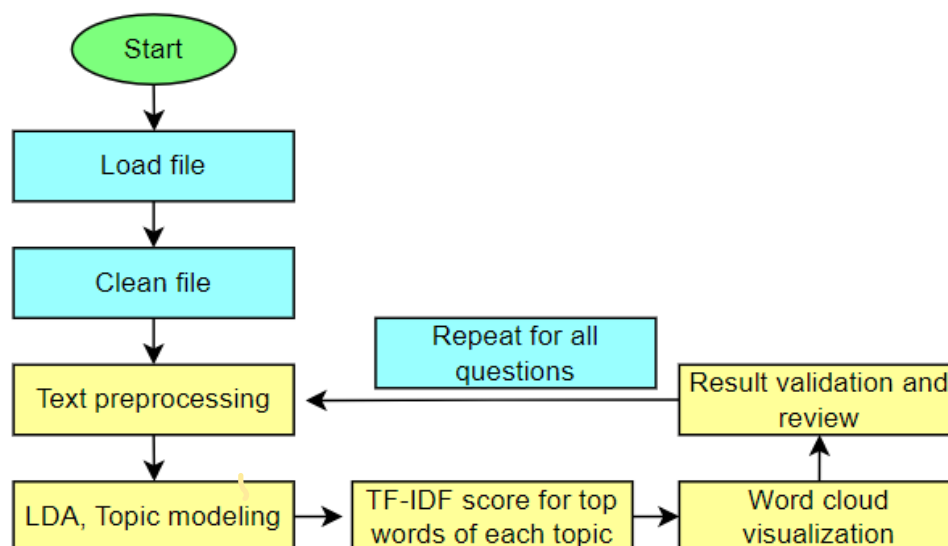


Figure 3.5: Text mining process

Raw responses were collected in Excel sheet. The responses were cleaned and pre-processed by text normalization, which include case conversion, punctuation removal, tokenization and lemmatization. Text normalization reduce the randomness of texts to improve quality and suitable for machine processing (Yze, 2021). Tokenization reduces text to its smallest tokens or units, whereas lemmatization only removes word's inflectional endings and restores a word's base form.

In LDA topic modeling process, model parameters were set according to need. The alpha and beta value were pre-set by Scikit-learn where alpha is  $(1/n\_topics)$  while beta is  $(1/n\_words)$  (ScikitLearn, n.d.). CountVectorizer module was used as feature extraction technique to draw data, converting to numerical representation and create document-term matrix where each row of the matrix represents a document while each column represents a word (Greer, 2018). Figure 3.6 depicts LDA topic modeling process.

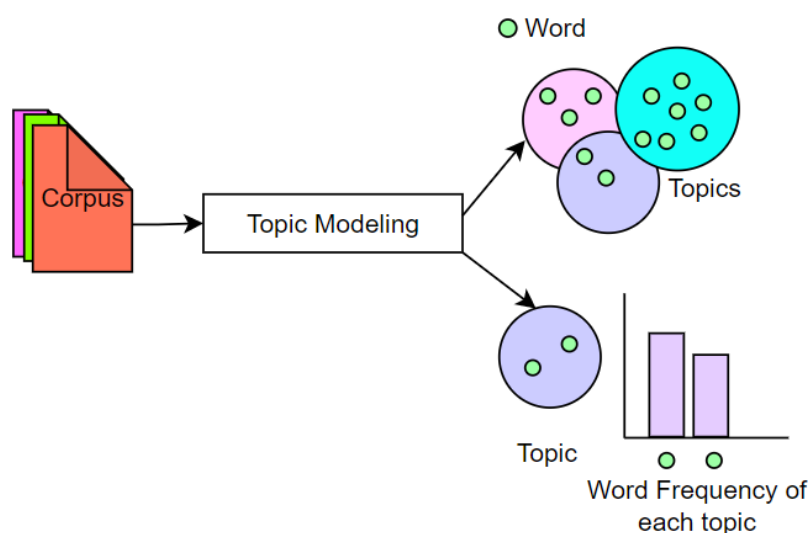


Figure 3.6: LDA topic modeling process

The importance and weight of words within topics were determined by calculating the TF-IDF scores for each word in each topic. A word's term frequency (TF) in a document or corpus is calculated. IDF, or inverse document frequency is the logarithm of ratio of a word's document frequency to the total number of documents in a corpus (Chiusano, 2022). The more important and pertinent the word, the higher its TF-IDF score.



### **3.6 Summary**

Survey form was finalized and developed upon pilot study completion. Convenience sampling method was used to collect responses. The responses were analysed by statistical tests and topic modeling.

## CHAPTER 4

### RESULTS AND DISCUSSION

#### 4.1 Introduction

The data analysis results were thoroughly reported and examined in this chapter. Principal findings and implications were discussed in light of the published literature.

#### 4.2 Socio-demographic characteristic of study respondents

Table 4.2 presents the baseline data. A total of 162 responses were collected from the study, made up by male as majority (52.5%). Age was grouped by interval scale and it has been of particular interest of this study. Normality test was performed for age group as 63.58% of respondents were aged between 18 to 24. Table 4.1 shows the normality test result for age group. The Kolmogorov-Smirnov test ( $p < 0.001$ ) resulted in rejection of null hypothesis, the age group data is not normally distributed. The majority of respondents identified as being of Chinese descent. A small percentage of respondents (3.7%) were either divorced or separated. 91.4% have attained pre-university, diploma, degree or higher education. A diverse sample was indicated by the wide variations in the distribution of respondents by working industry or field of study and monthly income.

Table 4.1: Test of normality and statistics for age

Age		
N	Valid, Missing	162, 0
Kolmogorov-Smirnov <sup>a</sup>		
	Statistic	.376
	df	162
	Sig.	.000
Mean		1.72
Median		1.00
Mode		1
Skewness		1.617
Kurtosis		1.645
Std. Error of Kurtosis		.379

Table 4.2: Sociodemographic Characteristics

Parameters		(n=162)	n (%)
Gender	Male	85	52.5%
	Female	76	46.9%
	Others	1	0.6%
Age	18-24	103	63.6%
	25-34	29	17.9%
	35-44	9	5.6%
	45-54	15	9.3%
	55-64	5	3.1%
	≥ 65	1	0.6%
	Ethnicity	Chinese	112
Malay		24	14.8%
Indian		20	12.3%
Bumiputera		3	1.9%
Others		3	1.9%
Marital status	Single	129	79.6%
	Married	27	16.7%
	Widowed	0	0.0%
	Divorced/ Separated	6	3.7%
Place of residence	Klang Valley	105	64.8%
	West Malaysia (exclude K.V.)	53	32.7%
	East Malaysia	4	2.5%
Highest educational attainment	Pre and secondary	14	8.6%
	Pre-U, diploma and degree	132	81.5%
	Master's and PhD	16	9.9%
Employment status	Unemployed	68	42.0%
	Employed	8	4.9%
	Student	86	53.1%
Working industry / field of study	Healthcare, Pharmaceutical	40	24.7%
	Finance and economics	23	14.2%
	Computer and technology	25	15.4%
	Others	74	45.7%
Monthly income	<1500	94	58.0%
	1500-4849	35	21.6%
	4850-10959	22	13.6%
	>10959	11	6.8%

### **4.3 Closed-ended responses result analysis**

The results of the analysis for Section A, B and C were reported as below.

#### **4.3.1 General perception towards telehealth**

Out of 162 respondents, 54.9% (n= 89) indicated they were aware of telehealth services, while 62.3% (n=101) were unaware that Malaysia offered such services. Gender and knowledge of telemedicine's existence were not significantly associated ( $p>0.05$ ). However, a pattern was observed where those who worked or studied in the healthcare and banking industries knew that Malaysia offers these services better than others. The findings showed a very strong association ( $\phi = 0.356$ ) between working industry or field of study and acknowledgment of local telemedicine services.

Besides, 64.2% (n=104) were aware that telemedicine requires collection of PHI but only 44.2% (n=46) of them read the policies prior to receiving healthcare services. There was a relatively strong association between these two variables ( $\phi = 0.217$ ). Results revealed that there was no relationship between respondents' educational attainment and if they read healthcare policies. It is intriguing to learn that, although 91.35% (n = 148) have educational qualifications above elementary and secondary school, 62.16% (n = 92) of respondents did not read the healthcare policies. From here, it can be presumed that although the public were aware of data sharing requirement, people often consent to legal terms without reading them. Privacy policy should be always read especially prior receiving healthcare treatments to understand data collection purposes. Table 4.3 presents results on Malaysian's knowledge of telemedicine.

Table 4.3: Knowledge towards telemedicine

Have you heard the services			Yes	No	Fisher's Exact / X <sup>2</sup> Test Value	p-value	Phi/ Cramer's V value
Gender	Male	n (% Within Gender)	48 (56.5%)	37 (43.5%)	1.009	0.801	-
	Female	n (% Within Gender)	40 (52.6%)	36 (47.4%)			
	Non-binary	n (% Within Gender)	1 (100.0%)	0			
Total		n (% Within Gender)	89 (54.9%)	73 (45.1%)			
Know Malaysia provides the services							
Gender	Male	n (% Within Gender)	27 (31.8%)	58 (68.2%)	3.415	0.138	-
	Female	n (% Within Gender)	34 (44.7%)	42 (55.3%)			
	Non-binary	n (% Within Gender)	0	1 (100.0%)			
Total		n (% Within Gender)	61 (37.7%)	101 (62.3%)			
Working Industry/ Field of Study	Healthcare , pharmaceutical	n (% Within Field)	24 (60.0%)	16 (40.0%)	20.549	<0.001 ***	0.356
	Finance, economics	n (% Within Field)	13 (56.5%)	10 (43.5%)			
	Computer and technology	n (% Within Field)	4 (16.0%)	21 (84.0%)			
	Others	n (% Within Field)	20 (27.0%)	54 (73.0%)			
Total		n (% Within Field)	61 (37.7%)	101 (62.3%)			
Read policy							
Aware collection PHI	Yes	n (% within Aware collection)	46 (44.2%)	58 (55.8%)	7.654	0.006**	0.217
	No	n (% within Aware collection)	13 (22.4%)	45 (77.6%)			
Total		n (% within Aware collection)	59 (36.4%)	103 (63.6%)			
Educational attainment	Pre and secondary	n (% within attainment)	3 (21.4%)	11 (78.6%)	1.758	0.415	-
	Pre-U, Diploma and Degree	n (% within attainment)	49 (37.1%)	83 (62.9%)			
	Masters and PhD	n (% within attainment)	7 (43.8%)	9 (56.3%)			
Total		n (% within attainment)	59 (36.4%)	103 (63.6%)			

Note: \*p<0.05, \*\*p<0.01, \*\*\*p<0.001

Figure 4.1 shows the Likert scale bar chart of general perception towards telemedicine. Over half of the respondents agreed to the ease of use and usefulness of telemedicine (56.8% and 59.9%). Majority were neutral on its reliability and efficiency. However, total of 18.6% strongly and somewhat disagreed on telemedicine's reliability. Klang Valley residents expressed greater disagreement towards the reliability and efficiency of services as compared to those in other parts of Malaysia. Nevertheless, this difference was statistically insignificant ( $p>0.05$ ). Besides, it is found that poor level of understanding on this topic was high at 17.9%, which is similar to results of study conducted by Aggarwal et al. (2021) in the UK and (Lee, Wong and Lee, 2020) in Selangor, Malaysia. Data indicated that Malaysians showed optimism towards its usefulness although unfamiliar with telehealth. Table 4.4 shows the perceived reliability and efficiency by region.

