



# MONASH University

## **Health Information Seeking and Online Health Information Searching Behaviours of University Students in Developing Countries**

Tesfahun Melese Yilma

A thesis submitted in fulfilment of the requirements for the degree of

*Doctor of Philosophy*

Monash University

Faculty of Information Technology

July 2019

## **Copyright Notice**

© Tesfahun Melese Yilma (2019). I certify that I have made all reasonable efforts to secure copyright permissions for third-party content included in this thesis and have not knowingly added copyright content to my work without the owner's permission.

## ABSTRACT

Knowledge and skill gaps exist between people from developing and developed countries in locating and using health information. These gaps can be minimized through improved Health Information Seeking Behaviour (HISB) and Online Health Information Searching Behaviour (OHISB) which have the potential to reduce knowledge gaps across social groups and to educate individuals. This research study aims to examine the HISB and its associated factors among university students from developing countries. It also aims to explore the OHISB of university students and their challenges in online environments. University students are selected because they are often presumed to be a neglected target group for health promotion activities. Moreover, previous research studies show high risks of chronic diseases, stress, and risky sexual behaviour among university students. Improving their HISB and OHISB has the potential to reduce these risks.

A quantitative cross-sectional study is employed in order to examine the HISB of university students. In addition, an interactive information retrieval experiment type methodology is used to explore university students' OHISB. Two sample groups, one for the survey and another for the experiment, are considered. A total of 901 university students (354 from Malaysia and 547 from Ethiopia) participated in the survey. For the experiment, a total of 176 university students from Malaysia who already involved in the survey have participated. Two data collection methods, a survey and an experiment, are used in this study. In the survey, a web-based survey is used as a data collection tool to gather information on HISB and its associated factors. The experiment involves users using a computer to search on three simulated and one personal tasks.

Results reveal that a substantial number of university students from developing countries have sought health information mostly from the Internet. The percentages of health information seeking from middle and low-income countries are found to be 77.1% and 61.8%, respectively. In this research study, the Internet is identified as the primary source of health information in a middle-income country. In addition, television or radio is identified as the main source of health information in a low-income country. The results could help health promoters to pass their health messages via the source that university students prefer. The factors associated with health information seeking are identified to

be health literacy, perceived susceptibility to health problems, and alcohol consumption. Results of the experimental study indicate question-based queries as a popular type of query among students when searching for health information online. Including a question-answering technique in health information retrieval systems could help users to express their information need in natural language questions and retrieve specific answers for their questions. Furthermore, specification is found to be popular query reformulation pattern which could indicate a lack of understanding the level of specificity of search tasks. Hence, providing query refining feature to health information retrieval systems could help users to refine their search and retrieve specific answers for their information needs. Overall, this research study provides an understanding of HISB and OHISB of university students in developing countries.

## **Declaration**

This thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

Name: Tesfahun Melese Yilma

Date: 16<sup>th</sup> July 2019

## LIST OF PUBLICATIONS

1. Yilma, T.M., Inthiran, A., Reidpath, D.D. & Orimaye, S.O. (2019). *Context-based interactive health information searching*. Information Research, 24(2), paper 815. Retrieved from <http://InformationR.net/ir/24-2/paper815.html> (Archived by WebCite® at <http://www.webcitation.org/78mjB8yl7>).
2. Yilma, T. M., Inthiran, A., Reidpath, D., & Orimaye, S. O. (2017). *Health Information Seeking and its Associated Factors among University Students: A Case in a Middle-Income Setting*. Paper presented at the 21st Pacific-Asia Conference on Information Systems, Langkawi, Malaysia.
3. Yilma, T. M., Inthiran, A., & Reidpath, D. (2016). *College Students from Developing Countries: Where do they Get Health Information?* Paper presented at the MEDIR 2016, SIGIR 2016, Pisa, Tuscany, Italy.
4. Yilma, T. M. (2016). *Health Information Seeking Behavior among College Students: A case in a Developing Country*. Paper presented at the Proceedings of the 2016 ACM on Conference on Human Information Interaction and Retrieval, Carrboro, North Carolina, USA.

## **ACKNOWLEDGEMENTS**

First and foremost, I would like to extend my infinite heartfelt gratitude to my main supervisor, Dr Anushia Inthiran, for helping me begin this journey, an amazing journey that has taken me on a full ride to where I am today. Thank you, Dr Anushia, for the opportunity you gave me, and for being encouraging and inspiring throughout this research work. No distance lessened your continuous advice and support, and I am deeply grateful for this. This work would not have been possible without your wise guidance and constant support, thank you so much, Dr Anushia.

I would also like to express my deepest appreciation to my associate supervisor Prof. Daniel Reidpath for your valuable guidance and insightful comments. Your incredible help and support during my difficult situations and throughout the past four and half years are sincerely appreciated. A heartfelt thanks to Dr Sylvester Orimaye for your constructive feedback and encouragement since you ever became my external supervisor during my second year. My deepest appreciation also goes to Associate Prof. Wong Kok Sheik for giving so much care and support, especially during data collection time. Ever since you became my supervisor during my third year, you have been kind and supportive to me, and I am grateful for that.

This research would not have been possible without the Monash University Malaysia Merit Scholarship that I received for three and a half years, from December 2014 to May 2018. I also would like to express my gratitude to the academic and administrative staff members of the School of Information Technology for their encouragement and support during various parts of my candidature. Special thanks to Ms Wan Nurul Rukiah for her understanding and kind support, especially during my data collection period.

Multitude of thanks and appreciation to my friends, Kokum, Catherine, Vinesha, Najini, Kuan Yew, Yakub, Kasun, Somayeh, Judy, Alena and other colleagues at the School of Information Technology for always being there, keeping me socially alive, creating an interactive and enjoyable work environment in the lab, and providing an unforgettable and enjoyable life outside the work environment. My journey of pursuing a PhD would be less rewarding without your presence.

I extend a heartfelt thank you to my study participants from Monash University Malaysia, University of Malaya, Swinburne University of Technology, Sarawak Campus and the University of Gondar, Ethiopia.

Last but not least, I give my dearest thanks to my family for their continuous support and endurance throughout my research work.



# TABLE OF CONTENTS

<b>ABSTRACT</b> .....	I
<b>LIST OF PUBLICATIONS</b> .....	IV
<b>ACKNOWLEDGEMENTS</b> .....	V
<b>LIST OF TABLES</b> .....	X
<b>LIST OF FIGURES</b> .....	XI
<b>CHAPTER 1: INTRODUCTION</b> .....	1
1.1. Health Information Seeking Behaviour .....	1
1.2. Online Health Information Searching Behaviour.....	3
1.3. Research Motivation .....	3
1.4. Research Objectives and Questions .....	9
1.5. Research Contributions .....	10
<b>CHAPTER 2: LITERATURE REVIEW</b> .....	11
2.1. Health Information Seeking Behaviour .....	12
2.1.1. Source of Health Information .....	13
2.1.2. Type of Health Information .....	21
2.2. Factors Associated with Health Information Seeking Behaviour .....	24
2.2.1. Socio-demographical Factors .....	25
2.2.2. Psychological Factors .....	28
2.2.3. Behavioural Factors .....	30
2.2.4. Health Literacy and Health Status .....	30
2.3. Online Health Information Searching Behaviour.....	35
2.3.1. Challenges of Health Information Searching.....	36
2.3.2. Health Information Searching Behaviour .....	37
2.4. Theoretical Framework.....	44
2.4.1. Wilson's 1999 Model of Information Seeking and Information Searching Research Areas .....	44
2.4.2. Longo's Expanded Model of Health Information Seeking Behaviour .....	45
2.4.3. Health Belief Model.....	48
2.4.4. Sutcliffe and Ennis' Process Model of Information Searching .....	50
2.4.5. Proposed Conceptual Framework .....	52
2.5. Chapter Summary .....	54

<b>CHAPTER 3: RESEARCH METHODOLOGY .....</b>	<b>56</b>
<b>3.1. Research Methods in Information Seeking Behaviour Studies.....</b>	<b>56</b>
3.1.1. Survey .....	57
3.1.2. Information Retrieval Experimental Methodologies .....	58
<b>3.2. Research Design and Setting .....</b>	<b>61</b>
<b>3.3. Participants.....</b>	<b>62</b>
<b>3.4. Sampling Technique.....</b>	<b>63</b>
<b>3.5. Data Collection Tools and Techniques .....</b>	<b>64</b>
3.5.1. Web-based Survey .....	64
3.5.2. Experiment.....	66
<b>3.6. Data Gathering Procedure .....</b>	<b>72</b>
<b>3.7. Data Management and Analysis.....</b>	<b>73</b>
<b>3.8. Validity and Reliability .....</b>	<b>75</b>
<b>3.9. Chapter Summary .....</b>	<b>76</b>
<b>CHAPTER 4: RESULTS .....</b>	<b>78</b>
<b>4.1. Demographic Information of Participants .....</b>	<b>78</b>
<b>4.2. Health Information Seeking Behaviour .....</b>	<b>80</b>
4.2.1. Source of Health Information.....	81
4.2.2. Type of Health Information .....	84
<b>4.3. Factors Associated with Health Information Seeking.....</b>	<b>85</b>
4.3.1. Factors Associated with Source Of Health Information .....	87
4.3.2. Factors Associated with Frequency of Health Information Seeking.....	90
<b>4.4. Online Health Information Searching Behaviour.....</b>	<b>92</b>
4.4.1. Demographic Information .....	92
4.4.2. Query Behaviour .....	94
4.4.3. Query Reformulation.....	99
4.4.4. Result Viewing Behaviour.....	101
<b>4.5. Challenges of Health Information Searching.....</b>	<b>102</b>
<b>4.6. Chapter Summary .....</b>	<b>103</b>
<b>CHAPTER 5: DISCUSSION AND RESEARCH CONTRIBUTION.....</b>	<b>106</b>
<b>5.1. Health Information Seeking Behaviour .....</b>	<b>106</b>
5.1.1. Source of Health Information.....	107
5.1.2. Type of Health Information .....	108

<b>5.2. Factors Associated with Health Information Seeking</b>	<b>109</b>
5.2.1. Factors Associated with the Source of Health Information	110
5.2.2. Factors Associated with the Frequency of Health Information Seeking	112
<b>5.3. Online Health Information Searching</b>	<b>114</b>
5.3.1. Querying Formulation	114
5.3.2. Contextual Factors of Query Formulation	116
5.3.3. Query Reformulation	118
5.3.4. Search Domain	118
<b>5.4. Challenges of Health Information Searching</b>	<b>119</b>
<b>5.5. Research Contribution</b>	<b>120</b>
5.5.1. Theoretical Contribution	120
5.5.2. Practical Contribution	126
<b>5.6. Chapter Summary</b>	<b>128</b>
<b>CHAPTER 6: CONCLUSION</b>	<b>131</b>
<b>REFERENCE</b>	<b>135</b>
<b>APPENDIX</b>	<b>148</b>
Appendix A: Operational Definitions	148
Appendix B: Web-based Survey on Health Information Seeking Behaviour	151
Appendix C: Simulated Situation	165
Appendix D: Observational Checklist	167
Appendix E: Post-experiment Survey	168
Appendix F: Explanatory Statement	171
Appendix G: Consent Form	174
Appendix H: Instruction Letter	175

## LIST OF TABLES

Table 1: Summary of Research Articles on Source of Health Information.....	20
Table 2: Summary of Research Articles on the Purpose of Health Information Searching .....	23
Table 3: Summary of Research Articles on Factors Associated with Health Information Seeking Behaviour .....	34
Table 4: Summary of articles on online health information searching behaviour.....	43
Table 5: Health information needs of 15 participants from the pilot study .....	68
Table 6: Summary of the Research Methodology .....	77
Table 7: Socio-demographic characteristics of participants from Malaysia and Ethiopia .....	79
Table 8: Multivariate logistic regression showing factors affecting health information seeking among participants located in Malaysia .....	86
Table 9: Multivariate logistic regression showing factors affecting the source of health information amongst participants from Malaysia and Ethiopia .....	90
Table 10: Poisson regression analysis showing the factors affecting the frequency of health information seeking amongst participants from Malaysia and Ethiopia .....	92
Table 11: Socio-demographic characteristics of participants from Malaysia .....	93
Table 12: Basic query features across tasks.....	96
Table 13: Number of queries and query length .....	97
Table 14: Poisson and negative binomial regression analysis showing the factors associated with the number of queries and query length across all the tasks .....	99
Table 15: Query reformulation pattern.....	101
Table 16: Top ten most websites visited by participants .....	102
Table 17: Summary results of health information seeking and searching behaviours .	104

## LIST OF FIGURES

Figure 1: Internet users (per 100 people) in Southeast Asia, data source: (The World Bank, 2015b).....	7
Figure 2: Wilson's 1999 nested model of information seeking and information searching. Source: (Wilson, 1999).....	45
Figure 3: Longo Health Information Model: information seeking, passive receipt, and use. Source: (Longo et al., 2010) .....	47
Figure 4: The Health Belief Model as a predictor of preventive health behaviour. Source: (Rosenstock, 1974) .....	50
Figure 5: The process model of information searching activities and knowledge sources. Source: (Sutcliffe & Ennis, 1998) .....	52
Figure 6: Conceptual Framework of Health Information Seeking Behaviour. A) Searcher ends up with a satisfactory result or unsatisfactory with a decision to abandon searching. B) Searcher reformulate a query in case of unsatisfactory results.....	53
Figure 7: Diagrammatic representation of data collection methods and procedure.....	73
Figure 8: Schematic diagram of the data analysis procedure and methods .....	75
Figure 9: Participants' reasons for not seeking health information .....	80
Figure 10: Primary source of health information amongst participants from Malaysia and Ethiopia .....	81
Figure 11: Primary online searching source of health information among participants from Malaysia and Ethiopia .....	82
Figure 12: Use of social media to look for health information amongst participants located in Malaysia (left) and Ethiopia (right) .....	83
Figure 13: Participants' activities on social media .....	84
Figure 14: Type of health information sought among participants from Malaysia and Ethiopia .....	85
Figure 15: Challenges faced by participants when performing searching tasks .....	103

## **CHAPTER 1: INTRODUCTION**

This research study provides information about the health information seeking behaviour and online health information searching behaviour of university students from two developing countries. In this chapter, a brief introduction to health information seeking behaviour is given. Then, the characteristics of health information seeking behaviours, such as the source of health information, the type of health information, and the actions used to seek health information are described. Next, information about online health information searching behaviour is given. Following the online health information searching behaviour, the motivations behind this research study are discussed followed by a presentation of the objectives and research questions. Finally, a brief summary of the contributions of this research study is described.