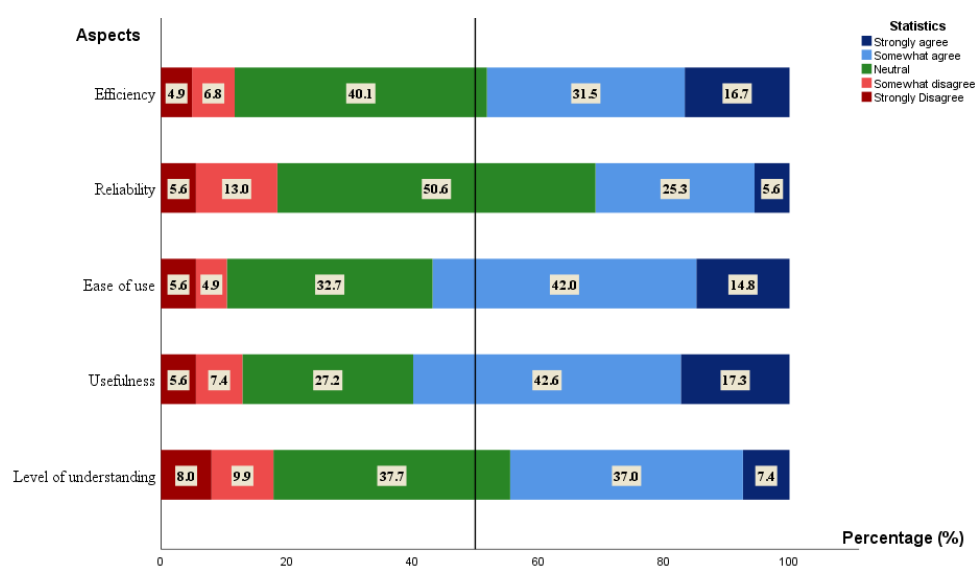


Figure 4.1: Likert scale bar chart of general perception towards telemedicine

Table 4.4: Perceived reliability and efficiency by region

Perceived reliability and efficiency by region		Klang Valley		West Malaysia (exclude K.V.)		East Malaysia		Fisher's Exact / X <sup>2</sup> Test Value	p-value
		n	Percentage (%)	n	Percentage (%)	n	Percentage (%)		
Reliability	Strongly disagree	7	6.70	2	3.80	0	0.00	12.241	0.378
	Somewhat disagree	14	13.30	6	11.30	1	25.00		
	Total	21	20	8	15.10	1	25.00		
Efficiency	Strongly disagree	6	5.70	2	3.80	0	0.00	7.315	0.845
	Somewhat disagree	7	6.70	4	7.50	0	0.00		
	Total	13	12.40	6	11.30	0	0.00		

Note: \* $p<0.05$ , \*\* $p<0.01$ , \*\*\* $p<0.001$

Respondent count for divorced or separated was only 6 out of 162 responses, standing only 3.7% of the total responses. Lower sample size could contribute to higher degree of random variation in data and affect the statistical power of the test (Deziel, 2018). Table 4.5 presents the results of perceived efficiency by marital status which exclude divorced and separated group. Results revealed there is a strong association between marital status and perceived efficiency ( $\phi=0.244$ ). Married patients tend to agree more on telemedicine's efficiency, hectic work life and having time constraints for family might be reasons to find virtual healthcare appointments timesaving.

Table 4.5: Perceived efficiency by marital status

Marital status		Single		Married		Fisher's Exact / X <sup>2</sup> Test Value	p-value	Phi/ Cramer's V value
		n	Percentage (%)	n	Percentage (%)			
Efficiency	Strongly Disagree	8	6.20	0	0.00	9.303	0.049*	0.244
	Somewhat disagree	6	4.70	5	18.50			
	Neutral	50	38.80	11	40.70			
	Somewhat agree	42	32.60	9	33.30			
	Strongly agree	23	17.80	2	7.40			

Note: \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$

### 4.3.2 Data sharing acceptability

Figure 4.2 shows the Likert scale bar chart for data sharing willingness by information type.

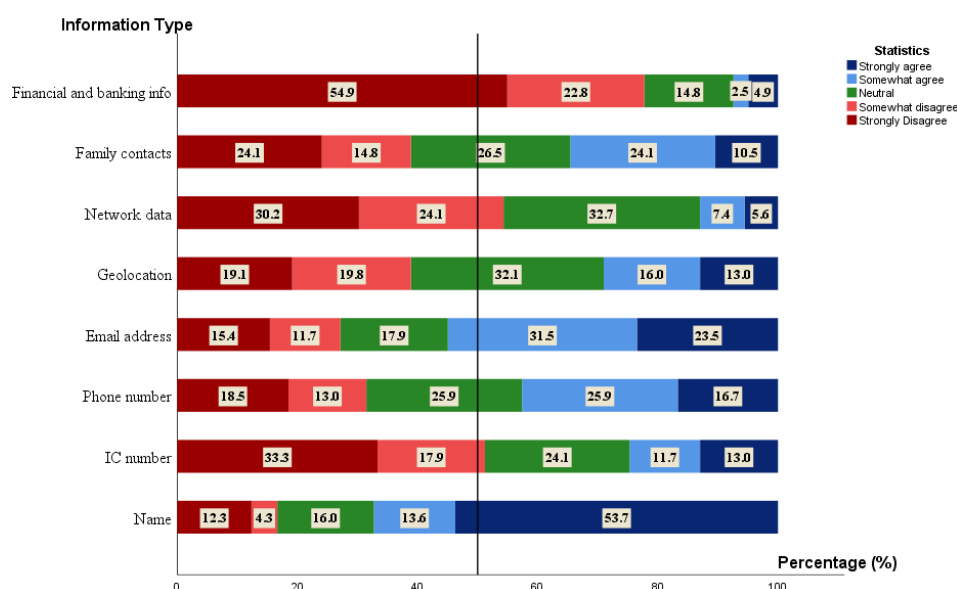


Figure 4.2: Likert scale bar chart of different data sharing willingness

Among all information, patients were less willing (somewhat and strongly disagree) to share IC number (51.2%), network data (54.3%), and financial information (77.7%). Similar result was obtained from O'Brien et al. (2019) as more than half of patients in the States refused to disclose SSN. For financial information, over half (54.9%) of patients strongly opposed on sharing such information. This result aligns with Seltzer et al. (2019) as patients in the States were least open to share financial and location data. Contrary to the mentioned study, Malaysians were open to share geolocation and this may due to different cultural norms. Table 4.6 presents financial data sharing willingness by working industry or field of study and monthly income. Data revealed that patients who worked or studied in healthcare and pharmaceutical industries were more reluctant to share financial data (85% vs 73.9% vs 76% vs 75.7%). In fact, patients with higher monthly income ( $\geq$  RM 4850) tend to strongly oppose disclosing such information. However, this difference was statistically insignificant ( $p > 0.05$ ). Malaysians involved in information and technology (IT) and healthcare industries were more vigilant of breaches, failures, and crimes. More evidence to demonstrate clinical benefits of telehealth services will be essential to address the perceived reliability and efficiency limitations. Besides, clear communication channels should be provided to allow patients to enquire or voice concerns on the purpose of financial data collection, as suggested by (O'Brien et al., 2019).

Table 4.6: Financial data sharing willingness by working industry or field of study and monthly income

Working Industry or Field of Study		Healthcare & Pharmaceutical		Finance & Economics		Computer & Technology		Others		Fisher's Exact / X <sup>2</sup> Test Value	p-value
		n, Percentage (%)									
Financial and banking information	Strongly Disagree	27	67.50	13	56.50	10	40.00	39	52.70	12.548	0.325
	Somewhat disagree	7	17.50	4	17.40	9	36.00	17	23.00		
	Neutral	3	7.50	3	13.00	6	24.00	12	16.20		
	Somewhat agree	2	5.00	1	4.30	0	0.00	1	1.40		
	Strongly agree	1	2.50	2	8.70	0	0.00	5	6.80		



Table 4.6 (continued)

Monthly income (RM)		<1500		1500-4849		4850-10959		>10959		Fisher's Exact / X <sup>2</sup> Test Value	p-value
		n, Percentage (%)									
Financial and banking information	Strongly Disagree	51	54.30	17	48.60	14	63.60	7	63.60	6.077	0.914
	Somewhat disagree	21	22.30	9	25.70	4	18.20	3	27.30		
	Neutral	16	17.00	6	17.10	1	4.50	1	9.10		
	Somewhat agree	2	2.10	1	2.90	1	4.50	0	0.00		
	Strongly agree	4	4.30	2	5.70	2	9.10	0	0.00		

Table 4.7 shows the phone number sharing willingness by age. Results revealed that those above 24 years old were more strongly disagreed to disclose their phone numbers than those below (25.4% vs 14.6%). There is a very strong association between age and phone number sharing willingness ( $\phi=0.252$ ).

Table 4.7: Phone number sharing willingness by age

Age		18-24		above 24		Fisher's Exact / X <sup>2</sup> Test Value	p-value	Phi / Cramer's V value
		n, Percentage (%)						
Phone number	Strongly disagree	15	14.6	15	25.4	10.324	0.035*	0.252
	Somewhat disagree	11	10.7	10	16.0			
	Neutral	30	29.1	12	20.3			
	Somewhat agree	33	32.0	9	15.3			
	Strongly agree	14	13.6	13	22.0			

Note: \* $p<0.05$ , \*\* $p<0.01$ , \*\*\* $p<0.001$

Figure 4.3 shows the Likert scale bar chart for medical data sharing willingness. Majority presented positive attitudes toward sharing medical information. Nevertheless, genetic data and insurance details received slightly more negative feedback (18.5% and 14.2% strongly disagreed). Results revealed that patients were more inclined to keep genetic and insurance information private. It could be to avoid discrimination, genome research or commercial purposes. Initiatives should be made to lead way to open science yet protecting patient's privacy and rights.

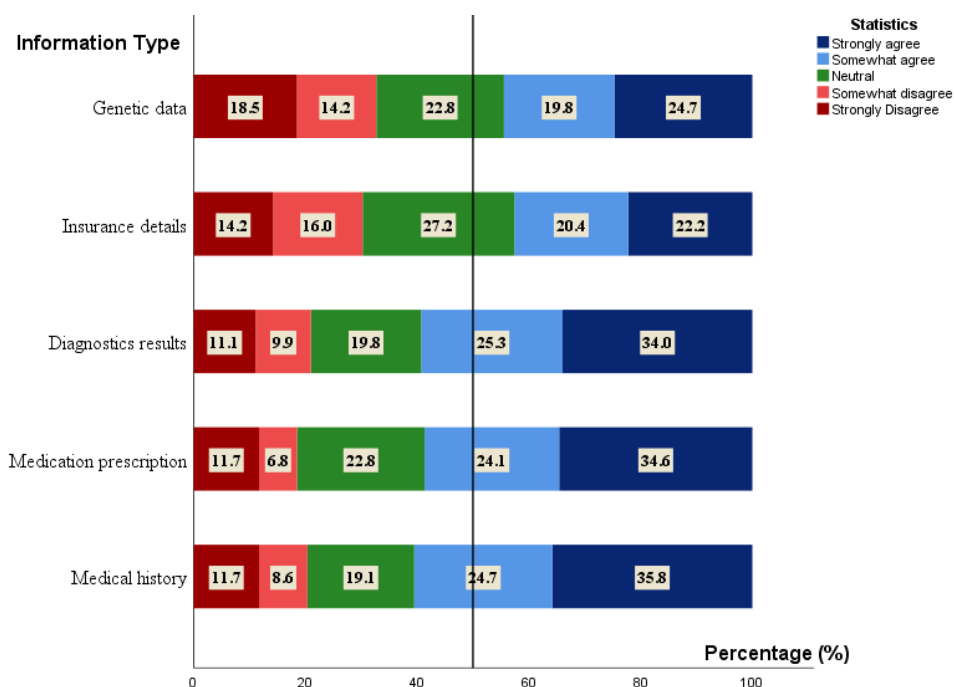


Figure 4.3: Likert scale bar chart for medical data sharing willingness

Malaysians who were aged above 24 were less willing to disclose diagnostic results to healthcare organizations. The result was statistically significant ( $p < 0.05$ ), there was a strong association between these variables ( $\phi = 0.246$ ). Table 4.8 shows diagnostic data sharing willingness results by age.

Table 4.8: Diagnostic data sharing willingness by age

Age		18-24		above 24		Fisher's Exact / $\chi^2$ Test Value	p- value	Phi / Cramer's V value
		n	Percentage (%)	n	Percentage (%)			
Diagnostic results	Strongly Disagree	8	7.8	10	16.9	9.838	0.043*	0.246
	Somewhat disagree	7	6.8	9	15.3			
	Neutral	19	18.4	13	22.0			
	Somewhat agree	32	31.1	9	15.3			
	Strongly agree	37	35.9	18	30.5			

Note: \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$

Figures 4.4 and 4.5 demonstrate Malaysians' willingness to disclose and grant permission to various organizations to keep PHI respectively. Results found all organizations received nearly symmetry votes on sharing and safeguarding healthcare information except for IT companies. Over half of the patients (50.6% and 51.9%) somewhat and strongly disagreed with granting the

IT industry authority. Majority of Malaysians did not trust IT companies with their private data, let alone PHI. They were hesitant to share PHI with them to aggregate and monetize on the information since the firms largely dominate the advertisement and social media platforms. The findings have similar pattern with studies by Aggarwal et al. (2021) and Bell, Ohno-Machado and Grando (2014), where patients were uncomfortable disclosing data to for-profit entities. Malaysians were rational in selecting “data custodians”, and hence trust healthcare and government authorities way more than academic researchers and not-for-profit companies in handling PHI which makes good sense.

Patients have the most confidence in the local public healthcare departments to own (55.6%) and safeguard (48.2%) these data. In terms of data sharing and handling, respondents also believed that insurance corporations, pharmaceutical companies, and academic researchers were more reliable. From this, it may be inferred that these healthcare-related parties held high levels of trust, which could be linked to how well organizations are at present providing their services.

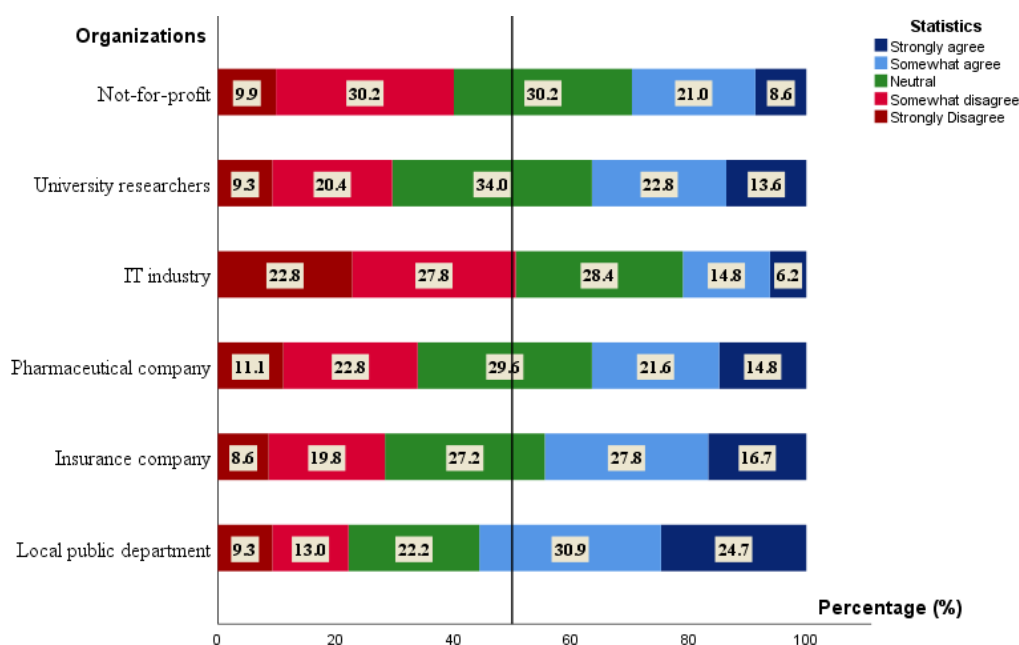


Figure 4.4: Likert scale chart of willingness to share PHI to organizations

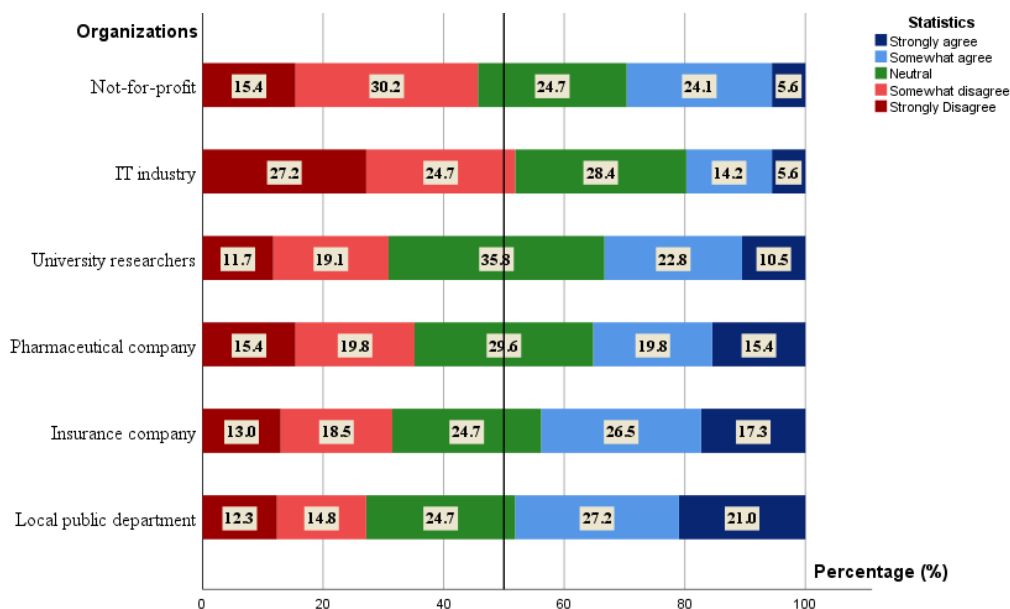


Figure 4.5: Likert scale chart of willingness to allow organizations to keep PHI

Age group and patient data sharing willingness to organizations were statistically significant ( $p < 0.05$ ). Malaysians aged over 24 typically were more reluctant to share PHI and grant authority to local public healthcare departments, pharmaceutical companies and IT industries to handle and safeguard this information. The results have similar trend in the study conducted by Kim and Choi (2019) in South Korea. Table 4.9 below shows the Cramer's V values from measure of strength of association. The findings showed a substantial difference in PHI privacy awareness among Malaysians over the age of 24. This is encouraging as it could potentially prevent serious privacy risks. Besides, awareness must be raised to allow younger generation beware of what they signed up for before disclosing personal data and PHI.