### **1.1. Health Information Seeking Behaviour**

Health Information Seeking Behaviour (HISB) is defined as the ways by which people obtain information about their health, disease, health risk, and health promotion (Lambert & Loiselle, 2007). It is an essential strategy in health promotion campaigns, and psychological and social adjustment to illnesses. Health information seeking provides a number of benefits, such as increasing health knowledge (Medlock et al., 2013; Weaver et al., 2010; Willis & Royne, 2017), improving ability to comfortably discuss health issues with health professionals in order to have enhanced shared decision making (Feinberg et al., 2015; Jung, 2014), improving the ability to cope with stresses, increasing the ability for self-care management skills and commitment to treatments (Jung, 2014; Willis & Royne, 2017), and medical treatment satisfaction (Jung, 2014). As a result of the benefits, research studies identified HISB as a key component of health communication for health outcomes (European Commission, 2014; Medlock et al., 2015; Wong et al., 2014).

HISB is characterized by the sources of health information, the type of health information, and the actions used to search health information (Lalazaryan & Zare-Farashbandi, 2014; Lambert & Loiselle, 2007; National Cancer Institute, 2014). The Internet, people (healthcare professionals, family, and friends), and traditional media (television, radio,

newspaper, magazine), are considered as the sources of health information (Lalazaryan & Zare-Farashbandi, 2014). Most notably, using the Internet to look for health information has increased because of easy access to enormous information (Bratucu et al., 2014; Cutilli, 2010; Osei Asibey et al., 2017). For example, 80% of Internet users in the United States of America have looked for health information online (Fox, 2011b), six out of ten Europeans go online to look for health information (European Commission, 2014), and 67.7% of university students in the Ghana looked for health information (Osei Asibey et al., 2017).

In terms of the type of health information, the most common topics that the general population look for are about chronic diseases, fitness, nutrition and diet, and mental health information (Fox, 2011b). Health professionals go online to look for general medical and clinical trial information, mostly from academic societies and medical journals (Noguchi et al., 2015). Parents and caregivers look for information about nutrition, fitness, and diseases prevention and control from the Internet, printed books, television, newspapers, and from their families and friends (Lwoga & Mosha, 2013). On the other hand, university students use the Internet to seek out information about nutrition, fitness, sexual/reproductive health, and specific disease (e.g. HIV/AIDS, malaria, sore throat, mental health) (Obasola & Agunbiade, 2016). All these types of health information can help health information seekers develop preventive health behaviour, deal with uncertainties and anxieties, and recover from medical surgery (Yuli et al., 2012). The type of health information can also enable information seekers to make decisions about treatment options and disease control (Longo et al., 2010).

With regard to the action used to obtain health information, strategies or techniques employed to search health information from either manual or computerized health information systems are considered (Lambert & Loiselle, 2007). Specifically, the actions used to obtain health information include asking or discussing health information with people (healthcare professionals, family, and friends), watching health programs on the television, listening health programs on the radio, reading health information from magazines and the newspaper, and using the Internet to search for health information.

## **1.2. Online Health Information Searching Behaviour**

Online Health Information Searching Behaviour (OHISB) is a subset of health information seeking behaviour (Wilson, 1999). It is concerned about the interaction of users with an information retrieval system (Wilson, 1999). OHISB includes users' querying, result viewing and browsing behaviours, search outcomes, and their searching strategy when users interact with an information retrieval system (Goeuriot et al., 2016; Sutcliffe & Ennis, 1998). Understanding users' OHISB can help the design of effective health information retrieval systems. Specifically, it can lead to better query input methods and recommendation techniques (Lopes & Ribeiro, 2015; Zhang et al., 2012) and better search results display mechanisms (Zhang, 2013; Zhang et al., 2012).

The field of online health information searching has grown as increasing numbers and varieties of health information systems are being developed for users. Despite the advancement of the field, online health information searching is not always efficient and effective as a result of the huge amount of information available and the unplanned nature of information searching by users. The search strategies users are using to obtain health information are often unsuccessful (Kim et al., 2011; Sharit et al., 2015) due to the inadequate specification of search tasks or queries and improper use of search features (Stommel & Wills, 2004). The ability of users at specifying their need, the type of users' information need, and the context in which users' queries are issued significantly affect the success of users' interaction with information retrieval systems and users' searching behaviour (Kelly, 2006; Lwoga & Mosha, 2013). Since some users are using the information they obtain to influence health decisions, it is increasingly important to identify strategies that aid consumer access to quality information to address their needs.

## **1.3. Research Motivation**

This research study is motivated by 1) the lack of both theoretical and empirical research studies on HISB in developing countries, 2) the low percentage of health information seeking in developing countries, 3) the effect of health information seeking in preventing risky behaviours of university students, and 4) the increased Internet penetration rate



globally and in the study areas which is an opportunity for youths to go online and seek health information. The next few paragraphs of this chapter explain the motivation behind this research study in detail.

### **1.3.1. Lack of Research Studies on HISB in Developing Countries**

Many research studies on HISB focus on the developed countries (Lalazaryan & Zare-Farashbandi, 2014; Li et al., 2015) where many healthcare services are available (The World Bank, 2013). Moreover, most people in developed nations presumably demonstrate an adequate level of literacy (Nutbeam, 2000) and awareness of online health information sources (Health Information and Quality Authority, 2011). Consequently, many people from developed countries are more active health information seekers and have more opportunities to make shared decisions with their physicians on health conditions and treatment options (Anker et al., 2011; Dickinson & Raynor, 2003). Conversely, in developing countries, there are limited literate human resources as well as infrastructures such as telecommunication, electricity, and transportation (The World Bank, 2013) to help people engaged actively in health information seeking.

Moreover, both theoretical and empirical research studies on HISB are lacking in developing countries (Hamzah et al., 2015). This could be because the field is underdeveloped in many developing countries. For example, a bibliometric study by Li et al. (2015) on PubMed<sup>1</sup> about publication trends of HISB research studies identifies a very small number of publications from developing countries. Li et al. (2015) find that the major research contribution on HISB is from developed countries in North America, Europe, and Australia. The very small number of publications in many developing countries indicates the needs for more research studies on HISB in developing countries. The same observation is also reported by Lalazaryan and Zare-Farashbandi (2014) which indicates the need for research studies on HISB in developing countries. Lalazaryan and Zare-Farashbandi (2014) find out that there is no model of HISB for people in developing countries considering their socioeconomic characteristics. Such research studies could

---

<sup>1</sup>PubMed (<http://www.ncbi.nlm.nih.gov/pubmed>) is a free search engine including more than 25 million citations from MEDLINE, life science journals, and online books.

give crucial information for health planning and decision making. Hence, it is necessary to give emphasis to this area in developing countries.

### **1.3.2. Low Percentage of Health Information Seeking in Developing Countries**

In developing countries, only limited research studies have been done, mostly among patients and the general population (Mohd-Nor et al., 2013; Yuli et al., 2012). These studies have identified the magnitude of HISB as low (e.g. 46% in Malaysia (Mohd-Nor et al., 2013) and 59% in China (Yuli et al., 2012)). The low magnitude of health information seeking could be attributed to the complex nature of health information seeking in developing countries. The possible reasons for the complexity could be because 1) health systems are not sufficient to reach out to the entire population (World Health Organization, 2015d); 2) people in developing countries have difficulties in accessing and understanding health information and technologies (Patel & Dowse, 2013; Raj et al., 2015); 3) people in developing countries often have limited online information searching skills (Raj et al., 2015); and 4) many people in developing nations lack proper health knowledge, influencing their ability to seek out and understand health information (Nasrabadi et al., 2015). Given the low magnitude of health information seeking among patients and the general population as well as the complexity of health information seeking, it is important to investigate how university students from developing countries seek health information. Understanding health information seeking behaviour as well as its challenges could help to promote health and improve the health status of university students from developing countries.

### **1.3.3. Effect of Health Information Seeking in Preventing Risky Behaviours of University Students**

Previous HISB studies focus on individual groups other than university students. Some research studies focus on parents and caregivers (Lwoga & Mosha, 2013) while other research studies focus on older adults (El-Attar et al., 2005; Eriksson-Backa et al., 2012; Medlock et al., 2012). There are also research studies among health managers (Edwards et al., 2013), rural community (Mohd-Nor et al., 2013), and patients (Cutilli, 2010; Patel & Dowse, 2013; Wong et al., 2014). On the other hand, limited research studies have been

done on HISB among university students. Li et al. (2015) suggest students as a potential target group for HISB research studies.

University students are recognized as a neglected target group for health promotion (Hunt & Eisenberg, 2010; Stewart-Brown et al., 2000) while they are at risk of chronic diseases due to sedentary behaviour (Deliens et al., 2015; Graham et al., 2014; Larouere & Winters, 2013; Oman Global School-based Student Health Survey, 2012) which increases the chance of getting a disease that may kill youths prematurely (Biswas et al., 2015). Stress is also prevalent among university students (Graham et al., 2014; Griffin, 2014) due to academic load, sleeping problem, and homesickness (Beiter et al., 2015; Graham et al., 2014; Jena et al., 2015; Larouere & Winters, 2013; Reang & Bhattacharjya, 2013). University students are also at risk of risky sexual practices that they could be exposed to sexually transmitted diseases, unintended pregnancy, and chronic pelvic pain and inflammatory diseases (Rubens et al., 2014; Turchik, 2012). Besides, the group is at high risk of using untrustworthy health information through social media networking due to lack of knowledge to judge health information (Banas, 2008). In order to prevent chronic diseases among young people, the World Health Organization has recommended a strategy to ensure that health information is widely available and easily understood (World Health Organization, 2003). Nevertheless, availability alone may not ensure that young people seek out and use health information. There is a need to precisely understand their HISB and examine its associated factors so that it can be used for health promotional activities. Health promotion professionals can also use the result of this research study to promote health and improve the health status of the university population.

#### **1.3.4. Increased Internet Penetration Rate: An Opportunity for Health Information Seeking**

In the past years, Internet penetration rate was poor in developing countries. However, now it has improved (The World Bank, 2018a). The World Bank data (Figure 1) shows that the average Internet penetration rate in Southeast Asia has improved from 14.8% in 2005 to 47.7% in 2017. The increased Internet penetration rate enables developing countries to have their own health information portals (e.g., [www.myhealth.gov.my](http://www.myhealth.gov.my) in Malaysia and [www.nhp.gov.in](http://www.nhp.gov.in) in India). At the same time, there is increased interest to

search for health information on the Internet due to the ease and prompt access (Cutilli, 2010). These create great opportunities for people to have the chance to keep and improve their health (Allen et al., 2013). The improved online access further provides an ecosystem for researchers to analyse and understand users' online health information searching. Moreover, young users under 35 years old are the most web users in Southeast Asia. For example, in Malaysia, Vietnam, and Thailand, users of age under 35 years spend more time online when compared to users aged over 35 years (Brown, 2013). Although the heaviest web users from Southeast Asia are young people, little is known about their health searching activities.