Table 4.9: Results on data sharing willingness by age group

Age		18-24		above 24		Fisher's Exact / X <sup>2</sup> Test Value	p- value	Phi / Cramer's V value
		n	Percentage (%)	n	Percentage (%)			
(share) Local public healthcare department	Strongly disagree	4	3.9	11	18.60	15.468	0.004**	0.309
	Somewhat disagree	13	12.6	8	13.60			
	Neutral	21	20.40	15	25.40			
	Somewhat agree	32	31.10	18	30.50			
	Strongly agree	33	32.00	7	11.90			

Table 4.9 (continued)

(share) Pharmaceutical company	Strongly disagree	5	4.90	13	22.00	16.505	0.002**	0.319
	Somewhat disagree	21	20.40	16	27.10			
	Neutral	31	30.10	17	28.80			
	Somewhat agree	26	25.20	9	15.30			
	Strongly agree	20	19.40	4	6.80			
(keep) Local public healthcare department	Strongly disagree	8	7.80	12	20.30	10.283	0.036*	0.036
	Somewhat disagree	14	13.60	10	16.90			
	Neutral	25	24.30	15	25.40			
	Somewhat agree	28	27.20	16	27.10			
	Strongly agree	28	27.20	6	10.20			
(keep) Pharmaceutical company	Strongly disagree	7	6.80	18	30.50	17.391	0.002**	0.328
	Somewhat disagree	21	20.40	11	18.60			
	Neutral	32	31.10	16	27.10			
	Somewhat agree	24	23.30	8	13.60			
	Strongly agree	19	18.40	6	10.20			
(keep) IT industry	Strongly disagree	24	23.30	20	33.90	9.961	0.041*	0.248
	Somewhat disagree	23	22.30	17	28.80			
	Neutral	34	33.00	12	20.30			
	Somewhat agree	13	12.60	10	16.90			
	Strongly agree	9	8.70	0	0.00			

Note: \*p<0.05, \*\*p<0.01, \*\*\*p<0.001

#### 4.4 Open-ended responses result analysis

Main topics were identified and described in detail, highlighting the main themes, ideas or contexts emerged from the responses assigned to each topic. The percentage distribution of responses was assigned to each topic out of the total of 162 responses.

**Q :** Please describe reasons of reading or not reading the privacy policies of healthcare institution before sharing PHI.

Results :

- Topic 1 (19.75%):

The most important terms in this topic are “lengthy”, “lazy”, “complicate”, and “ensure\_understand”. It seems to be related to the complexity, length and purpose of reading privacy policies. TF-IDF scores suggested that respondents are reluctant to read policies because they are too long and difficult to understand. Those who read the policies aimed to ensure they understand the terms better before using telemedicine services.

- Topic 2 (17.9%):

The most important terms in this topic are “privacy”, “personal”, “understand” and “policy”. It seems to be relevant to the importance of privacy and need to protect personal information. TF-IDF scores suggested that respondents read privacy policy to ensure they understand them and own the rights to protect their private information from telemedicine services.

- Topic 3 (16.67%):

The most important terms in this topic appear to be similar in Topic 1. However, this topic targets mainly the length of the healthcare policies. They often were lazy and usually agreed to the terms as they would like to ensure that they can receive medical care and services.

- Topic 4 (25.31%):

The most important terms in this topic are “read”, “data”, “ensure\_understand” and “long”. It seems to be related to the importance of understanding personal data collection purposes and privacy policies. TF-IDF scores suggested that respondents took time to read long policies to ensure they understand the terms prior sharing their personal data to healthcare organizations.

- Topic 5 (20.37%):

This topic seems to have the same context as Topic 4. Although the important terms in this topic appear to be similar to others, some respondents did not read the policies as they failed to understand the policies and doubted to trust in healthcare service providers.

### **Key Takeaways**

Overall, respondents who read the privacy policies want to ensure they understand the terms prior sharing personal data to healthcare organizations. 36.42% of the respondents were lazy to read them felt that the contents are too

lengthy, they did not understand well and often agreed to receive medical and telemedicine services. Besides, 20.37% have poor trust in healthcare organizations in collecting personal and healthcare data.

**Q:** *What are your thoughts/ feelings/ opinions after reading facts below? You may describe your answer.*

*Telemedicine and Telehealth Legal Framework Overview (2022) by Baker McKenzie had reported that the Poisons (Amendment) Bill 2022 ("Bill") was recently tabled in the Malaysian Parliament. It seeks to, among others, regulate the use of electronic documents relating to the dispensation of drugs. It is unclear whether and when the Bill will become law. In other words, while the Malaysian Parliament has passed the Telemedicine Act 1997, it is still not yet in force.*

Results:

- Topic 1 (20.62%):

The most important terms in this topic are “telemedicine”, “enforce”, “information”, and “think”. It seems to be related to the concerns about legal aspects of telemedicine use. TF-IDF scores suggested that respondents felt the need to enforce the Telemedicine Act 1977 for personal data protection. They believed that the enforcement of the law can protect public’s rights.

- Topic 2 (21.88%):

The most important terms in this topic are similar to Topic 1. However, there are words suggesting that the law should be enforced very soon to ensure healthcare organizations will use the information for good. TF-IDF scores suggested that respondents were optimistic about the implementation of Telemedicine Act 1977.

- Topic 3 (19.38%):

The most important terms in this topic are “telemedicine”, “people”, “government”, and “good”. It seems to be related to the respondents looking forward to the implementation of law towards telemedicine. TF-IDF scores suggested that respondents wished to ensure the electronic dispensation of drugs and prescription processes are well documented, ensuring ethical and lawful dispensation.

- Topic 4 (22.5%):

This topic is very similar to Topic 1 and 2 as well. Respondents felt that there is a need to enforce the act to ensure telemedicine services' proper use and medical care for people are well.

- Topic 5 (15.62%):

The most important terms in this topic resemble those in others. However, it includes words such as "regulate" and "data". It seems to be related to the respondents being concerned about data privacy and its security in the use of telemedicine services thus there is a need to regulate it.

### **Key Takeaways**

Overall, the question resulted in concerns regarding the lack of enforcement of Telemedicine Act 1977. All of them wished and advocated its implementation, expressing optimism toward the benefits of the law. They believed enforcement of the act can address data privacy and security issues during the process of receiving telemedicine services such as electronic prescriptions of medications.

**Q:** *What are your thoughts/ feelings/ opinions after reading facts below? You may describe your answer.*

*E-pharmacy is available in Malaysia, medications can be procured online and delivered at doorstep. Over-the-counter drugs can be purchased easily but prescription drugs must be prescribed by physician upon e-consultation with electronic signature before dispensing. Information being collected and recorded include drug information (e.g., quantity, dosage) and patient information (e.g., address, ID number). However, without appointment of licensed courier company, delivery of drugs directly to a patient's residence can be done.*

### **Results:**

- Topic 1 (20.37%):

The most important terms in this topic are "courier", "drug", "delivery", and "company". It appears that respondents are repeating the words stated in question. However, TF-IDF scores of "license" and "need" suggested that respondents felt concerned on the safety and legality issues that will be emerged when drug delivery can be done through unlicensed couriers. They may also feel the need of company to appoint licensed courier for drug delivery.



- Topic 2 (17.9%):

The most important terms in this topic are “drug”, “medicine”, “pharmacy”, and “risk”. It seems to be related to the risk associated with drug prescription and purchase on e-pharmacy. The TF-IDF scores suggested that respondents felt that risks can be brought easily if they purchase drugs online.

- Topic 3 (34.57%):

The important terms in this topic are similar to those in Topic 1. This topic is related to the optimism respondents showed towards the idea of having drug delivery straight to doorsteps through the delivery company. Respondents were discussing on e-pharmacy medicine prescription and drug delivery services provided.

- Topic 4 (12.35%):

The important terms in this topic are “courier”, “delivery”, “company” and “pharmacy”. These top words do not seem to give meaningful interpretation, but TF-IDF scores suggested that patients have concerns regarding privacy issues and the reliability of e-pharmacy on drug delivery. The scores for words “license”, “convenient” and “privacy” indicate that despite having drug delivery through courier delivery companies present to be convenient, personal information disclosure may raise privacy concerns.

- Topic 5 (14.81%):

The most important terms in this topic are “drug”, “risk”, “dangerous”, and “make”. This topic mainly focuses on the potential risk associated with e-pharmacy although the process may be easy. TF-IDF scores suggested that respondents felt worried about the possibility for delivery of medications without licensed courier personnel may bring negative consequences.

### **Key Takeaways**

Overall, respondents understood that purchase of medication or drug prescription through e-pharmacy may be convenient. Nevertheless, 65.43% still have personal information and privacy concerns pertaining delivery of drug to doorstep without a licensed courier company.

*Q: Malaysia is expected to be an aged society in 2044 ( with 14% of population aged 65 and over), there may be growth in geriatric telerehabilitation and telemedicine use. Do you think Malaysia's current telehealth guidelines and framework are sufficient to face the coming future ? Please express your concerns and thoughts on privacy issues on geriatric telemedicine and telerehabilitation utilization.*

Results:

Of all 162 respondents, 24.69% believed Malaysia's current framework was insufficient to face the growth in geriatric telehealth use.

- Topic 1 (20.37%):

The most important terms in this topic are “medical”, “data”, “information”, and “people”. It seems to be related to the concerns about legal aspects of telemedicine use. TF-IDF scores suggested that respondents were expressing general concerns about the personal information protection of telemedicine use by geriatric population.

- Topic 2 (12.35%):

The most important terms in this topic are “elder”, “understand”, “private”, and “information”. It seems to be related to the lack of understanding by geriatric patients towards telemedicine use. There are also words such as “misuse” and “health” which could be suggesting that incomprehension of telemedicine use can lead to misuse of private healthcare information in the context of geriatric care.

- Topic 3 (12.96%):

The most important terms in this topic are “lot”, “understand”, “information”, and “people”. It appears that there is too much information for elderlies to understand about telemedicine services. TF-IDF scores for other words also suggested that the elderly population lacks digital literacy, which makes it tougher for them to understand advanced technology such as telemedicine services.

- Topic 4 (34.57%):

The most important terms in this topic are “telemedicine”, “issue”, “data”, and “information”. This topic is similar to Topic 1 on the concerns regarding personal healthcare data being used and protected. TF-IDF scores for other

words such as “company” and “need” also suggested that it is necessary for companies involved to address this issue. From another perspective, the words might indicate that geriatric patients may require company during their telemedicine use.

- Topic 5 (19.75%):

The most important terms in this topic are “telemedicine”, “geriatric”, “information”, and “privacy”. This topic is similar to Topic 1 and 3. TF-IDF scores in other words suggested that unfamiliarity with telemedicine and telerehabilitation use by elderly would cause privacy issues and healthcare information to be leaked or disclosed.

### **Key Takeaways**

Overall, this question raised concerns about the lack of digital literacy among geriatric population and the potential for healthcare information leaks and privacy issues when receiving telemedicine and telerehabilitation services. Besides, 34.57% suggested that elderly require someone to accompany during the process. The same outcome on the unreadiness of elderly was presented by Lam et al. (2020) and Fadzil et al. (2022).

### **Q: Do you think sharing digital health information benefit you in any way?**

Analysis was performed based on groups. Responses were grouped by “yes”, “yes, but...”, “no, but...” and “no”.

Results for “yes”:

- Topic 1 (36.36%):

The most important terms in this topic are “ensure\_understand”, “medical”, “useful”, and “doctor”. This topic is related to respondents considering sharing PHI is beneficial to them as they can understand how this information can be useful for doctors while providing medical services.