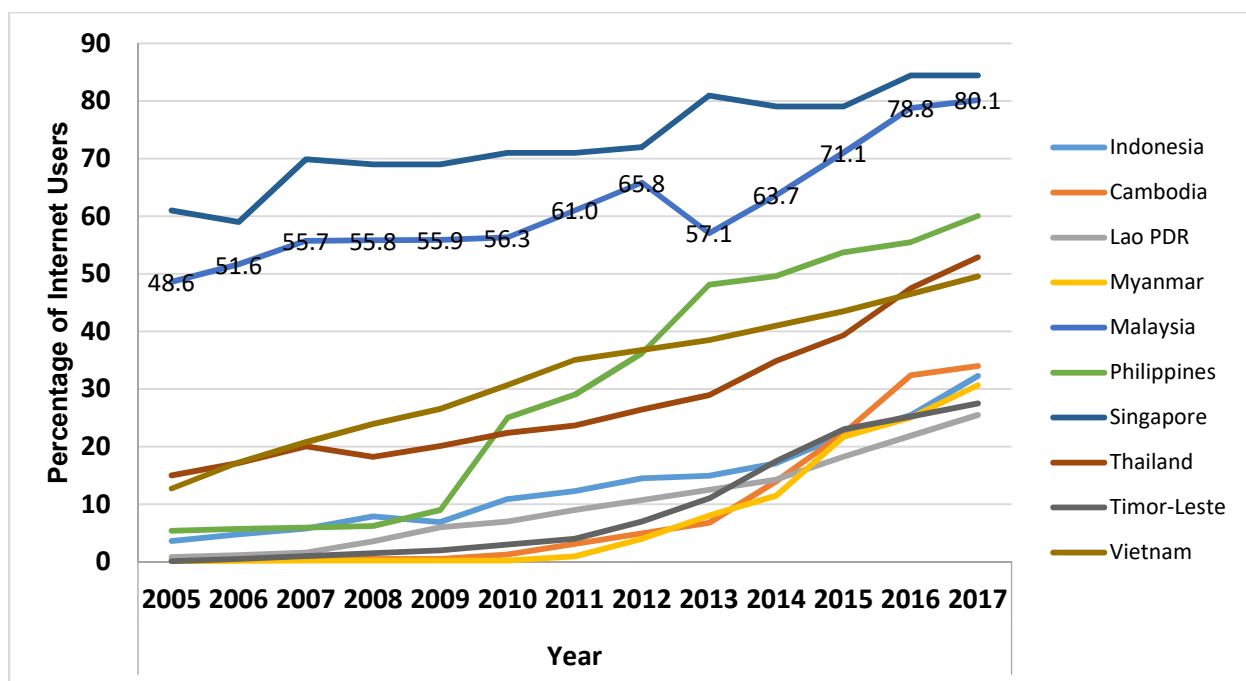


Figure 1: Internet users (per 100 people) in Southeast Asia, data source: (The World Bank, 2015b)

In Malaysia, the Internet penetration rate has improved from 21.4% in the year 2000 to 80.1% in the year 2017. Thirty-seven percent of Internet users in the country belongs to young people in the age group of 15-24 years. These young people contribute to more than a quarter of Malaysia population (Ministry of Health Malaysia, 2012), and they are the future productive group for the country. However, the health problem of this young population has become a significant issue for the country. According to the Malaysian

National Health Morbidity Report, diabetes, hypertension, and obesity are getting serious among these young population (Ministry of Health Malaysia, 2011). A research study pointed out that the reasons for Malaysia's youth attached to these diseases included low health literacy and lack of health consciousness which lead to unhealthy dietary and lifestyles (Norris et al., 2014). Hence, it is important to increase the health literacy and health awareness amongst young age group through improving their HISB. Therefore, to raise health awareness and make health information available online is essential to increase the health literacy of these group of users. Conducting research to study their HISB is critical for policymakers and health information promoters to ensure these young population encouraged and engaged in health information seeking.

In Ethiopia, although the overall Internet penetration rate is very low, university students are presumed to have access to the Internet on university campuses. These students use the Internet mostly for academic, communication, and entertainment purposes (Moly, 2014). Ethiopian youths also visit social networking sites frequently and spend considerable time to stay connected with their network friends (Hussain & Arasad, 2015). However, far too little is known about this young age group engagement in health information seeking. University students in Ethiopia are at high risk of risky sexual practices (Alamrew et al., 2013; Mengistu et al., 2013), depression (Dachew et al., 2015; Terasaki et al., 2009), and substance use (Gebreslassie et al., 2013; Tesfaye et al., 2014). Providing appropriate health information could help youths to get informed about their health matters so as to make informed decisions on their health and improve their health status.

Therefore, this research study focuses on examining health information seeking and searching behaviours of university students from developing countries. This research study attempts to create a better understanding of HISB and OHISB of university students thereby making it available for health promotional activities. It has also a potential contribution to suggest design strategies to improve health information retrieval systems.

## **1.4. Research Objectives and Questions**

This research study consists of two research objectives.

### ***Research Objective 1:***

The first research objective is to examine health information seeking behaviour of university students from developing countries. This research study attempts to describe the sources of health information used and the types of health information sought by university students from developing countries. The influence of socio-demographic, behavioural, psychological, and health-related factors on HISB is also identified.

The research questions associated with this research objective are:

1. What sources of health information do university students from developing countries use?
2. What type of health information do university students from developing countries search for?
3. Do socio-demographic, psychological, behavioural, and health-related factors affect the health information seeking behaviour of university students' from developing countries?

### ***Research Objective 2:***

The second research objective is to explore online health information searching behaviour and its challenges among university students from developing countries.

The research questions associated with this research objective are:

1. What searching behaviour do university students from developing countries exhibit when they search for health information online?
2. What challenges do university students from developing countries face while searching for health information online?

## 1.5. Research Contributions

In this section, a summary of the major contributions of this research study is highlighted.

- From the literature review of this research study, a comprehensive theoretical framework that could help to guide future research studies on HISB and OHISB is developed from four basic theoretical models, such as Wilsons' 1999 nested model of information seeking and information searching research areas, Longo's 2010 expanded model of health information seeking behaviour, the Health Belief Model, and Sutcliffe and Ennis' Process Model of Information Searching.
- The findings of this research study identify the magnitude of health information seeking among university students from middle-income and low-income countries as 77.1% and 61.8%, respectively. Regarding the source of health information, the Internet is identified as the primary source of health information among university students from a middle-income country whereas television or radio is found to be the main source of health information among university students from a low-income country. Moreover, the factors associated with health information seeking among university students from a middle-income country are identified as health literacy, perceived susceptibility to health problems, and alcohol consumption. The findings could be used for health promotion activities.
- Students commonly issue question-based queries to search engines when searching for health information. Including a question-answering technique in health information retrieval systems could help users to express their information need in natural language questions. Information extraction and natural language processing techniques could be employed to provide answers to questions posed by users.
- Results of query reformulation analysis reveal 'specification' as a popular query reformulation pattern. The popularity of 'specification' query reformulation pattern indicates a lack of defining search strategies and understanding the level of specificity of search tasks. Providing query refining features (e.g. refined by keywords and health topics) to information retrieval systems could help users to refine their search and retrieve specific answers for their information needs.

## CHAPTER 2: LITERATURE REVIEW

This chapter aims to review previous research findings on health information seeking and searching behaviours so as to have a thorough understanding of the research coverage and gaps in these areas. The literature review covers research studies from both developing and developed countries. However, since there are limited research studies done on health information seeking and searching behaviours in developing countries (Lalazaryan & Zare-Farashbandi, 2014; Li et al., 2015), voluminous information is provided mostly from developed countries. While the focus of this research study is on health information seeking and searching behaviours among university students from developing countries, research studies which are done among people other than university students are included. In addition, research studies which are done from developed countries are reviewed. This is because there are limited research studies done on the field among university students from developing countries (Li et al., 2015). The research studies from developed countries help to understand the theoretical background of the topic under study and guide to developing the conceptual framework of this research study. The studies are also helpful to adapt appropriate research methodologies to this research study.

In this chapter, the following key sections are covered: 1) Health information seeking behaviour (HISB), 2) Factors associated with HISB, 3) Online health information searching behaviour (OHISB), and 4) Theoretical framework. The chapter begins with a discussion on the concept of HISB. The concept is examined to have an understanding of how HISB is characterized. Understanding the concept helps to set the scope of this research and then develop the conceptual framework of this research study which is presented at the end of this chapter. Under the section HISB, the following subsections are covered: the sources of health information (online, physicians or health professionals, television or radio, family, and friends), the type of health information seeking, the use of social media for health information, and search engine preferences for health information searching. The subsections are discussed because they are key elements that makeup HISB (Lalazaryan & Zare-Farashbandi, 2014; Lambert & Loiselle, 2007; National Cancer Institute, 2014).



Next, the following factors associated with HISB are presented: the demographic, psychological, behavioural, and health-related factors. The purpose of reviewing the literature about the factors associated with HISB is to identify possible factors that may affect the HISB of university students from developing countries. Following the discussion on the factors associated with HISB, OHISB is described. Under the section OHISB, query formulation, query reformulation, result viewing behaviour, and the challenge of health information searching are discussed. Reviewing literature on OHISB helps to have an understanding of users' searching behaviour and challenges, and frame this research study. Finally, after reviewing four theoretical models of health information seeking and searching behaviours, a comprehensive conceptual framework that helps to guide this research study is developed.

## **2.1. Health Information Seeking Behaviour**

Commonly, the term HISB is used as “information seeking behaviour”, implying the term “health” as the type of information sought (Lambert & Loiselle, 2007). Some researchers conceptualize HISB as an intentional or purposive acquisition of health information from an information system (Case, 2012; Longo, 2005). Other researchers consider that it can also include passive or accidental exposure to information, acquiring and using the information (Eheman et al., 2009; Gavgani et al., 2013; Longo et al., 2009; Longo et al., 2010). Lambert and Loiselle (2007) review the literature on the concept of HISB in order to define and identify its essential characteristics. Based on the review of 100 scientific articles and 5 books published from 1982 to 2006, the authors state that the concept of HISB is not unified into a single and explicit definition. The definition is usually recognized as being obvious; and it is often taken as how people obtain health information (Lambert & Loiselle, 2007). After this critical review, Lambert and Loiselle have coined HISB as “the ways in which individuals go about obtaining information, including information about their health, health promotion activities, risks to one's health, and illness”.

Lambert and Loiselle (2007) have also characterized HISB by 1) the health information sources used, 2) the type and amount of health information seeking, and 3) the actions used to search for health information. This characterization is supported by the United

States National Cancer Institute (2014), which collects data about health information seeking in terms of the source, type, and amount of health information. Another literature review supports that HISB can be characterized by the source, type, and extent of health information sought (Lalazaryan & Zare-Farashbandi, 2014). Lambert and Loiselle (2007) have grouped the characteristics of HISB into two views: the information and method views. The information view of HISB focuses on the type and amount of health information sought. The type relates to the variety of health information searched with respect to, for example, specific diseases or health problems. The amount relates to the detail, the depth, or the frequency of health information sought. For example, some people might search for general information while others might prefer detailed and specific health information.

The method view of HISB includes the health information source used and the actions followed to obtain health information. The action used could be the strategies or techniques employed to search for health information from either manually or from a health information system. The sources of health information could be listening radio, watching television, reading newspapers or magazines, searching from the Internet, and discussing with physicians, family members, or friends (Lambert & Loiselle, 2007).

### **2.1.1. Source of Health Information**

Health information is obtained from different sources. Based on the literature review, the following health information sources are identified:

#### **2.1.1.1. Online: Source of Health Information**

The use of the Internet as a source of health information has increased in developed countries (Bargh & McKenna, 2004; Escoffery et al., 2005; European Commission, 2014; Fox, 2011a; Medlock et al., 2015; Medlock et al., 2012; Medlock et al., 2013) due to its easiness for immediate access (Miller & Bell, 2012; Sharif et al., 2015). For example, a study done among college students in the United States of America shows that 78% of the students report having looked for health information online in the past year, and more than 37% report that they search for health information three or more times from the

Internet (Percheski & Hargittai, 2011). Results of the Hong Kong Family and Health Information Trends Survey shows that the use of the Internet for health information has increased from 30.6% in 2009 to 35.6% in 2012 (Wang et al., 2013).

However, from the perspective of developing countries, the use of the Internet for health information is presumed to be low. For example, a university-based study in Nigeria among undergraduate students indicates that 22.2% of the students use the Internet to obtain health-related information on a weekly basis (Obasola & Agunbiade, 2016). The main reason for the low health information seeking could be due to limited access to the Internet (The World Bank, 2018a). In the study by Obasola and Agunbiade (2016), 42% of students perceive that the cost of the Internet is high. Similarly, a school-based study which is done among adolescents in Ghana identifies that only 35% of youths use the Internet to search for health information (Borzekowski et al., 2006). Another study which is conducted among the general public in China indicates that 59% of the participants use the Internet very often for preventive health information seeking (Yuli et al., 2012). Similarly, in the rural community of Malaysia, the prevalence of HISB is considered low. Results of the study in Malaysia indicate that 46% of the rural community seek online health information (Mohd-Nor et al., 2013). The Internet is also found to be a useful source of health information by few urban Peruvians (Garcia-Cosavalente et al., 2010) and public library users from Iran (Gavgani et al., 2013).

Results of the above mentioned research studies indicate that people from developed countries use the Internet as a common source of health information (Bargh & McKenna, 2004; Escoffery et al., 2005; European Commission, 2014; Fox, 2011a; Medlock et al., 2015; Medlock et al., 2012; Medlock et al., 2013), which enable them to actively engage in health information seeking. The reason for their active engagement could be due to instant and easy access to the Internet (Cutilli, 2010), and their adequate level of literacy (Nutbeam, 2000) and awareness on online health information sources (Health Information and Quality Authority, 2011). On the other hand, the use of the Internet to look for health information among people in developing countries is presumed to be low (Borzekowski et al., 2006; Mohd-Nor et al., 2013; Obasola & Agunbiade, 2016; Yuli et al., 2012) due to limited infrastructure and resources that can support access to the Internet (Obasola &

Agunbiade, 2016; Raj et al., 2015), and limited skill and knowledge in understanding and using technologies and the Internet (Nasrabadi et al., 2015; Patel & Dowse, 2013; Raj et al., 2015).

The next two subsections discuss the use of social media and search engines for health information seeking.