- Topic 2 (13.22%):

The most important terms in this topic are “beneficial”, and “people”. It appears that respondents have positive attitude towards sharing of digital health information. TF-IDF scores for other words such as “maybe”, “health” and “help” suggest that they believed sharing of this information to hospitals or

healthcare organizations is beneficial for improving healthcare treatments of people in general.

- Topic 3 (14.88%):

The most important terms in this topic are “hospital”, “digital”, “yes” and “data”. This topic is about respondents agreed to sharing digital health information with hospital. Other words such as “health” and “information” did not seem to show distinctive context.

- Topic 4 (8.26%):

The most important terms in this topic are “condition”, “time”, and “yes”. This topic reveals that respondents agreed to sharing PHI as they recognize that sharing digital healthcare information is useful in assisting doctors diagnose and treat conditions online. Besides, other words such as “helpful” and “track” reveal that sharing digital data helps to keep track of patients’ health online efficiently.

- Topic 5 (27.27%):

The most important terms in this topic are “yes”, and “information”. Other words such as “use”, “research”, “health” and “better” imply that sharing such information digitally can facilitate healthcare research which will eventually benefit the respondents.

### **Key Takeaways**

Out of 162 respondents, 74.69% of them considered sharing digital health information personally benefit them. Overall, 49.58% of these respondents believed that it could help healthcare providers in improving patient care. They believed that this could ease telemedicine processes and allow healthcare practitioners to provide treatments accurately. 8.26% considered that this could save both parties’ time to provide and receive medical care. Additionally, 27.27% of respondents considered that their PHI could aid medical research, ultimately being beneficial to themselves.

Results for “yes, but...”:

- Topic 1 (50%):

The most important terms in this topic are “share” and “people”. Other words such as “public”, “medication” and “doctor” could imply that this group of

respondents felt fine to share PHI digitally, as it could benefit them and the public. However, this does not apply to those with mental health-related issues. This topic also suggests that respondents with mental health issues opined that they would be negatively impacted when the information is being shared. For instance, they were worried to be treated differently, affecting their social and spiritual health.

- Topic 2 (12.5%):

The most important terms in this topic are “share”, “medication” and “mental”. This topic is similar to Topic 1. It appears that respondents have positive attitude towards sharing digital health information but for assisting in better medication prescription and treatments only. They have concerns and disagree to the context if it is for those with mental health issues.

- Topic 3 (37.5%):

The most important terms in this topic are “share”, and “information”. This topic does not present a distinctive context on the topic. Other words such as “correct” and “think” suggest that this topic might be about respondents agreed to the question but only if sharing such information could assist healthcare practitioners in providing the correct treatments to them.

### **Key Takeaways**

Out of 162 respondents, 4.94% of respondents conditionally agreed that sharing digital health information could benefit them. 62.5% of them disagreed that it is advantageous for those with mental disorders. Additionally, 55% believe it can benefit them only in assisting healthcare providers to administer proper treatments.

Results for “no, but...”:

- Topic 1 (38.46%):

The most important terms in this topic are “really”, “contribute” and “info”. This topic implies that respondents did not think sharing digital information can be helpful for them but may contribute to other aspects. Other words such as “research”, and “help” reveal that respondents did not agree on the question but it may support healthcare research.

- Topic 2 (23.08%):

The most important terms in this topic are “use”, “really” and “depend”. Other words such as “beneficial” and “research” suggest that this topic is similar to Topic 1. It appears that respondents thought it depends on individuals but disagreed that it is beneficial to them personally but great for research purposes.

- Topic 3 (38.46%):

The TF-IDF scores for the words in this topic are equal. However, words such as “directly”, “benefit” and “research” suggest that respondents disagreed that it could directly benefit them but can facilitate healthcare research purposes.

### **Key Takeaways**

Out of 162 respondents, 8.02% of respondents denied that sharing digital health data can benefit them but could leave positive impacts on healthcare research and studies.

Results for “no”:

- Topic 1 (55%):

The most important terms in this topic are “benefit”, “unless”, “party” and “information”. This topic reveals that respondents disagreed on sharing digital information can be helpful for them as they perceived the information shared will not only be in the hands of physicians. They wanted to keep this information as confidential as possible.

- Topic 2 (20.0%):

The most important terms in this topic are “information”, “think”, “dont” and “benefit”. It appears that respondents strongly opposed the idea that sharing PHI would benefit them. Other words such as “personal” and “data” suggest that respondents viewed the data shared as very personal and private information thus sharing it do not benefit them. For instance, data misuse might be their worried.

- Topic 3 (25.5%)

The most important terms in this topic are “information”, “think”, “people” and “ensure\_understand”. This topic is related to respondents being pessimistic as they wished to ensure to understand clearly before sharing such information to other people. Other words such as “digital” and “health” do not give strong

meaning but might be related to some concerns that respondents have towards telemedicine.

### **Key Takeaways**

Out of 162 respondents, 12.35% of respondents disagreed that sharing digital health data can benefit them. 75% were worried about putting such confidential data in third party's hands and leading to crime while 25.5% have low confidence towards telemedicine and would make sure they fully understand it.

**Q:** *Do you wish to know the purpose of healthcare organizations in keeping your personal health information? (E.g., research purposes, commercial purposes, defense of legal claims, administration)*

*Please describe why you would / wouldn't wish to the above question.*

Analysis was performed based on groups. Responses were grouped by “yes”, “yes”, “insignificant” and “no”.

Results for “yes”:

- Topic 1 (35.25%):

The most important terms in this topic are “information”, “use”, and “right”. This topic appears to be related to respondents wished to ensure the healthcare information shared to organizations is being used rightfully. Other words such as “purpose”, “ensure\_understand” and “data” imply that respondents were deeply interested to understand the PHI data collection purpose.

- Topic 2 (26.46%):

The most important terms in this topic are “make”, “scam”, “expose”, “like” and “use”. This topic appears to be related to respondents' concerns on the illegal use of PHI such as frauds and scams that could expose their data. Word such as “ensure\_understand” emphasizes respondents' desire to understand the data collection purpose to avoid being crime victims.

- Topic 3 (10.07%):

The most important terms in this topic are “record”, “issue”, “case” and “disease”. This topic is about respondents who owned some diseases, medical issues and cases wishing to know the reasons for PHI being recorded. Other words such as “ensure\_understand”, “info”, “organization” and “personal” emphasize that respondents were keen to understand the PHI data collection purpose by related organizations.

- Topic 4 (11.51%):

The most important terms in this topic are “patients”, “transparency”, “course”, “information” and “better”. Respondents considered transparency of healthcare providers in collecting PHI of patients in a serious level, similar to the study conducted by Bell et al. (2014). Other words such as “need”, “future” and “research” suggest that they wished to know how the data will be used in the future such as for research purposes.

- Topic 5 (18.71%):

The most important terms in this topic are “healthcare”, “way”, “use”, “ensure” and “information”. This topic appears to be related respondents wanting to protect their PHI and know the collection purpose in general.

### **Key Takeaways**

Overall, 85.8% of respondents responded “yes” when asked if they wish to know the purpose of healthcare organizations in keeping PHI. According to the results, 26.46% of respondents agreed as they wish to avoid fraud and 11.51% of them would want to know if their PHI is used for research purposes.

### Results for “insignificant”:

Overall, 5.56% of respondents responded did not have clear idea of the context. This makes 5.56% of the total responses insignificant for analysis.

### Results for “no”:

- Topic 1 (50%):

The most important terms in this topic are “trust”, “patients”, and “ensure\_understand”. This topic is related to respondents trusting the healthcare organizations and they understood that they should understand the data collection purpose well. The results revealed that this group of respondents has excellent faith on healthcare providers.

- Topic 2 (21.43%):

The most important terms in this topic are “long”, and “patients”. Both words do not directly explain the topic. However, other words such as “work”, “care”, and “wish” seem to imply that respondents did not wish to care so much about



the data collection purpose by healthcare organizations as long as it works well for them.

- Topic 3 (28.57%):

The most important terms in this topic are “use”, “long”, and “health”. These words do not show distinctive context but other words such as “personal” and “information” suggest that the topic is similar as Topic 2, respondents felt that it is all right as long as it would be used for good and improve patients' health.

**Key Takeaways**

Overall, 8.64% of all 162 respondents did not wish to know PHI data collection purposes by healthcare providers. Half of these respondents have great trust in healthcare organization on handling the data collected while the others felt that they do not need to know the data collection purpose as long as it is for the good of their health and work well for them.

*Q: Do you think anonymization approach (removal of identifier such as names, IC numbers) help increase your data sharing acceptability ? From privacy perspective, apart from anonymization, what are other ways and aspects to improve patient data willingness?*

Of 162 responses, 83.33% believed anonymization could increase patient data-sharing willingness. This result aligns with the study published by Aggarwal et al. (2021), in which anonymization of data was strongly supported (83.8%).

Results:

- Topic 1 (16.67%):

The most important terms in this topic are “share”, “make”, “care”, and “protection”. This topic is related to ways in ensuring data privacy and safety to improve data sharing willingness. Other words such as “need”, “improve”, “privacy” and “health” suggested that the need to enhance healthcare data privacy presents a critical factor in increasing data sharing willingness.

- Topic 2 (26.54%):

The most important terms in this topic are “don't know”, “share”, “think”, and “use”. Respondents do not seem to have idea of answering to the question, they may not have a clear understanding on the context. However, TF-IDF scores of words such as “education”, “explain”, and “purpose” indicate that respondents

suggested healthcare organizations or parties related to promoting education on healthcare data sharing, explaining how the data will be utilized, as well as clarifying the data collection purpose.

- Topic 3 (15.43%):

The most important terms in this topic are “information”, “patients”, “protect”, and “use”. This topic appears to be related to respondents who hoped to know that the patient’s information will be protected. Other words such as “good”, “organization” and “data” indicate that healthcare organizations should ensure the healthcare data collected are used for good only.

- Topic 4 (30.86%):

The most important terms in this topic are “healthcare”, “share”, “security”, and “data”. Respondents seemed to suggest that improving patient PHI security during collection may improve data sharing willingness. TF-IDF scores of these words reveal the importance of data security to patients. Hence, healthcare organizations or authorities involved should prioritize it to improve willingness.

- Topic 5 (37.04%):

The most important terms in this topic are “purpose”, “protect”, “ensure\_understand”, and “use”. Respondents suggested that to improve patient data sharing willingness, authorities and healthcare organizations should inform and ensure patients understand the data collection purposes. Additionally, parties involved should find ways to enhance data protection.