#### **2.1.1.1.1. Social Media: Online Source of Health Information**

In recent times, people have started using social media to obtain health information as it provides a platform for people to communicate and share information (Parvanta et al., 2011). People in social networks can post, share, and interact with others about specific health information, such as wellness, illnesses, treatments, and medications (Parvanta et al., 2011). A longitudinal study on the role of social networking sites in seeking health information online is conducted in the years 2006, 2008, and 2010 in the United States of America (Feng & Xie, 2015). The results show that 2.23% in 2006, 20.37% in 2008, and 40.05% in 2010 of the respondents acquire health information from social network sites, such as MySpace, Facebook, and Friendster. This shows that the use of social media as a source of health information is increasing in the United States of America. Another study is conducted on the contribution of social media to seek information about food-related risks in eight European countries (Kuttschreuter et al., 2014). The results identify social media as a complementary channel for consumers to seek information about food-related risks. A study from a developed country indicates that university students actively seek health information from social networking sites, such as Facebook, Twitter, LinkedIn, Instagram, YouTube, and Reddit (Hackman & Pember, 2016). This shows social media could be used as a potential health promotion tool for the university population.

From the perspective of developing countries, a small percentage of public library users (8.2%) in Iran has used social media as a source of health information. In Vietnam, about 73% of youths show interest in health information on Facebook (Zhang et al., 2017). The study from Vietnam reports the perception of health information seeking on Facebook but it fails to determine Vietnamese youths' actual engagement in social media for health

information seeking. A report shows that the percentage of web users visiting social media is 88.2% in Vietnam, 93.9% in Thailand, 96.1% in the Philippines, 91.6% in Malaysia, and 86.9 in Indonesia (Brown, 2013). The figures indicate that Southeast Asians are active in using social media. In Ethiopia, although there are frequent social media blocking and Internet censorship (Amnesty International, 2016), youths are found to be actively using social media, consuming 30 to 60 minutes per session (Hussain & Arasad, 2015). However, their engagement in health information seeking using social media is not explored. Therefore, in this study, the use of social media as a source of health information among university students from Malaysia and Ethiopia is examined.

#### **2.1.1.1.2. Use of Search Engine: Online Source of Health Information**

Search engines have become a significant means by which people access information. Generally, search engines can be broadly categorized into two, general and domain-specific search engines (Minnie & Srinivasan, 2011; Pang et al., 2014). While general search engines provide information on various information domains, domain-specific search engines, as its name indicate, provide information on specific disciplines (e.g. medicine, pharmacy, nursing, and economics) (Minnie & Srinivasan, 2011; Pang et al., 2014). General search engines are usually useful to support people engaged in basic search tasks in order to locate information. However, they may not be helpful for people who need a deeper understanding of a topic with multifaceted search tasks engagement which requires multiple search sessions and continuous interactions. Furthermore, general search engines may not be helpful for people engaged in exploratory searching where people further learn and investigate to have a deeper understanding of a topic (Marchionini, 2006).

A university-based study on online HISB is conducted to understand searching approaches of university students from a developed country (Pang et al., 2014). Results indicate that general search engines are found to be unsatisfactory to handle searches and health queries of university students. Using a general search engine, users perform surface search and face problem in accurately specifying health terms to formulate queries. They also face difficulty to filter and narrow search results due to limited medical knowledge and analytical skills. However, exploratory searching allows people to learn

health terms which enable them to filter, and narrow search results by exploring contents and links (Pang et al., 2014).

Research studies show that people in the developing world mainly use general search engines to look for health information (Gavgani et al., 2013; Inthiran et al., 2013; Yuli et al., 2012). For example, a study done in Iran identifies that most Internet users use search engines (64%), such as Google and Yahoo whereas only limited users (16.5%) use specific websites and databases to look for health information (Gavgani et al., 2013). Another university-based study done in Malaysia shows that more people use general search engines (46%) compared to health-specific search engines (23%) to search for health information (Inthiran et al., 2013). Similarly, a community-based study in China identifies that general search engines, such as Google and Baidu (Chinese language-search engine) are the most frequently used sources (65.5%). However, the use of health-specific search engines (17.9%) is identified as low in the study (Yuli et al., 2012).

The above-mentioned research studies show that people from developing countries search health information using general search engines although health-specific search engines support exploratory search which enables users to acquire a deeper understanding of the needed health information. This could be because searchers in developing countries may not be aware of health-specific search engines (Health Information and Quality Authority, 2011).

#### **2.1.1.2. Physicians or Health Professionals: Source of Health Information**

Despite the increased access to health information through the Internet, patients from developed countries prefer healthcare providers or physicians as the main sources of health information (Gutierrez et al., 2014; Hesse et al., 2005; Oh et al., 2012). For example, youths (40%) in Singapore prefer doctors for health information consultation (Mokhtar et al., 2009). This is because youths in Singapore consider doctors as a trusted source of health information (Mokhtar et al., 2009). Moreover, about 58% of the youths in Singapore indicate that doctors provide sufficient health information. However, about 62% of youths indicate that doctors are not worth the cost for health information consultation.

Although doctors are not worth the cost, they are still preferred and trusted source of health information (Mokhtar et al., 2009).

In developing countries, doctors are found to be a useful source of health information. A literature review on preventive HISB in developing countries identifies that people from developing countries use doctors as useful sources of health information (Muhammed & D'Souza, 2014). In Pakistan, doctors are found to be the third choice among adolescents for health information seeking (Thaver & Kamal, 2010). The first and second choices of young Pakistanis are parents and television, respectively. In Peru, both urban and rural Peruvians use health professionals as the main source of health information (Garcia-Cosavalente et al., 2010).

The above research studies from developed (Gutierrez et al., 2014; Hesse et al., 2005; Mokhtar et al., 2009; Oh et al., 2012) and developing (Garcia-Cosavalente et al., 2010; Muhammed & D'Souza, 2014; Thaver & Kamal, 2010) countries indicate that health professionals remain as a preferred source of health information. The reason to choose health professionals as a primary source of health information could be due to the trustworthiness of health professionals for health information (Hesse et al., 2005; Oh et al., 2012).

#### **2.1.1.3. Television, Radio, Family, and Friends: Source of Health Information**

The other sources of health information are television, radio, family members, and friends which are the commonly preferred source of health information among people from developing countries. For example, in Malaysia, the majority of participants from rural area prefer television and radio as main sources of health information (Mohd-Nor et al., 2013). Another research study done in Iran on HISB among library users shows that television, family members, friends, and books are found to be the most common resources to seek health information (Gavgani et al., 2013). The authors claim that one of the reasons for library users to use sources other than the Internet is due to the lack of people's consciousness towards the existence of health websites. Another community-based study which is done in India also indicates television and interpersonal

communication with friends, relatives and doctors as the most common sources of tuberculosis-related information (Prasad et al., 2013). A review paper on preventive health information seeking among the general population in China also indicates television as the major source of health information channel (Muhammed & D'Souza, 2014).

People from developing countries use radio and television as main sources of health information. Previous research studies indicate that people who use radio and television as sources of health information are more likely to be passive or unintentional seekers (Garcia-Cosavalente et al., 2010; Gavgani et al., 2013; Mohd-Nor et al., 2013; Muhammed & D'Souza, 2014). These passive seekers may not be able to obtain health information specific to their need as they do not seek health information intentionally. However, nowadays, the Internet penetration rate is growing fast in developing countries (The World Bank, 2015b). This creates an opportunity for people to actively engage in health information seeking. Hence, in this study, the use of the Internet for health information seeking is explored in developing countries.



Table 1: Summary of Research Articles on Source of Health Information

Information Source	Specific Source Used	Studies from Developing Countries	Studies from Developed Countries
<b>Online</b>	<b>Google</b>	Gavgani, V.Z., et al. 2013;	Escoffery, C., et al. 2005; Yuli, Y., et al. 2012
	<b>Yahoo</b>	Gavgani, V.Z., et al. 2013;	Escoffery, C., et al. 2005
	<b>WebMD</b>	-	Escoffery, C., et al. 2005; Fox, S. 2011
	<b>Social Media</b>	Gavgani, V.Z., et al. 2013;	Feng and Xie, 2015; Kuttischreuter et al., 2014; Hackman and Pember, 2016
	<b>*Unspecified</b>	Borzekowski, D.L., et al. 2006; Mohd-Nor et al., 2013; Sharif, S.P. et al. 2015; Obasola and Agunbiade, 2016	Weaver, J.B., et al., 2010; Medlock, S., et al. 2012; Medlock, S., et al. 2015; Miller, L.M. and R.A. Bell 2012; Hesse, B.W., et al. 2005; European Commission 2014
<b>Personal Communication</b>	<b>Family</b>	Gavgani, V.Z., et al. 2013; Muhammed KM and D'Souza, 2014	European Commission 2014
	<b>Friends</b>	Prasad, B.M., et al. 2013; Muhammed KM and D'Souza, 2014	European Commission 2014
	<b>Health Providers</b>	Garcia-Cosavalente H.P., 2010; Muhammed KM and D'Souza, 2014	Hesse, B.W., et al. 2005; Gutierrez, N., et al. 2014; Oh, K.M., et al. 2012; Wang, M.P., et al. 2013
<b>Mass Media</b>	<b>Television/ Radio</b>	Gavgani, V.Z., et al. 2013; Prasad, B.M., et al. 2013; Mohd-Nor et al., 2013	Wang, M.P., et al. 2013

\* Specific search engines or websites as a source of online health information are not specified in the reviewed articles

As shown in Table 1, the majority of the articles show that people from developed countries mainly use the Internet as a source of health information while television, radio, and personal communication (such as communication with doctors, family, and friends) are being used by people from developing countries. Only few research studies are done among university students from developing countries, showing the Internet as a main source of health information (Borzekowski et al., 2006; Obasola & Agunbiade, 2016).

### **2.1.2. Type of Health Information**

People seek health information to satisfy their needs (Fox, 2011b). Some search for illness (curative purposes), some for wellness (preventive measures), and others for both illness and wellness information (Weaver et al., 2010). Wellness information includes information about personal and public hygiene, environmental protection, proper diet, and adequate exercise whereas that of illness information includes information about disease or health problem control and treatment (Weaver et al., 2010). The classification of wellness and illness information can help to determine distinct health motives for health information seeking (Weaver et al., 2010).

Research studies show that people from developed countries look for both curative and preventive health information (Fox, 2011b; Medlock et al., 2012; Weaver et al., 2010). The Pew Internet survey report indicates that Internet users in the United States of America look for both wellness and illness information. In the Pew Internet survey, 55%, 43%, 27%, 25%, 19%, and 16% of the participants searched for specific diseases or health problems, medical treatments or procedures, weight control, health insurance, food safety, and drug safety, respectively (Fox & Duggan, 2013). Results of this study indicate there is a higher proportion of illness related information users. There is also a relatively significant number of wellness related information seekers in the United States of America. In another study, 40.6% of adults look for both illness and wellness related information online while an almost equal proportion of illness only-related (28.6%) and wellness-only (30.8%) related are being searched (Weaver et al., 2010).

However, it is worthwhile to note that people in developing countries mostly seek illness-related information (Gavgani et al., 2013; Muhammed & D'Souza, 2014; Prasad et al., 2013). They seek out information about specific disease control, illnesses, treatment choices, and drug and medication information. They also show interest to have a healthier lifestyle and diet, to lose weight, and to stay fit through physical activities. A community-based study done in China shows that illness and treatment-related information are the most popular information for the respondents to search for information from the Internet (Yuli et al., 2012). This shows that the respondents give less attention to preventive or

wellness related information. In the study in China, the most important information needs for the majority of the participants are causes and symptoms of diseases, and diagnoses and treatments. However, less focus is given towards information related to cholesterol, smoking cessation, blood pressure, diabetes, and stress management. Especially, information about psychosocial and stress management were considered less important by participants. A university-based study in Nigeria shows that undergraduate students seek illness-related information (Obasola & Agunbiade, 2016). Similarly, a school-based study done in Ghana among adolescents indicates that the most popular information searched for is sexually transmitted diseases. Information on diet or nutrition, fitness or exercise, sexual activity, drug use, sexual abuse, pregnancy, and heart disease are also the common health topics searched by the participants (Borzekowski et al., 2006). This shows that adolescents in Ghana go online to search for mostly illness-related information.

Table 2 shows that people from developed countries seek both wellness and illness related information, whereas people from developing country focus on illness-related information. One reason for this difference could be attributed to the difference in health knowledge. People from developed country exhibit an adequate level of literacy and knowledge on health issues (Nutbeam, 2000) which could lead them to pay attention to both wellness or preventive mechanisms of diseases and illness information. However, people in developing countries have limited health knowledge (Nasrabadi et al., 2015) which could lead them to pay attention to health information after they acquire health problems. The other reason for the difference could be because people from developing countries focus on mostly illness-related information due to the high levels of communicable and non-communicable diseases (World Health Organization, 2009a). In this study, the type of health information that university students from developing countries sought is explored.

Table 2: Summary of Research Articles on the Purpose of Health Information Searching

Purpose of health information searching		Studies from Developing Countries	Studies from Developed Countries
Wellness information	Nutrition	Borzekowski, D.L., et al. 2006; Yuli, Y., et al. 2012;	Escoffery, C., et al. 2005; Medlock, S., et al. 2012; Fox, S. 2011; European Commission 2014
	Physical Activity	Borzekowski, D.L., et al. 2006; Yuli, Y., et al. 2012;	Escoffery, C., et al. 2005; Weaver, J.B., et al., 2010; Medlock, S., et al. 2012; European Commission 2014
Illness information	Disease control and treatments	Borzekowski, D.L., et al. 2006; Yuli, Y., et al. 2012; Prasad, B.M., et al. 2013; Muhammed KM and D'Souza, 2014; Mohd-Nor et al., 2013; Sharif, S.P. et al. 2015; Obasola and Agunbiade, 2016	Weaver, J.B., et al., 2010; Medlock, S., et al. 2012; Fox, S. 2011
	Medication	Sharif, S.P. et al. 2015	Weaver, J.B., et al., 2010; Medlock, S., et al. 2012; Fox, S. 2011; Hesse, B.W., et al. 2005

### Summary: HISB, Sources and Purpose of Health Information Seeking

HISB has a substantial contribution to disease prevention and control activities and plays a major role in keeping good health status of a population, especially in developing countries. It has also the potential to reduce health knowledge gaps among societies through education. Individuals' health knowledge can be increased through disseminating health information in a well-designed online health information system. Hence, HISB has significant implication for health promotion campaigns. Despite the benefits, the magnitude of HISB remains low (Mohd-Nor et al., 2013; Yuli et al., 2012) in the general population of developing countries due to inadequate health literacy, low digital literacy, limited awareness on the sources of health information, and limited resources (Nasrabadi et al., 2015). Although university students' health issues are very important, but they are at risk of chronic diseases (Deliens et al., 2015; Graham et al., 2014; Larouere & Winters, 2013; Oman Global School-based Student Health Survey, 2012), their HISB is unexplored and not well understood in developing countries.

The most common sources of health information in developing countries are physician, television, radio, interpersonal communications, and magazines and newspapers. Even though the usage is very low (Borzekowski et al., 2006; Mohd-Nor et al., 2013; Obasola & Agunbiade, 2016), many research studies in developing countries identify the Internet as useful resource for health information seeking (Garcia-Cosavalente et al., 2010;

Gavgani et al., 2013). Only very few research studies explore the use of the Internet for health information seeking in developing countries.

The purpose of health information that people look for are generally classified as wellness (e.g. healthy lifestyle, weight loss, diet, physical fitness, and specific disease prevention) and illness (e.g. specific disease control, illnesses, treatment choices, and drug and medication) information (Weaver et al., 2010). People in developing countries seek most for illness-related information (Gavgani et al., 2013; Muhammed & D'Souza, 2014; Prasad et al., 2013). Most of the above studies mentioned are done among the general population where there may not be adequate Internet access. Given that university students are at risk of having health problems and considering the opportunity that the Internet is everywhere on university campuses, conducting HISB research among university students gains importance.

## **2.2. Factors Associated with Health Information Seeking Behaviour**

Previous research studies from developed countries have explored the factors associated with HISB. These studies show that HISB is affected by socio-demographical (Gavgani et al., 2013; Kim, 2015), psychological (Lalazaryan & Zare-Farashbandi, 2014), behavioural (Wang et al., 2013; Weaver et al., 2010), and health-related (health literacy and health status) factors. However, these factors are unexplored among university students from developing countries. In this study, socio-demographical (age, gender and Internet use experience), psychological (self-efficacy, perceived susceptibility to health problems and perceived severity of health problems), behavioural (alcohol consumption, smoking and physical activity), and health-related (health literacy and health status) factors are explored. Identifying the factors could help to promote health and educate individuals on university campuses.

The factors are selected because of their health influence in the university population. For example, self-efficacy, which is a psychological factor, is affected by stress which is prevalent among university students (Graham et al., 2014; Griffin, 2014). Torres and Solberg (2001) identify that low self-efficacy together with stress can negatively affect the health status of university students. Results of a study in Malaysia identify a significant

proportion of low self-efficacy to do a task among university students (Goje et al., 2014). The same study finds that students with low self-efficacy tend to be physically inactive. In addition, self-efficacy together with perceived susceptibility to and severity of health problems are considered as the driving force behind health information seeking (Rosenstock, 1974).

Behavioural factors, such as smoking, unhealthy diet, and physical inactivity are also prevalent among the young population (World Health Organization, 2014a). Smoking, unhealthy diet, and physical inactivity are the most risk factor for cardiovascular diseases which are the number one cause of death that contribute to premature deaths globally (World Health Organization, 2014a) and in Malaysia (Ministry of Health Malaysia, 2008; World Health Organization, 2014b). The World Health Organization (2014a) estimates that more than 75% of deaths from cardiovascular diseases are from middle and low-income countries.

Health-related factors, such as health literacy and health status are the other factors influencing HISB in developed countries. However, little is known about the factors among people in developing countries, especially university students. People from developing countries have limited health knowledge, influencing their understanding of health information (Raj et al., 2015). Therefore, there is a need to precisely understand university students' HISB and examine the associated factors so as to provide valuable information for use in health promotion campaigns. Results can also be used by health education and promotion professionals to promote health and improve the health status of the university population.

Therefore, in the following sections below, the underlying factors that are associated with HISB from previous studies are described.

### **2.2.1. Socio-demographical Factors**

Research studies from developed countries indicate that age is the most significant determinants of health information seeking (Fox, 2011a; Kim, 2015; Oh et al., 2012). Results of the Pew Internet Project indicates young people of age 18-29 years as the

most health information seekers in the United States of America. Even among health information seekers, youths are found to be more active information seekers in comparison to elderly (Pálsdóttir, 2008). Likewise, older people are found to be less likely to search for health information when compared to the younger people (Medlock et al., 2012; Wang et al., 2013; Wong et al., 2014). This is due to shame and embarrassment and lack of interest to manage their own health among older people (Cotten & Gupta, 2004).

On the other hand, a literature review on preventive health information seeking in developing countries shows that younger people seek health information more from television and the Internet while older ones seek more from newspapers and magazines (Muhammed & D'Souza, 2014). However, a community-based study done in China identifies no significant association between the level of age and preventive information need (Yuli et al., 2012). In this study, though not statistically significant, participants who never used the Internet for health information are identified to be elderly people aged over 65 years. A reason for this could be due to the age distribution of the sample in the study, which only six percent of the participants are over the age of 60 years. The above two articles show no clear relationship between HISB and age. However, among health information seekers, youths prefer the Internet to search for health information than older people. This could be due to the fact that young people are close to digital technologies and the Internet in comparison to older people (Fox & Duggan, 2013).

Research studies from developed countries indicate women as more popular health information seekers than men (Escoffery et al., 2005; Manierre, 2015; Medlock et al., 2012; Wang et al., 2013). Results of a study done on gender difference in health information seeking indicate that females with good health condition are more likely to look for cancer information than males. This is because of the masculinity and the factors (e.g. perceived cancer risk and fear of environmental exposures which men are less worry about) that may lead men to be less prone to seek information (Manierre, 2015). Even among health information seekers, women are found to be active information seekers in comparison to men, by acquiring health information intentionally (Pálsdóttir, 2008). On the other hand, men receive health information passively or unintentionally from

traditional media (TV, magazines, and radio) and interpersonal communication (Pálsdóttir, 2008).

Gender difference in HISB is also present among university students from developed countries. Results of a study in the United States of America indicates that significantly more female students seek out health information in comparison to male students (Escoffery et al., 2005). In Singapore, male youths are found to be more likely concerned about their health in comparison to females. However, they are less likely to seek out health information from magazines and to consult physicians for their illness (Mokhtar et al., 2009). The reason why females look for health information more than males could be due to the trend in staying healthy which may be published in health magazines, as the authors claimed. However, the gender difference in HISB is not well known among people in developing countries.

Other factors affecting HISB include online information searching experience. Results of a study from a developed country show that university students with more information searching experience are more likely to look for health information when compared to university students with less information searching experience (Escoffery et al., 2005). Furthermore, university students with more information searching experience are the major online health information seekers and will likely to continue to use the Internet for health information. Due to their university education, students are likely to have good Internet skills. Internet skills can motivate university students to look for health information (Janeice et al., 2013). However, the relationship between online information searching experience and HISB of university students from developing countries is unexplored.

In conclusion, research studies from developed countries identify relationship between HISB and socio-demographical variables (Fox, 2011b; Jung, 2014; Kim, 2015; Medlock et al., 2012; Oh et al., 2012; Pálsdóttir, 2008; Wang et al., 2013; Wong et al., 2014). Results of the studies show that females and youths are more likely to look for health information in comparison to male and older people, respectively. However, the relationship is not clearly known among university students from developing countries.



Hence, in this research study, an attempt is made to examine the association between HISB and socio-demographical variables from developing countries point of view.

### **2.2.2. Psychological Factors**

HISB can also be influenced by psychological factors, such as self-efficacy, belief, perceived susceptibility to and perceived severity of health problems. Self-efficacy is “the extent to which people believe they are capable of performing specific behaviours in order to attain certain goals” (Hensel et al., 2013). Results of research studies from developed countries indicate that people with high health self-efficacy are more likely to look for health information in comparison to people with low self-efficacy (Chen & Feeley, 2014; Lee et al., 2008; McMullan, 2006; Pálsdóttir, 2008). High self-efficacy can also contribute to a shift from passive to more active health information seeking. Results of a research study from a developed country show that individuals with higher health self-efficacy tend to be more motivated for new information and they seek health information actively (Chen & Feeley, 2014). The study also states that higher health self-efficacy can minimize health concerns and distress, which may motivate individuals to seek health information. This conclusion is supported by a study done among women diagnosed with breast cancer in a developed country (Lee et al., 2008). The study states that believing to control disease and high negative emotion seem to encourage patients to find more information. This indicates that HISB can be negatively influenced by low health self-efficacy and negative emotion. In addition, the study concludes that patients are not always motivated to obtain health information when they have negative emotions, such as feeling worried, being nervous, or terrified about their disease.

In another study from a developed country, the highest self-efficacy beliefs are identified among active health information seekers while the lowest self-efficacy beliefs are observed among passive seekers (Pálsdóttir, 2008). In the study, it is found that passive seekers have slightly lower self-efficacy beliefs as compared to moderately active seekers. Hence, high self-efficacy can lead people to better seek out health information so that they can have participatory decision making on their health together with their physician. This can result to have a more patient-centred interaction with health

professionals. High self-efficacy has a positive influence on people to look for health information (Chen & Feeley, 2014; Lee et al., 2008; McMullan, 2006; Pálsdóttir, 2008). However, the impact of these variables is not explored in developing countries where many people lack proper health knowledge, influencing their ability to seek out and understand health information (Nasrabadi et al., 2015).

Health information seeking can occur as a result of belief influences. A qualitative study which is done in Iran among cardiovascular patients identifies that health information seeking in patients occurs as a result of the influence of beliefs. According to this study, believing death as a cycle of life and giving little value to disease prevention and control, believing in traditional medicine, thinking that health information seeking is not the responsibility of patients, and lacking belief in the information provided by staffs other than doctors, such as nurses are the factors negatively affecting health information seeking (Gholami et al., 2014).

Perceived susceptibility to and perceived severity of health problems have a direct influence on HISB. A research study integrates the health belief model and technology acceptance model to investigate health-related Internet use among women in Malaysia (Ahadzadeh et al., 2015). The result indicates that perceived health risk, which consists of perceived susceptibility to and perceived severity of a health problem, has a positive effect on Internet use for health information. This implies women with a higher level of perception about health risks are more likely to use the Internet for health-related information than women with a low level of perceived health risk. The perceived benefit of the Internet has also a positive influence on health-related information seeking among Malaysian women. Similarly, a study has also measured perceived health risk through perceived seriousness and perceived susceptibility scale to study consumers' disease information seeking behaviour on the Internet in Korea (Yun & Park, 2010). The study has found that consumers with a higher level of health risk perception are more likely to seek disease information on the Internet.

As mentioned in section 1.3.3, university students are at risk of chronic diseases (Biswas et al., 2015; Oman Global School-based Student Health Survey, 2012), stress (Griffin,

2014), and risky sexual behaviour (Rubens et al., 2014). Understanding their risk perception in terms of perceived susceptibility and severity could help to avoid the risks through appropriate health information. Therefore, it is important to understand how these perception impact HISB of the university population.

### **2.2.3. Behavioural Factors**

Research studies from developed countries show that behavioural factors, such as physical activity and smoking affect HISB. For example, smokers and physically inactive people are less likely to look for health information (Wang et al., 2013; Weaver et al., 2010). Weaver et al. (2010) indicate that there is a strong association between type of health information (wellness and illness) and physical activity (Weaver et al., 2010). Wellness information seekers are most likely to engage in physical activities. Wang et al. (2013) also identify the relationship between the frequency of health information seeking and physical activity. Wang et al. (2013) state that being physically active is related to monthly health information seeking from the Internet, television or radio, and newspapers or magazines. Although smokers and physically inactive groups have greater health information need to restore their healthy behaviours, infrequent health information seeking is observed among these groups. In another study in Iceland, it is found that physically inactive people tend to be passive (unintentional) health information seekers (Pálsdóttir, 2008). However, the effect of behavioural factors on HISB is not explored from the developing countries point of view. As mentioned in section 2.2, smoking and physical inactivity are prevalent among young adults in developing countries. These factors are the main risk factors for number one killers, that is, cardiovascular diseases which affect middle and low-income countries most. Hence, it is necessary to study the effect of smoking and physical inactivity on HISB from the perspective of developing countries.