### **Key Takeaways**

26.54% of the respondents did not have idea about the context. Overall, respondents suggested that education (26.54%), clear explanation (37.04%) on healthcare data collection can bring positive impacts in improving patient data sharing willingness. The result of this topic is similar to published papers by Karampela, Ouhbi and Isomursu (2019) and Lee, Wong and Lee (2020). Besides, authorities should develop strategies to enhance data privacy of PHI upon collection, ensuring the data is used lawfully. This context is relevant to the study conducted by Grande et al. (2022) which concluded that new protections can potentially increase data-sharing confidence.

*Q: Digital technology gives convenience but may shed light on potential emotional, reputational, economic and physical harms. In the context of sharing personal and health information for telemedicine and telerehabilitation, privacy versus convenience, would you choose the former or latter ? Please elaborate your answer.*

Analysis was performed based on groups. Responses were grouped by “privacy over convenience”, “neither”, “convenience over privacy”.

Result for “privacy over convenience”:

- Topic 1 (4.11%):

The most important terms in this topic are “number”, “person”, “information”, and “misuse”. This topic seems to reveal that more respondents chose privacy over convenience as they want to protect their personal information such as phone number from misuse. TF-IDF scores of other words such as “issue” and “protect” reveal that respondents sought to protect themselves from privacy issues.

- Topic 2 (24.66%):

The most important terms in this topic are “party”, “protect”, “convenience”, and “need”. This topic seems to reveal that respondents chose privacy over convenience as they believed despite convenience often provided, patients should protect their personal data from any party possible. TF-IDF scores of other words such as “security”, “need” and “important” reveal that there is a need to ensure personal data is always secured. Hence, authorities or healthcare organizations should ensure safeguarding this information.

- Topic 3 (52.05%):

The most important terms in this topic are “information”, “share”, “ensure\_understand”, and “health”. This topic seems to reveal that respondents would choose privacy as they wished to ensure they understand the healthcare services terms fully before sharing such information. TF-IDF scores of other words such as “breach”, “choose” and “data” could indicate that respondents are worried about data breaches and hence prefer if they selectively choose the data to be shared. This result supports the study conducted by Bell et al. (2014) in United Kingdom where respondents preferred to have data-sharing control.

- Topic 4 (8.22%):

The most important terms in this topic are “information”, “better”, “medical”, and “solve”. This topic seems to be related to respondents’ belief sharing healthcare data to authorities could help to solve medical issues better. TF-IDF scores of other words such as “think”, “online” and “understand” may indicate that respondents thought they should understand the information clearly when it comes to receiving healthcare services online.

- Topic 5 (10.96%):

The most important terms in this topic are “information”, “convenience”, “feel”, “individual” and “patient”. This topic seems to suggest that respondents felt that choosing between privacy or convenience is depending on individuals and patients. However, there are words such as “secure”, “doctor”, “company” and “share” suggesting that healthcare companies and practitioners hold the responsibility to protect the information shared.

### **Key Takeaways**

Overall, 45.06% of respondents chose privacy over convenience. Among those, 56.16% have concerns about privacy information breaches. Besides, 35.62% of respondents considered healthcare providers to hold the responsibility to protect patients’ healthcare information and personal data shared with them. Additionally, 52.05% of respondents chose privacy over convenience as they believed they ought to comprehend the healthcare services provided well before receiving them.

Results for “neither”:

- Topic 1 (50%):

The most important terms in this topic are “technology”, and “don’t know”. Respondents seemed to not have a clear understanding of the question asked or healthcare technology. TF-IDF scores of other words such as “provide”, “balance” and “certain” revealed that respondents were uncertain about the decision and wish a balance between convenience and privacy can be provided.

- Topic 2 (26.92%):

The most important terms in this topic are “respect”, “think”, and “patients”. This topic seems to be related to respondents’ opinions on respecting patients’

decisions. TF-IDF scores of other words such as “improve”, “privacy” and “convenience” suggested that healthcare providers should enhance privacy security while offering convenience to patients.

- Topic 3 (23.08%):

The most important terms in this topic are “purpose”, “data”, and “ensure\_understand”. This topic seems to be related to respondents wanting to ensure they understand the purpose of sharing PHI or data related before receiving healthcare services. However, TF-IDF scores of other words such as “depends”, “health” and “believe” could indicate that these respondents perceived the choice between privacy and convenience depending on whether they trust the healthcare providers and services.

### **Key Takeaways**

Overall, 16.05% of respondents neither chose privacy nor convenience. Among those, 85.62% thought a balance between both is essential while 50% did not have a clear understanding on the context. Besides, 23.08% of respondents considered that the choice is depending on the trust of patients put in healthcare providers and services. Understanding the primary purpose of data sharing and collection is essential in the process. In the survey conducted by Kato-Lin and Thelen (2022), patients in the States did not consider privacy as a concern during pandemic or other environmental risk factors despite agreeing to its importance.

Results for “convenience over privacy”:

- Topic 1 (20.63%):

The most important terms in this topic are “make” and “ensure\_understand”. These words do not seem to provide meaningful information. TF-IDF scores of other words such as “easier”, “time” and “believe” suggest that prioritizing privacy would take longer time. Hence, respondents believed that convenience made work easier and they believed the healthcare providers.

- Topic 2 (11.11%):

The most important terms in this topic are “ease”, “prefer”, “future”, and “record”. This topic seems to reveal that respondents chose convenience as they prefer to ease their future healthcare processes. TF-IDF scores of other words such as “telemedicine”, “good”, “convenience” and “privacy” suggest that

despite respondents did agree to the good of prioritizing privacy, they opted for convenience especially in use of telemedicine.

- Topic 3 (28.57%):

The most important terms in this topic are “healthcare”, “share”, and “digital”. This topic seems to be related to sharing of healthcare information digitally. TF-IDF scores of other words such as “convenience”, “technology”, “data” and “choose” could indicate that respondents chose convenience as they agreed to the advancement of healthcare technology and sharing of personal data.

- Topic 4 (15.87%)

The most important terms in this topic are “patient”, “fast”, and “world”. This topic seems to be related to the advancement and speed of healthcare services to treat patients. TF-IDF scores of other words such as “technology”, “help” and “time” may indicate that respondents chose convenience as the healthcare technology today are meant to facilitate patients efficiently.

- Topic 5 (23.81%):

The most important terms in this topic are “handle”, “agree”, and “want”. This topic seems to be about respondents agreed and wanted a party to handle their PHI. There are words such as “service”, “need”, “use” and “information” which suggest that respondents allowed healthcare providers to handle and utilize PHI as they need the healthcare services as patients.

### **Key Takeaways**

Overall, 38.89% of respondents chose convenience. Among those, 47.61% chose convenience mainly due to the trust they put in the efficiency and services provided by the current healthcare technology. They believed prioritizing convenience over privacy can help save time optimize the services. Additionally, 23.81% of respondents opted for convenience as they agreed to let healthcare providers to manage their PHI when receiving the services.

#### **4.5 Summary**

Analysis has been performed on all 162 collected responses which were categorized into closed and open-ended responses. More than 61% of respondents responded that they were open to share their family contacts, geolocation, phone number, email address and name. Meanwhile, more than 58 % of respondents were willing to disclose medication prescriptions, medical history and diagnostic results. Malaysians were least willing to disclose financial and genetic data. Age, work or study field, income and marital status were key determinants of data-sharing willingness. They tried to comprehend privacy policies but length and complexity of the policies are the issues. Malaysian advocated the need to enforce Telemedicine Act 1977. Malaysians were concerned about the elderly's digital literacy and the public's mistrust of healthcare organizations. Sharing PHI could benefit most Malaysian and improve healthcare services. However, some were concerned about privacy issues. When it comes to sharing and safeguarding PHI, the public has the least trust in the IT industry. Besides, data privacy enhancement, need for education and clear communication on data collection were emphasized.

## CHAPTER 5

### CONCLUSIONS AND RECOMMENDATIONS

#### 5.1 Conclusions

The problem statement and objectives of this study have been addressed and achieved. A survey questionnaire on patient perception and willingness to share personal data and healthcare information for telemedicine services use in the Malaysian setting was successfully developed. The survey was conducted on 162 respondents. Privacy issues, geriatric use and the public's perspectives on telehealth were discussed by providing room for discussion through this study. Closed-ended responses were analyzed using SPSS with descriptive and inferential statistics. Meanwhile, open-ended responses were analyzed using Python by topic modeling with LDA and TF-IDF scores. All findings were analyzed and reported accordingly.

#### 5.2 Recommendations for future work

For the current study, only bivariate analysis was conducted. More comprehensive analysis using multivariate analysis between variables is highly recommended. Further analysis of different sub-groups such as age, marital status and health status can be done with open-ended responses to examine factors and trends holistically. For instance, consider the relationship between age group and perception or concerns towards geriatric telemedicine use. The combination of these variables can obtain deeper insights in the topic of interest. Additionally, longitudinal study can be employed in future study to examine changes in attitudes and perceptions over certain period such as a year. LDA with time component and Structural Topic Modeling (STM) can be performed to analyze such time-domain responses.



## REFERENCES

Abdul Aziz, M.F., 2020. Telemedicine in Malaysia during Covid-19 pandemic: assessing legal protection for patient's privacy. *Centre For Law and Ethics in Science and Technology*, pp.3–10.

Aggarwal, R., Farag, S., Martin, G., Ashrafian, H. and Darzi, A., 2021. Patient Perceptions on Data Sharing and Applying Artificial Intelligence to Health Care Data: Cross-sectional Survey. *Journal of Medical Internet Research* 2021;23(8):e26162 <https://www.jmir.org/2021/8/e26162>, [online] 23(8), p.e26162. <https://doi.org/10.2196/26162>.

Akoglu, H., 2018. User's guide to correlation coefficients. *Turkish Journal of Emergency Medicine*, [online] 18(3), pp.91–93. <https://doi.org/10.1016/J.TJEM.2018.08.001>.

Ali, N., Khalifa, O., Manaf, A.A. and Olanrewaju, R.F., 2013. ICT in Telemedicine: Conquering Privacy and Security Issues In Health Care Services | electronic Journal of Computer Science and Information Technology. *electronic Journal of Computer Science and Information Technology*, [online] 4(1). Available at: <<https://ejcsit.uniten.edu.my/index.php/ejcsit/article/view/39>> [Accessed 26 June 2022].

Anon. 2023a. *Naluri: The Digital Health Tech Company Of The Future - MedTech Innovator*. [online] Available at: <<https://medtechnovator.org/naluri-the-digital-health-tech-company-of-the-future/>> [Accessed 2 April 2023].