### **2.2.4. Health Literacy and Health Status**

The World Health Organization defines Health Literacy as “the cognitive and social skills which determine the motivation and ability of individuals to gain access to, understand and use information in ways which promote and maintain good health” (World Health

Organization, 2009b). One of the reasons for people to look for health information is to increase their health literacy level, and hence, health literacy is a very important factor to improve individual's access to health information and is crucial to empowerment (World Health Organization, 2009b). Health literacy affects people's ability to communicate their health history with health practitioners. While university students presume to have an adequate level of health literacy, they may face problems in understanding and obtaining health information (Yi, 2015). For example, in a study done in the United States of America, nearly 1 in 11 university students have less than adequate health literacy (Mackert et al., 2014). Hence, misassumption of university students to have an adequate level of health literacy may mislead healthcare givers to provide detail information on treatment follow-ups and prescriptions.

A research study shows that low levels of health literacy are associated with problems in providing an accurate medical history, completing medical forms, and understanding instructions for drug prescriptions (Manganello, 2008). These problems may influence disease treatment decisions and outcomes, as there is a strong association between health literacy and health outcomes. Results of another study support that people with inadequate health literacy are 68% less likely to look for health information in comparison to people with adequate literacy (Gutierrez et al., 2014). Similarly, other studies done in Australia and Finland identify that people with limited health literacy are less likely to look for health information (Ellis et al., 2012; Eriksson-Backa et al., 2012). Results of the qualitative study in Australia (Ellis et al., 2012) indicates that participants with a high level of health literacy exhibit the most engagement in HISB. However, participants with a low level of health literacy show little or no engagement in HISB as they simply accept information from health providers. While the participants with a high level of health literacy advise other people in their social network, participants with a low level of health literacy receive advice from people in their social network.

The above research studies from developed countries identify the relationship between health literacy and HISB (Ellis et al., 2012; Eriksson-Backa et al., 2012; Gutierrez et al., 2014; Mackert et al., 2014; Manganello, 2008). The studies indicate that people with better health literacy level are more likely to look for health information. However, little is

known from developing countries point of view where there is a lack of proper health knowledge (Nasrabadi et al., 2015). Results of a study from Iran find a significant association between literacy and information seeking behaviour of patients (Gavgani, 2010). In the study, a statistically significant association has been found between literacy and sources of health information. Literate patients look for health information from the Internet and library while less literate patients seek from friends. This study uses the general term 'literacy' to measure the literacy of patients, which might not substitute the term 'health literacy'. In addition, the study is conducted among patients which may have different health knowledge when compared to university students. Therefore, in this study, the health literacy of university students is measured, and its impact on HISB is examined.

Another health-related factor affecting HISB includes individuals' health status. Research studies show that having chronic diseases (Wang et al., 2013), diminished mental and physical health, having a medical prescription, poor quality of life are associated with frequent health information seeking (Weaver et al., 2010). Similarly, another study indicates that healthier people are less likely to seek out health information than their counterparts (Kim, 2015). However, in a different study, least healthy people are found to be passive health information seekers who have the lowest frequency of information seeking with the lowest motivation (Pálsdóttir, 2008). Pálsdóttir (2008) concludes that more healthy behaviour is related to more frequent health information seeking, which is the characteristic of active and moderately active health information seekers. This is because more health information seeking tends to increase health knowledge thereby improving the health status of individuals (Medlock et al., 2013; Weaver et al., 2010).

### **Summary: Factors Associated with Health Information Seeking Behaviour**

Limited research studies in developing countries are found in relation to the determinant factors affecting HISB. As a result, research studies from developed countries are reviewed to find out the important factors which may influence HISB in developing countries to serve as indications for further research. Results of the research studies from developed nation show that female and younger people are more likely to search for health information when compared to male and older people (Fox, 2011a; Kim, 2015; Oh

et al., 2012). However, in developing countries, the factors associated with HISB are not clearly known. Very few studies have shown discrepant results. Some researchers find that older people are more likely to search from magazine and newspaper while young people from television and the Internet (Muhammed & D'Souza, 2014). Other researchers have found no statistically significant association between HISB and age (Yuli et al., 2012).

Other factors affecting HISB include self-efficacy, online information searching experience, health status, health literacy, physical activity, and smoking. Generally, higher self-efficacy (Chen & Feeley, 2014; Lee et al., 2008; McMullan, 2006; Pálsdóttir, 2008), being physical active (Wang et al., 2013; Weaver et al., 2010), poor health condition (Kim, 2015; Wang et al., 2013; Weaver et al., 2010), non-smokers (Wang et al., 2013; Weaver et al., 2010), adequate health literacy level (Ellis et al., 2012; Eriksson-Backa et al., 2012; Gutierrez et al., 2014; Mackert et al., 2014; Manganello, 2008), and information searching experience (Escoffery et al., 2005) are found to seek health information more than their counterparts. However, there is an inconsistent study that shows healthier people to seek out health information more frequently than less healthy people (Pálsdóttir, 2008).

In the context of developing countries, it is very constructive to conduct a study that investigates the factors that can contribute to engaging university students to become health information seekers. The result might be useful not only for health promotion activities that can create awareness for students but also for academic societies and industrial consumptions to conduct further related research studies. A summary of the articles reviewed in relation to factors associated with HISB is presented in Table 3.

Table 3: Summary of Research Articles on Factors Associated with Health Information Seeking Behaviour

Factors affecting Health Information Seeking Behaviour		Findings	Studies Conducted in Developing Countries	Studies Conducted in Developed Countries
Demographic factors	Gender	Female look for health Information more than Male	-	Manierre, M.J. 2015; Kim, S. 2015; Wang, M.P., et al. 2013; Medlock, S., et al. 2012; Mokhtar I.A et al. 2009; Escoffery, C., et al. 2005; Pálsdóttir, Á., 2008
		No association between gender and HISB	Gavgani V.Z. 2010; Muhammed KM and D'Souza, 2014; Yuli, Y., et al. 2012	-
	Age	Young people look for health information more than older people	Muhammed KM and D'Souza, 2014;	Fox, S. 2011; Oh, K.M., et al. 2012; Kim, S. 2015; Pálsdóttir, Á., 2008; Wang, M.P., et al. 2013; Medlock, S., et al. 2012; Wong, C., et al. 2014
		No association between age and HISB	Yuli, Y., et al. 2012	-
	Education	People with higher education level seek health information more than people with lower education level	-	Fox, S. 2011; Oh, K.M., et al. 2012; Kim, S. 2015; Pálsdóttir, Á., 2008; Wang, M.P., et al. 2013; Jung, M. 2014; Rider, T. et al 2014
		People with lower education level seek health information more than people with higher education level	Muhammed KM and D'Souza, 2014	-
		No association between education and HISB	Yuli, Y., et al. 2012	-
	Online information searching experience	People with more information searching experience are more likely to seek out health information	-	Escoffery, C., et al. 2005
Psychological Factors	Self-efficacy	People with high health self-efficacy	-	Pálsdóttir, Á., 2008; Chen, Y. and T.H.

		are more likely to look for health information		Feeley 2014; Lee, S.Y., et al., 2008; McMullan, M., 2006
	<b>Perceived susceptibility to disease</b>	perceived susceptibility to disease positively affects HISB	Ahadzadeh et al, 2015	Yun & Park, 2010
	<b>Perceived severity of disease</b>	perceived severity of disease positively affects HISB	Ahadzadeh et al, 2015	Yun & Park, 2010
<b>Behavioural Factors</b>	<b>Physical activity</b>	Physical activity is associated with engagement in HISB	-	Weaver, J.B., et al. 2010; Wang, M.P., et al. 2013; Pálsdóttir, Á., 2008
	<b>Smoking</b>	Smokers are less likely to engage in HISB	-	Weaver, J.B., et al. 2010; Wang, M.P., et al. 2013
<b>Health-Related Factors</b>	<b>Health status</b>	Less healthy people are less likely to engage in HISB	-	Weaver, J.B., et al. 2010; Wang, M.P., et al. 2013; Kim, S. 2015;
		More healthy people are more likely to engage in HISB	-	Pálsdóttir, Á., 2008
	<b>Health Literacy</b>	Low health literacy is associated with less health information seeking	Gavvani V.Z. 2010	Mackert, M., et al. 2014; Gutierrez, N., et al., 2014; Ellis, J., et al., 2012; Eriksson-Backa, K., et al., 2012; Manganello 2008

### 2.3. Online Health Information Searching Behaviour

Online health information searching behaviour is a subset of HISB. It focuses on the interactions of users with information retrieval systems. It includes users' searching behaviours and strategies when interacting with an information retrieval system. In this section, challenges faced by health information searchers in developed countries are presented. Since there are very limited studies from developing countries, previous literature from developed countries is reviewed to have an understanding of health information searching and its challenges. In addition, the following searching behaviours are reviewed: query behaviour (number of query, query length, and conceptual factors of query formulation), result viewing behaviour (result clicked and source opened), and search outcomes. These searching behaviours are reviewed because they are commonly researched in other similar research studies (Belkin et al., 2003; Broussard & Zhang, 2013; Lopes & Ribeiro, 2015; Zhang, 2013; Zhang et al., 2012). In addition, the searching behaviours often reflect tactical and conceptual changes that are experienced by users



during information searching (Belkin et al., 2003; Broussard & Zhang, 2013; Lopes & Ribeiro, 2015; Zhang, 2013; Zhang et al., 2012).

### **2.3.1. Challenges of Health Information Searching**

In this section, the challenges of health information searching are discussed with respect to the understanding of health terminology. Health terminology is selected because people in developing countries have limited health knowledge (Nasrabadi et al., 2015) which may affect the use of health terminology to articulate queries and search for health information. Use of appropriate health terminologies during health information searching is essential to retrieve the required information.

#### **2.3.1.1. Challenges in Understanding Health Terminologies**

University students are likely to have good Internet skills as their university education might provide them with the skills to use the Internet for academic purposes. However, Internet skills alone might not guarantee the successful retrieval of health information. Students are still likely to experience difficulties in understanding and using appropriate terms to search for health information (Schroeder, 2010). The difficulties can affect the success of health information searching. Especially, students from other than health field may not be able to identify medical or health terms for their query, or their query may not match with medical vocabularies. For example, a research study from the United States of America on health literacy and health information identifies difficult health terminology as one of the barriers to find health information (Yi, 2015). In the study, the respondents report that they have problems in understanding health terminologies. People face difficulty in articulating health queries mainly due to limited health literacy which is characterized as having necessary skills and competencies needed to find, understand, evaluate, and use health information (Zarcadoolas et al., 2003).

Unfamiliarity with health terminologies has the potential to impact users' searching behaviour and continues to be a problem during health information searching. Results of a research study identify the effects of topic familiarity on users' health information searching behaviour (Puspitasari et al., 2015). According to the study, users with lower health topic familiarity submit more queries than those users with higher health topic

familiarity. The users submit more queries in order to have a better understanding of the topics and then locate the required information. Puspitasari et al (2015) have measured the difficulty of information sources based on their readability scores using Simple Measure of Gobbledygook formula (Mc Laughlin, 1969), cited in (Puspitasari et al., 2015). The result from the study shows that users' selection of content is influenced by the health topic familiarity. Users who are more familiar with a particular health topic tend to use more difficult content while users who are less familiar with health topic tend to use easier content. Topic familiarity has also the potential to impact users' query reformulations and spelling errors in health information searching process. A study indicates that users with higher health topic familiarity are less likely to reformulate queries and commit spelling errors but more likely to use more specific terms when compared to lower health topic familiarity group (Hu et al., 2013).

Therefore, in this study, university students' understanding of health terminology as a challenge for health information searching are explored.

### **2.3.2. Health Information Searching Behaviour**

In this section, users' searching behaviour is discussed in relation to querying behaviour, result viewing behaviour, and search outcomes.

#### **2.3.2.1. Querying Formulation**

Query formulation is the process of transforming a user's information need into a request to an information retrieval system. In order to retrieve the required information and achieve their needs, users need to correctly phrase their queries to an information retrieval system (Yeganova et al., 2009). However, laypeople sometimes even use plain English sentences (Inthiran et al., 2011; Luo et al., 2008) or questions (Zhang, 2013) as a query to search for health information. These make existing web search engines to less likely to give query suggestions. When plain English sentence is used as a query, the results returned from the search engine may not be accurate or even it could be totally irrelevant. Consequently, users are commonly disappointed by the search results, and are unlikely to click and view the results, resulting in an unsatisfactory searching outcome (Inthiran et al., 2011). In addition, the use of abbreviations or acronyms, use of slang

expression, or misspellings during query formulation are the common barriers to retrieve health information (Boden, 2014). Hence, in this study, university students from developing countries are observed on how they formulate health queries.

#### **2.3.2.1.1. Query length and number of queries**

Query length is correlated with users' satisfaction with search results (Belkin et al., 2003). Users are commonly observed to use simple and short queries which might lead them to unsatisfactory results. For example, results of a research study from the United States of America find that the average query length issued by users is ranged from 1.79 to 4 terms per task with an average number of queries ranging 1 to 2.63 per task (Zhang, 2013). However, most of the terms that form the queries are found to be stop words which are not much helpful to characterize users' information need (Zhang, 2013). From 10, 257 queries extracted from Alltheweb.com, Spink et al. (2004) finds 2.3 average terms per medical or health query with a mean of 2.2 queries per medical or health session, which is short. This study also finds that users show few numbers of queries per search task which show succinct query issuance. In the point of views of these research studies from developed countries, the behaviour of using too short queries and stop words to formulate keywords seem to be the main obstacles of searching. However, these behaviours are not known among people in developing countries. Their querying behaviour could be different because of their limited health knowledge which could affect the use of health terms to formulate queries.