Anon. 2023b. *sklearn.decomposition.LatentDirichletAllocation — scikit-learn 1.2.2 documentation*. [online] ScikitLearn. Available at: <<https://scikit-learn.org/stable/modules/generated/sklearn.decomposition.LatentDirichletAllocation.html>> [Accessed 13 April 2023].

Anwyl-Irvine, A.L., Massonnié, J., Flitton, A., Kirkham, N. and Evershed, J.K., 2020. Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, [online] 52(1), pp.388–407. <https://doi.org/10.3758/S13428-019-01237-X>.

Baker McKenzie, 2022. *Telemedicine and Telehealth Legal Framework Overview*. Chicago .

Bassan, S., 2020. Data privacy considerations for telehealth consumers amid COVID-19. *Journal of Law and the Biosciences*, [online] 7(1), pp.1–12. <https://doi.org/10.1093/JLB/LSAA075>.

Batimalai, K., 2022. *Telemedicine Services In Private Hospitals Rose During Covid-19 - CodeBlue*. [online] CodeBlue. Available at: <<https://codeblue.galencentre.org/2022/05/05/telemedicine-services-in-private-hospitals-rose-during-covid-19/>> [Accessed 26 June 2022].

Bell, E.A., Ohno-Machado, L. and Grando, M.A., 2014. Sharing My Health Data: A Survey of Data Sharing Preferences of Healthy Individuals. *AMIA Annual Symposium Proceedings*, [online] 2014, p.1699. Available at: <<https://pubmed.ncbi.nlm.nih.gov/26419941/>> [Accessed 3 July 2022].

Bestsenny, O., Gilbert, G., Harris, A. and Rost, J., 2021. *Telehealth: A post-COVID-19 reality? | McKinsey*. [online] McKinsey & Company. Available at: <<https://www.mckinsey.com/industries/healthcare-systems-and-services/our-insights/telehealth-a-quarter-trillion-dollar-post-covid-19-reality>> [Accessed 25 June 2022].

Bevans, R., 2020. *Choosing the Right Statistical Test | Types & Examples*. [online] Scribbr. Available at: <<https://www.scribbr.com/statistics/statistical-tests/>> [Accessed 6 April 2023].

Cassel, C. and Bindman, A., 2019. Risk, Benefit, and Fairness in a Big Data World. *Journal of the American Medical Association*, 322(2), pp.105–106. <https://doi.org/10.1001/JAMA.2019.9523>.

Centers for Disease Control and Prevention, 2022. *Health Insurance Portability and Accountability Act of 1996 (HIPAA) | CDC*. [online] Centers of Disease Control and Prevention. Available at: <<https://www.cdc.gov/phlp/publications/topic/hipaa.html#two>> [Accessed 27 June 2022].

Chin, C., 2022. *Telemedicine – Overview of the law, medico-legal issues and case studies*. Raja, Darryl and Loh. Kuala Lumpur.

Chiusano, F., 2022. *Two minutes NLP — Learn TF-IDF with easy examples | by Fabio Chiusano | NLPlanet | Medium*. [online] NLPlanet. Available at: <<https://medium.com/nlplanet/two-minutes-nlp-learn-tf-idf-with-easy-examples-7c15957b4cb3>> [Accessed 13 April 2023].

Cint, 2021. *Survey Sampling Methods & Techniques | Lucid — Cint™*. [online] Cint. Available at: <<https://www.cint.com/blog/different-methods-of-survey-sampling>> [Accessed 15 August 2022].

Combs, V., 2021. *Telemedicine: Doctors and patients are both worried about privacy and data security | TechRepublic*. [online] TechRepublic. Available at: <<https://www.techrepublic.com/article/telemedicine-doctors-and-patients-are-both-worried-about-privacy-and-data-security/>> [Accessed 28 June 2022].

CompareHero, 2022. *T20, M40 And B40 Income Classifications in Malaysia*. [online] Available at: <<https://www.comparehero.my/budgets-tax/articles/t20-m40-b40-malaysia>> [Accessed 21 July 2022].

David, A., 2022. Malaysia's path to becoming ageing nation fast even by global standards. *New Straits Times*. [online] 4 Jan. Available at: <<https://www.nst.com.my/news/nation/2022/01/760251/malaysias-path-becoming-ageing-nation-fast-even-global-standards>> [Accessed 20 July 2022].

Department of Statistics Malaysia, 2021. *Statistics on Causes of Death, Malaysia, 2021*. [online] Available at: <[https://www.dosm.gov.my/v1/index.php?r=column/cthemByCat&cat=401&bul\\_id=R3VrRUhwSXZDN2k4SGN6akRhTStwQT09&menu\\_id=L0pheU43NWJwRWVVSZklWdzQ4TlhUUT09](https://www.dosm.gov.my/v1/index.php?r=column/cthemByCat&cat=401&bul_id=R3VrRUhwSXZDN2k4SGN6akRhTStwQT09&menu_id=L0pheU43NWJwRWVVSZklWdzQ4TlhUUT09)> [Accessed 21 July 2022].

Deziel, C., 2018. *The Effects of a Small Sample Size Limitation | Sciencing*. [online] Sciencing. Available at: <<https://sciencing.com/effects-small-sample-size-limitation-8545371.html>> [Accessed 16 April 2023].

Donges, N., 2023. *Types of Data in Statistics: A Guide | Built In*. [online] BUILTIN. Available at: <<https://builtin.com/data-science/data-types-statistics>> [Accessed 6 April 2023].

ECRI, 2021. *Videoconferencing Systems, Telemedicine*. [online] Plymouth Meeting. Available at: <<http://www.ecri.org>> [Accessed 28 June 2022].

Fadzil, N.H.M., Shahar, S., Rajikan, R., Singh, D.K.A., Ludin, A.F.M., Subramaniam, P., Ibrahim, N., Vanoh, D. and Ali, N.M., 2022. A Scoping Review for Usage of Telerehabilitation among Older Adults with Mild Cognitive Impairment or Cognitive Frailty. *International Journal of Environmental Research and Public Health* 2022, Vol. 19, Page 4000, [online] 19(7), p.4000. <https://doi.org/10.3390/IJERPH19074000>.

Friedman, A.B., Pathmanabhan, C., Glicksman, A., Demiris, G., Cappola, A.R. and McCoy, M.S., 2022. Addressing Online Health Privacy Risks for Older Adults: A Perspective on Ethical Considerations and Recommendations: *Sage Journal*, [online] 8, p.233372142210957. <https://doi.org/10.1177/23337214221095705>.

Grande, D., Mitra, N., Iyengar, R., Merchant, R.M., Asch, D.A., Sharma, M. and Cannuscio, C.C., 2022. Consumer Willingness to Share Personal Digital Information for Health-Related Uses. *Journal of the American Medical Association*, [online] 5(1), pp.e2144787–e2144787. <https://doi.org/10.1001/JAMANETWORKOPEN.2021.44787>.

Greer, B., 2018. *Sklearn LDA vs. GenSim LDA. One of my favorite, and most... | by Ben Greer | Medium*. [online] Available at: <<https://medium.com/@benzgreer/sklearn-lda-vs-gensim-lda-691a9f2e9ab7>> [Accessed 13 April 2023].

HABIBU, S., 2019. Health Ministry investigating leak of patient records | The Star. *TheStar*. [online] 19 Sep. Available at: <<https://www.thestar.com.my/news/nation/2019/09/19/health-ministry-investigating-leak-of-patient-records>> [Accessed 26 June 2022].

Hall, J.L. and McGraw, D., 2014. For telehealth to succeed, privacy and security risks must be identified and addressed. *Health Affairs*, [online] 33(2), pp.216–221. <https://doi.org/10.1377/HLTHAFF.2013.0997>.

Haque, M.M.A., Jahan, Y., Khair, Z., Moriyama, M., Rahman, M.M., Sarker,

M.H.R., Shaima, S.N., Chowdhury, S., Matin, K.F., Karim, I.J., Ahmed, M.T., Hossain, S.Z., Masud, M.A.H., Nabi, M.G., Aziz, A.B., Sharif, M., Chowdhury, M.F.I., Shams, K.L., Nizam, N.B., Ananta, T.T., Amin, M.R. and Hawlader, M.D.H., 2022. Perceptions about Telemedicine among Populations with Chronic Diseases amid COVID-19: Data from a Cross-Sectional Survey. *International Journal of Environmental Research and Public Health*, [online] 19(7), p.4250. <https://doi.org/10.3390/IJERPH19074250>.

Indeed, 2021. *19 Types of Industry Sectors*. [online] Indeed. Available at: <https://www.indeed.com/career-advice/finding-a-job/types-of-industry> [Accessed 12 July 2022].

Intan Sabrina, M. and Defi, I.R., 2020. Telemedicine Guidelines in South East Asia—A Scoping Review. *Frontiers in Neurology*, [online] 11, p.581649. <https://doi.org/10.3389/FNEUR.2020.581649>.

Jafni, T.I., Bahari, M., Ismail, W. and Hanafi, M.H., 2019. Exploring barriers that affect telerehabilitation readiness: A case study of rehabilitation centre in Malaysia. *Advances in Intelligent Systems and Computing*, 843, pp.761–771. [https://doi.org/10.1007/978-3-319-99007-1\\_70](https://doi.org/10.1007/978-3-319-99007-1_70).

Jansen-Kosterink, S., Dekker-van Weering, M. and van Velsen, L., 2019. Patient acceptance of a telemedicine service for rehabilitation care: A focus group study. *International Journal of Medical Informatics*, 125, pp.22–29. <https://doi.org/10.1016/J.IJMEDINF.2019.01.011>.

Karampela, M., Ouhbi, S. and Isomursu, M., 2019. Connected Health User Willingness to Share Personal Health Data: Questionnaire Study. *Journal Medical Internet Research* 2019;21(11):e14537 <https://www.jmir.org/2019/11/e14537>, [online] 21(11), p.e14537. <https://doi.org/10.2196/14537>.

Karin, K., 2023. *What is Data Analysis? Process, Types, Methods and Techniques*. [online] Simplilearn. Available at: <https://www.simplilearn.com/data-analysis-methods-process-types-article> [Accessed 6 April 2023].

Kaspersky, 2021. *Telehealth take-up: the risks and opportunities*. Moscow.

Kato-Lin, Y.-C. and Thelen, S.T., 2022. Privacy Concerns and Continued Use Intention of Telemedicine During COVID-19. *Telemedicine Journal and E-health : The Official Journal of The American Telemedicine Association*. [online] <https://doi.org/10.1089/TMJ.2021.0603>.