A study indicates that longer queries are significantly associated with increased users satisfaction in search results (Belkin et al., 2003). As users issue longer queries that characterize their information need, their searching effectiveness increases, resulting in a better interactive information retrieval performance. However, Jiang et al. (2015) identify that query length negatively affect search satisfaction. These two findings are conflicting with each other in terms of the effectiveness of using long queries. These contradicting findings are worth to further study. Therefore, this research study attempts to determine the query length issued by university students and examine its effect on search outcomes.

#### **2.3.2.1.2. Contextual factors for query formulation**

Information retrieval can be affected by contextual factors, such as searchers and search task characteristics (Kelly, 2006). Query formulation which is an essential part of the process of information retrieval can also be affected by contexts. Context remains as an indefinable concept in information retrieval (Sparck Jones, 1981) but is commonly referred to. Kelly (2006) states that context in information retrieval comprises individual and situational variables. The individual variables consist of user features (e.g. gender, age, knowledge structure, and cognitive and learning styles) whereas that of situational variables include search task features (e.g. task difficulty and task complexity). Context could help information retrieval systems to learn and predict information needs of a user, relate one piece information to another, display required information in a suitable manner to users, relate tasks performed by a searcher and enable other searchers to use it.

Existing research studies that explore contextual factors of query formulation are limited. These few studies focus mostly on task features and often ignore user features that possibly affect information searching on the web. Lopes and Ribeiro (2010) study the effects of task and user features on query formulation in health information searching. The authors find that gender, years of experience in web search, frequency of health information seeking, previous experience on search topics, and ease of tasks are associated with query length. That is, being female, less experience in web searching, frequent health information seeking, previous experience on a topic and difficult search tasks are associated with longer query length. In the study, the effect of user features on a number of queries is ignored which is a key feature in the query formulation process. Kelly (2009) points out that including users in information retrieval system evaluation and investigating users' information searching behaviour are essential to improve information retrieval systems. Hence, in this study, the effect of user features on the query length as well as the number of queries are explored.

Another research study has attempted to identify factors affecting query formulation in web information searching (Aula, 2003). This study identifies experience in using computers and web search as factors affecting the query formulation process, indicating that experienced users issue longer and more queries than those users with less

experience in using computers and web search. The study by Aula (2003) is conducted in the general information domain. In this study, the contextual factors in the health-specific domain are explored. The health domain is rich in context. In addition, its search process is based on well-defined scenarios (e.g. prevention, treatment, or medication) whose contexts could help to improve health information retrieval systems.

#### **2.3.2.2. Query Reformulation**

Research studies from developed countries show that when users are not satisfied with their initial search results the users make conceptual changes, such as modification of previous queries or switching to new concepts, and hence, spending a long time in a searching session. In a university setting, Zhang (2013) has studied the searching behaviour of students in the United States of America and identified four types of conceptual changes during the query reformulation process. These are: “(1) specification: (the query becomes more specific); (2) generalization (the query becomes more general); (3) parallel movement (the reformulated query has a partial overlap with the previous query); and (4) switching to new concepts (the reformulated query has no overlap with the previous query)”. Similarly, Broussard and Zhang (2013) observe users while they reformulate queries several times to get a relevant result and to explore more to get specific information. In the study, participants make the following conceptual changes: switching to new concepts, replacing concept, and making a concept more specific. Users are also observed while correcting misspellings. The above findings show that health information searching is a sense-making and learning process which enable users to explore health concepts. Understanding users’ pattern of query reformulations can help to suggest design strategies to support users to better interact with health information retrieval systems in a more natural way. Therefore, in this research study, university students from developing countries are observed for their query formulation and reformulation patterns.

#### **2.3.2.3. Result Viewing Behaviour**

One of the challenges of users to search for health information is to filter and select relevant results. A study identifies that users are uncertain to select relevant websites after search results are returned by a search engine (Broussard & Zhang, 2013). They

either select results through trial and error or select results from the top hits, believing that the more relevant results are located at the top of a result page. They also select results based on website familiarity simply because they feel comfortable with the website they are familiarized (Broussard & Zhang, 2013). In another study, it is found that users select results based on website reputability and advertisement (Fiksdal et al., 2014). Users are likely to select a website if it is reputable and does not have a lot of advertisements and popups. Users consider a website as reputable if its organization name is well known, if it gets trust through academics and research, and if it is a local website. In this study, how university students from developing countries select search results after a search engine returned the results are explored. People from developing countries do not have awareness of online health information sources (Nasrabadi et al., 2015; Raj et al., 2015). This may affect their selection of relevant websites after search engine returns search results. Profiling their result viewing behaviour may have design implication on health information retrieval systems in developing countries so as to better support users to select relevant results.

#### **2.3.2.4. Search Outcomes**

People may be frustrated while searching for health information on the Internet. The frustration could be during the search process or at the result viewing stage. For example, Janeice et al. (2013) observe and finds that participants are frustrated due to their limited searching skills. Users with better searching skills are more likely to specify their queries and commit fewer errors to get satisfactory search results (Hu et al., 2013). On the contrary, when users are inexperienced with the Internet the search process will take much time and in the end, they are frustrated. A systematic review identifies that the average time to complete a searching task is  $17.1 \pm 13.19$  (SD) minutes, with minimum and maximum completion time of 1 and 75 minutes, respectively (Kelly & Sugimoto, 2013). The other result that Janeice et al. (2013) identified is dealing with a large amount of health information available on the Internet and the quality of the information they obtain. This result is supported by another study done in China which indicates that searchers are unhappy with the quality of health information obtained on the Internet because the information obtained is too general and repetitive (Yuli et al., 2012). The

searchers, in this study, are also frustrated because they do not get immediate feedback from the information retrieval system. They also questioned the credibility of the information obtained.

Results of another study found that users could not find the desired result while searching for health information (Zeng et al., 2004). In the study, the major factors that contributed to the search failure are: “the information found is not new, information found is too general, confusing interface or organization of a website, information overload (too much information is retrieved, use of simple search strategies (browsing or short text queries) that do not characterize users’ information needs well, and the variety and quantity of health information available, which make web navigation complex”.

The above research studies identify search outcomes of people from developed countries when searching for health information. However, little has been said about these from developing countries. Therefore, in this study, search satisfaction of university students is explored from the perspective of developing countries.

### **Summary: Online Health Information Searching Behaviour**

In the literature review, a low level of health literacy and limited searching skills are identified as challenges of health information searching which lead searchers to unsuccessful results (Yi, 2015). Although university students are expected to have good Internet searching skills, they may not be familiarized with health terms. Hence, it is crucial to explore and understand the search challenges of university students from developing countries.

The reviewed literature finds that laypeople from developed countries use plain English language to formulate health queries (Inthiran et al., 2011; Luo et al., 2008; Zhang, 2013). They also commonly issue simple and short queries when searching for health information (Spink et al., 2004; Zhang, 2013). These could lead them to get unsatisfactory search results and then be frustrated. Users are also being dissatisfied by the huge amount of health information searched online and the credibility of the results (Janeice et al., 2013). When users unsatisfied with the search results they made reformulation effort with four types of conceptual changes: specification, generalization, parallel movement,

and switching to new concepts (Zhang, 2013). This shows that health information searching is not only formulating queries and finding results but also a learning process. A summary of the articles reviewed in relation to users' online health information searching behaviour is provided in Table 4.

Table 4: Summary of articles on online health information searching behaviour

<b>Health Information Searching Behaviour Characteristics</b>	<b>Findings</b>	<b>Research Studies</b>
<b>Challenges in understanding and using health terminologies</b>	Users face difficulties in understanding and using health terms to search health information	Schroeder, H., 2010; Yi, Y.J., 2014; Zarcadoolas, et al., 2003
	Spelling errors in unfamiliar topics	Hu, R., et al 2013;
	People with limited health literacy face difficulty in articulating health queries and finding health information	Zarcadoolas et al., 2003; Yi, 2015
<b>Query behaviour</b>	Use of plain English sentence as query	Luo, G., et al., 2008; Zhang, Y., 2013; Inthiran A. et al 2011
	Use of slang expression	Boden, 2014
	Use of short queries	Spink et al., 2004
	Use of longer queries	Belkin, N.J., et al., 2003; Zhang, Y., 2013; Inthiran A. et al 2011
	More reformulation effort	Broussard & Zhang, 2013; Zhang, 2013
	Less reformulation	Spink et al., 2004
<b>Result viewing behaviour</b>	Uncertain to select websites after searching	Broussard & Zhang, 2013;
	Selection based on website reputability	Fiksdal et al., 2014
	Use of hyperlink	Pang et al., 2014
<b>Search outcome</b>	Frustration with the search results	Janeice, Eileen, & Trauth, 2013; Kelly & Sugimoto, 2013; Inthiran A. et al 2011; Yuli et al., 2012; Zeng et al., 2004
	Having Internet searching skills leads to satisfaction with search results	Hu, R., et al 2013; Belkin et al., 2003



## **2.4. Theoretical Framework**

In order to develop a comprehensive conceptual framework for this research, four basic health information seeking models are used: Wilsons' 1999 nested model of information seeking and information searching research areas, Longo's 2010 expanded model of health information seeking behaviour, the Health Belief Model, and Sutcliffe and Ennis' Process Model of Information Searching. Therefore, in this section, detail information about these four theoretical models and the reasons to select each model is provided.

### **2.4.1. Wilson's 1999 Model of Information Seeking and Information Searching Research Areas**

Wilson's model is selected because it enables to view information searching behaviour as a subset of information seeking behaviour research area (Wilson, 1999). This allows looking at information searching as a process embedded in information seeking behaviour. The model also helps to define health information seeking and searching behaviours into: 'the ways by which people obtain information' and 'the interactions of searchers with an information retrieval system', respectively (Wilson, 1999). The definitions help to limit the scope of this research study.

Wilson reviewed various information behaviour models in 1999 and developed a new nested model of information seeking and information searching research areas (Figure 2) consisting of information behaviour, information seeking behaviour (which is a subset of information behaviour), and information searching behaviour (which is a subset of information seeking behaviour) (Wilson, 1999). Wilson stated that information seeking and searching behaviours are complementary. While the former focuses on the ways by which people use to find and obtain access to information, the latter concerned with the interactions of searchers with an information retrieval system. In his review, Wilson pointed out that models in information seeking attempt to describe information seeking activities, the motivations behind for the activities, or the links of phases in information seeking. Wilson also suggested for new research on how knowledge of modes of information seeking behaviour affects the information searching process. Hence, in this

research study, information searching behaviour is viewed in the wider perspective of information seeking behaviour in the context of health.

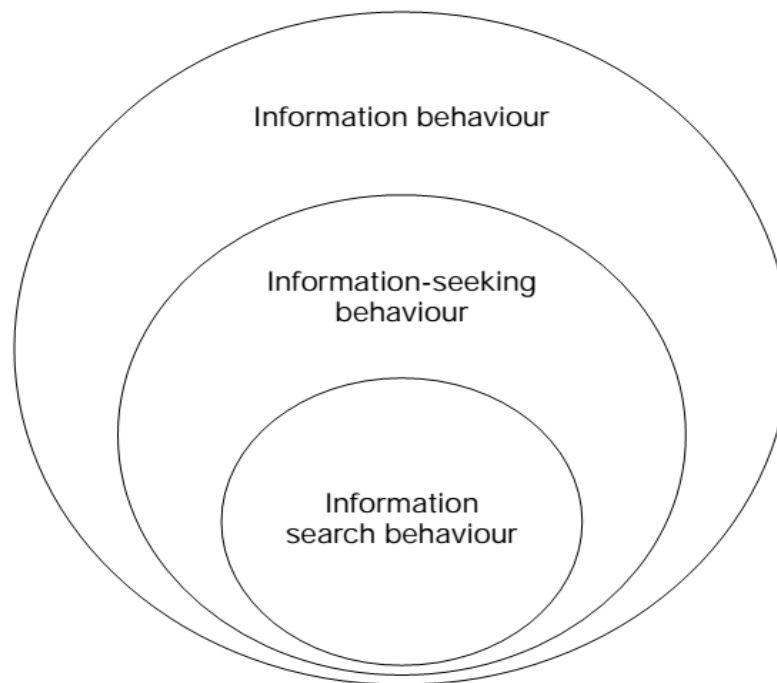


Figure 2: Wilson's 1999 nested model of information seeking and information searching.  
Source: (Wilson, 1999)

#### **2.4.2. Longo's Expanded Model of Health Information Seeking Behaviour**

Longo's 2010 model is chosen because it takes into account the list of the contextual and personal factors affecting HISB comprehensively (Longo et al., 2010). It also considers health literacy as one factor affecting HISB. Further, the model considers passive health information seeking in addition to active health information seeking. This fills the gaps in the current HISB literature in developing countries which states that people in developing countries are not just active information seekers as identified by previous research studies (Gavgani et al., 2013; Mohd-Nor et al., 2013; Muhammed & D'Souza, 2014). People can receive health information passively (unintentionally) in addition to actively (intentionally) seeking health information. Furthermore, the model is well known by the National Cancer Institute of the United States of America (Treiman & Squiers, 2005) cited in (Longo et al.,

2009). Although the model is initially developed for cancer-related information, it can be applied to other health information with minor modifications (Longo et al., 2010).

The Longo model was initially developed in the year 2005 in order to assess the information seeking behaviour of patients and consumers (Longo, 2005). The model was revised and used in the year 2009 with the inclusion of 'cross-cultural communication' and 'health literacy' variables (Longo et al., 2009). Longo used this model to conduct a quantitative study to understand breast-cancer patients' information seeking behaviour. The same model was used in the year 2010 (Figure 3) to investigate the information seeking behaviour of diabetic patients qualitatively (Longo et al., 2010).