Kim, H.-Y., 2017. Statistical notes for clinical researchers: Chi-squared test and Fisher's exact test. *Restorative Dentistry & Endodontics*, [online] 42(2), p.152. <https://doi.org/10.5395/RDE.2017.42.2.152>.

Kim, T.K. and Choi, M., 2019. Older Adults' Willingness to Share their personal and health information when adopting healthcare technology and services. *International Journal of Medical Informatics*, [online] 126, pp.86–94.

<https://doi.org/10.1016/J.IJMEDINF.2019.03.010>.

Lam, K., Lu, A.D., Shi, Y. and Covinsky, K.E., 2020. Assessing Telemedicine Unreadiness Among Older Adults in the United States During the COVID-19 Pandemic. *Journal of the American Medical Association Internal Medicine*, [online] 180(10), pp.1389–1391. <https://doi.org/10.1001/JAMAINTERNMED.2020.2671>.

Lee, J.Y., Wong, C.P. and Lee, S.W.H., 2020. m-Health views and perception among Malaysian: findings from a survey among individuals living in Selangor. *mHealth*, 6, pp.6–6. <https://doi.org/10.21037/MHEALTH.2019.09.16>.

Li, S., 2018. *Topic Modelling in Python with NLTK and Gensim | by Susan Li | Towards Data Science*. [online] Towards Data Science. Available at: <<https://towardsdatascience.com/topic-modelling-in-python-with-nltk-and-gensim-4ef03213cd21>> [Accessed 13 April 2023].

Lin, Z., 2017. *The Overall Perception of Telemedicine and Intention to Use Telemedicine Services: A comparison between frequent travelers and non frequent travelers*. [online] Available at: <<https://ecommons.cornell.edu/handle/1813/56813>> [Accessed 15 July 2022].

Majeed, A., 2019. Attribute-centric anonymization scheme for improving user privacy and utility of publishing e-health data. *Journal of King Saud University - Computer and Information Sciences*, 31(4), pp.426–435. <https://doi.org/10.1016/J.JKSUCI.2018.03.014>.

Matthew, G., 2021. *The 9 best free form builders and survey tools | Zapier*. [online] Available at: <<https://zapier.com/blog/best-free-survey-tool-form-app/>> [Accessed 9 August 2022].

McCoy, M.S., Joffe, S. and Emanuel, E.J., 2020. Sharing Patient Data Without Exploiting Patients. *Journal of the American Medical Association*, [online] 323(6), pp.505–506. <https://doi.org/10.1001/JAMA.2019.22354>.

Ministry of Health Malaysia, 1997. *Malaysia's Telemedicine Blueprint Leading Healthcare Into The Information Age*. Kuala Lumpur.

Ng, S.W., Hwong, W.Y., Husin, M., Rahman, N.A., Nasir, N.H., Juval, K. and Sivasampu, S., 2022. Assessing the Availability of Teleconsultation and the Extent of Its Use in Malaysian Public Primary Care Clinics: Cross-sectional Study. *Journal of Medical Internet Research*, [online] 6(5), p.e34485. <https://doi.org/10.2196/34485>.

O'Brien, E.C., Rodriguez, A.M., Kum, H.C., Schanberg, L.E., Fitz-Randolph, M., O'Brien, S.M. and Setoguchi, S., 2019. Patient perspectives on the linkage of health data for research: Insights from an online patient community questionnaire. *International Journal of Medical Informatics*, 127, pp.9–17. <https://doi.org/10.1016/J.IJMEDINF.2019.04.003>.

Office of the Australian Information Commissioner, 2022. *What is a privacy*

*policy?* [online] Australian Government. Available at: <<https://www.oaic.gov.au/privacy/your-privacy-rights/what-is-a-privacy-policy>> [Accessed 9 August 2022].

Ozeran, L., Solomonides, A. and Schreiber, R., 2021. Privacy versus Convenience: A Historical Perspective, Analysis of Risks, and an Informatics Call to Action. *Applied Clinical Informatics*, [online] 12(2), p.274. <https://doi.org/10.1055/S-0041-1727197>.

Darus, K.M., 2017. *Personal Data Protection Act 2010*. Government of Malaysia.

Pharmaceutical Services Division Ministry of Health Malaysia, 2016. *GUIDE TO GOOD DISPENSING PRACTICE*. Available at: <<https://www.pharmacy.gov.my/v2/sites/default/files/document-upload/gdsp-2016-final.pdf>> [Accessed 9 August 2022].

QuestionPro, 2022. *Close Ended Questions: Definition, Types with Examples / QuestionPro*. [online] QuestionPro. Available at: <[https://www.questionpro.com/close-ended-questions.html#Types\\_of\\_Close\\_Ended\\_Questions](https://www.questionpro.com/close-ended-questions.html#Types_of_Close_Ended_Questions)> [Accessed 10 August 2022].

Raja Eileen Soraya, R.A., Lai Ling, T., Teh, J., Adrienne Chin Raja, C., Malaysia, L., Eileen Soraya Raja Aman, R. and Adrienne Chin, C., 2021. *Definitive global law guides offering comparative analysis from top-ranked lawyers*. Kuala Lumpur .

Rouidi, M., Elouadi, A. and Hamdoune, A., 2022. Acceptance and use of telemedicine technology by health professionals: Development of a conceptual model: *Digital Health*. [online] <https://doi.org/10.1177/20552076221081693>.

Salleh, M.I.M., Abdullah, R. and Zakaria, N., 2021. Evaluating the effects of electronic health records system adoption on the performance of Malaysian health care providers. *BMC Medical Informatics and Decision Making*, 21(1), pp.5–11. <https://doi.org/10.1186/s12911-021-01447-4>.

Seltzer, E., Goldshear, J., Guntuku, S.C., Grande, D., Asch, D.A., Klinger, E. V. and Merchant, R.M., 2019. Patients' willingness to share digital health and non-health data for research: A cross-sectional study. *BMC Medical Informatics and Decision Making*, [online] 19(1), pp.1–8. <https://doi.org/10.1186/S12911-019-0886-9/TABLES/4>.

Steger, A., 2019. *Healthcare Data Breach: What Happens to Stolen Healthcare Data?* [online] HealthTech. Available at: <<https://healthtechmagazine.net/article/2019/10/what-happens-stolen-healthcare-data-perfcon>> [Accessed 8 July 2022].

Tan, Y.R., Tan, M.P., Khor, M.M., Hoh, H.B., Saedon, N., Hasmukharay, K., Tan, K.M., Chin, A.V., Kamaruzzaman, S.B., Ong, T., Davey, G. and Khor, H.M., 2022. Acceptance of virtual consultations among older adults and

caregivers in Malaysia: a pilot study during the COVID-19 pandemic. *Taylor & Francis Online*, [online] 134(2), pp.224–229. <https://doi.org/10.1080/00325481.2021.2004792>.

Taylor, D., 2023. *Data Mining Tutorial: What is Data Mining? Techniques, Process*. [online] Guru99. Available at: <<https://www.guru99.com/data-mining-tutorial.html>> [Accessed 7 April 2023].

Thong, H.K., Wong, D.K.C., Gendeh, H.S., Saim, L., Athar, P.P.B.S.H. and Saim, A., 2021. Perception of telemedicine among medical practitioners in Malaysia during COVID-19. *Journal of Medicine and Life*, [online] 14(4), p.468. <https://doi.org/10.25122/JML-2020-0119>.

Woo, K. and Dowding, D., 2018. Factors Affecting the Acceptance of Telehealth Services by Heart Failure Patients: An Integrative Review. *Telemedicine Journal and e-Health*, [online] 24(4), p.292. <https://doi.org/10.1089/TMJ.2017.0080>.

Yze, D.L., 2021. *Text Normalization for Natural Language Processing (NLP) / by Diego Lopez Yse / Towards Data Science*. [online] Towards Data Science. Available at: <<https://towardsdatascience.com/text-normalization-for-natural-language-processing-nlp-70a314bfa646>> [Accessed 13 April 2023].

## APPENDICES

### Appendix A: UTAR's Ethical Approval for Research Project



**UNIVERSITI TUNKU ABDUL RAHMAN** DU012(A)

Wholly owned by UTAR Education Foundation Co. No. 578227-M

Re: U/SERC/180/2022

29 August 2022

Dr Kwan Ban Hoe  
Head, Department of Mechatronics and BioMedical Engineering  
Lee Kong Chian Faculty of Engineering and Science  
Universiti Tunku Abdul Rahman  
Jalan Sungai Long  
Bandar Sungai Long  
43000 Kajang, Selangor

Dear Dr Kwan,

#### Ethical Approval For Research Project/Protocol

We refer to your application for ethical approval for your student's research project from Bachelor of Engineering (Hons) Biomedical Engineering programme enrolled in course UEGE4118. We are pleased to inform you that the application has been approved under Expedited Review.

The details of the research projects are as follows:

No.	Research Title	Student's Name	Supervisor's Name	Approval Validity
1.	A Study on Patient Data Sharing Acceptability Among Malaysian with Data Mining Techniques	Lee Xin Ying	Dr Goh Choon Hian	29 August 2022 – 28 August 2023

The conduct of this research is subject to the following:

- (1) The participants' informed consent be obtained prior to the commencement of the research;
- (2) Confidentiality of participants' personal data must be maintained; and
- (3) Compliance with procedures set out in related policies of UTAR such as the UTAR Research Ethics and Code of Conduct, Code of Practice for Research Involving Humans and other related policies/guidelines.
- (4) Written consent be obtained from the institution(s)/company(ies) in which the physical or/and online survey will be carried out, prior to the commencement of the research.

**Kampar Campus** : Jalan Universiti, Bandar Barat, 31900 Kampar, Perak Darul Ridzuan, Malaysia

Tel: (605) 468 8888 Fax: (605) 466 1313

**Sungai Long Campus** : Jalan Sungai Long, Bandar Sungai Long, Cheras, 43000 Kajang, Selangor Darul Ehsan, Malaysia

Tel: (603) 9086 0288 Fax: (603) 9019 8868

**Website**: www.utar.edu.my





Should the students collect personal data of participants in their studies, please have the participants sign the attached Personal Data Protection Statement for records.

Thank you.

Yours sincerely,



**Professor Ts Dr Faiz bin Abd Rahman**  
Chairman  
UTAR Scientific and Ethical Review Committee

c.c    Dean, Lee Kong Chian Faculty of Engineering and Science  
        Director, Institute of Postgraduate Studies and Research