The factors included in Long's model are: 1) **Contextual Factors**, such as "health status, health care structure, delivery of care, information environment, information seeking for self, family member, or friend at risk or with current medical problem, interpersonal social supports, networks"; and 2) **Personal Factors**, such as "demographic factors, socioeconomic factors, health history, genetics, stress, education, culture, language, attitudes, behaviours, current health status, cognitive ability, interpersonal communication". However, some of the variables are characterized to patients, not to university students, for example, health care structure and delivery of care. Hence, for this study purpose, Longo's model is adapted by including the following variables in the conceptual framework:

1. **Demographic factors:** age and gender
2. **Health-related factors:** health status and health literacy
3. **Behavioural factors:** physical activity, smoking, and alcohol consumption

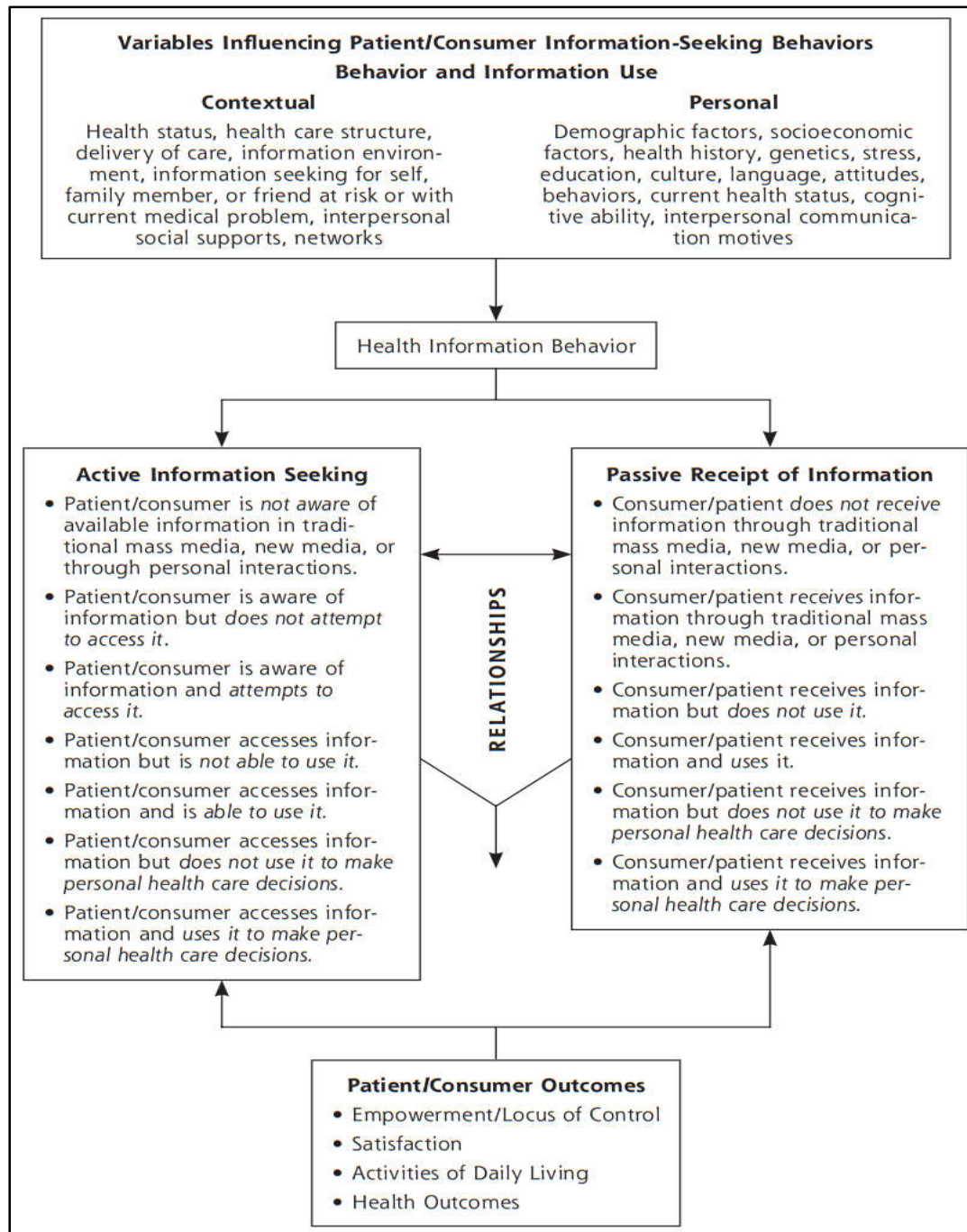


Figure 3: Longo Health Information Model: information seeking, passive receipt, and use. Source: (Longo et al., 2010)

### **2.4.3. Health Belief Model**

The limitation of Longo's model is that it does not consider psychological factors affecting HISB. The factors are perceived susceptibility to and severity of health problems and health self-efficacy, which are well described in the Health Belief Model (HBM) (Rosenstock, 1974). The factors in the HBM model are considered as driving forces behind health information seeking. Together with health information seeking, the factors are central to health-related perceptions and behaviour for health outcome. Combining Longo's model and HBM enables to have a comprehensive conceptual framework that includes the factors contributing to and the barriers inhibiting HISB.

Health information seeking through communication channels, such as the Internet, television, radio, magazines, newspaper, and interpersonal communication, can serve as a cue to take appropriate action for disease control and prevention (Rosenstock, 1974). In turn, perceived severity and susceptibility of health problems within the Health Belief Model can influence health information seeking (Yun & Park, 2010). These perceptions are considered as driving forces behind health information seeking. Hence, it is necessary to integrate these factors with other key variables in Longo's model to predict HISB. In the next few paragraphs, Rosenstock's work and the variables in the HBM are described in detail.

Health Belief Model (Figure 4) is a conceptual framework used to understand health and related behaviour. It consists of four major variables of perceptions to predict individuals' engagement in health-related behaviour (e.g. disease control and prevention). These are: (1) perceived susceptibility to health problems, (2) perceived severity of health problems, (3) perceived benefits of measures to reduce the negative health outcomes, and (4) perceived barriers that prevent individuals to take action to get a better health outcome.

Perceived susceptibility to a health problem refers to an individual's subjective perception of the risks of acquiring that particular health problem. According to Rosenstock (1974), people perceiving their susceptibility to a health problem are more likely to take action to reduce the risk of acquiring that health problem.

Perceived severity of health problems shows the perceived seriousness of a particular health problem, which may vary among people depending on the level of difficulties that individuals believe a given health problem will create for them. In this case, the HBM predicts that people perceiving a given health problem as serious are more likely to prevent developing the health problem or reduce the severity of the problem.

Perceived benefit refers to one's perception of values of preventing or reducing a particular health problem. To take action, an individual needs to believe that the action will provide him/her a positive health outcome (disease prevention or control).

The fourth variable is the perceived barrier which refers to one's perception of obstacles to prevent or reduce a specific health problem. Although an individual is threatened by a health problem and is aware of the benefit of taking actions, perceived barriers may prevent the person to take actions for better health outcome.

Perceived susceptibility and seriousness have a cognitive component and are dependent on knowledge on a given health problem. Both are believed to provide a force that leads to an action of direction that thought to be influenced by perceived benefits of measures to reduce the negative health outcomes and by perceived barriers that prevent individuals to take action to get a better health outcome.

With a weak level of susceptibility to or severity of a health problem, relatively strong stimuli could be required to trigger an action. Hence, an additional necessary variable called as "cue to action" is included to complete the model. The cues could be internal (e.g. signs and symptoms) or external, such as interpersonal communication, media, and social networks and the Internet. These stimuli can trigger health behaviour and make individuals aware of feelings about health problem so as to take appropriate action for better health outcome. For example, a 51-year-old woman decided to contact a physician after reading a health magazine who suggests that what she previously thought symptoms of aging could actually be signs and symptoms of ovarian cancer. Subsequently, it turned out that she actually had ovarian cancer for which surgery was

done that saved her life. In this case, the magazine served as a cue to take action (Longo, 2005).

Other potential variables, named as “modified factors”, are also included in the model. The variables consist of demographical (e.g. age, gender, education, and ethnicity), socio-psychological (personality, social class, and peer pressure), and structural variables (e.g. disease knowledge and disease history). The model suggested that these modified factors may have an indirect influence on the above mentioned four main variables.

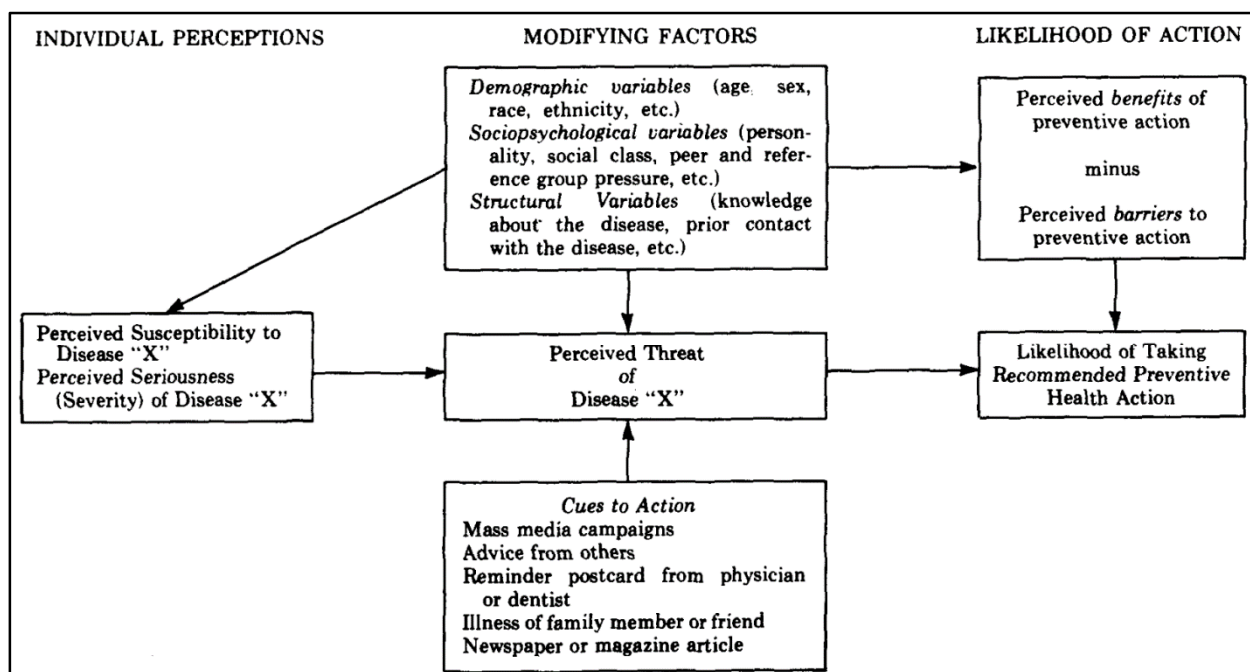


Figure 4: The Health Belief Model as a predictor of preventive health behaviour. Source: (Rosenstock, 1974)

#### 2.4.4. Sutcliffe and Ennis' Process Model of Information Searching

The fourth model used in the conceptual framework is the Sutcliffe and Ennis' process model of information searching (Figure 5). This model views information seeking as an interaction cycle consisting of identifying tasks, followed by formulating queries, viewing results, and if needed, reformulating queries (Sutcliffe & Ennis, 1998). The process is

repeated until a satisfactory result is found. The model is chosen because it describes the standard information searching process. The model enables us to investigate and understand users' querying and result viewing behaviours while interacting with an information retrieval system.

Generally, information seeking process is assumed to be an interaction cycle consisting of identifying an information need, followed by formulating a query, accessing and evaluating results, and then, if necessary, reformulating query and repeat the process until a satisfactory result is found. Sutcliffe and Ennis developed a theoretical model of investigating information retrieval by identifying two major components: "process model of information searching and knowledge representation to support the process" (Sutcliffe & Ennis, 1998). The knowledge representation encompasses the searcher's knowledge on a particular information retrieval system, information resources, domain, and information retrieval knowledge. Whereas, the process model includes four main activities, such as identifying problems, articulating information need, formulating a query, and evaluating results. Recognizing the problem helps to articulate and set an information need which usually is expressed by keywords or queries in order to search and find a result from an information retrieval system.

Sutcliffe and Ennis's model is helpful when users' problem is an external task or a general problem. Decomposition method might be used to divide the problem into smaller groups when the problem is complex. Once the problem has been identified, information need articulation will be made through the identification of main concepts or terms which will be utilized during query formulations. Query formulation can be complex or simple depending on the information retrieval system used and the users' skill in transforming the pre-identified main concepts or terms into query language with the help of Boolean operators. The next step that decides the continuity of the cycle is evaluating results which could be either accepting retrieved documents or continuing searching. If the decision is to continue searching, query reformulation will be made and a new result will be found. The cycle will continue until a satisfactory result is found. Evaluating retrieved results mainly depends on users' knowledge on a subject. The more familiarity with a domain area, the better to evaluate the relevance of the results.



Within the information searching process, query formulation and reformulation, number of queries issued, and the length of queries under the querying stage are looked into. In the result stage, result viewing behaviour is explored. Finally, search satisfaction is examined at the end of the searching process.

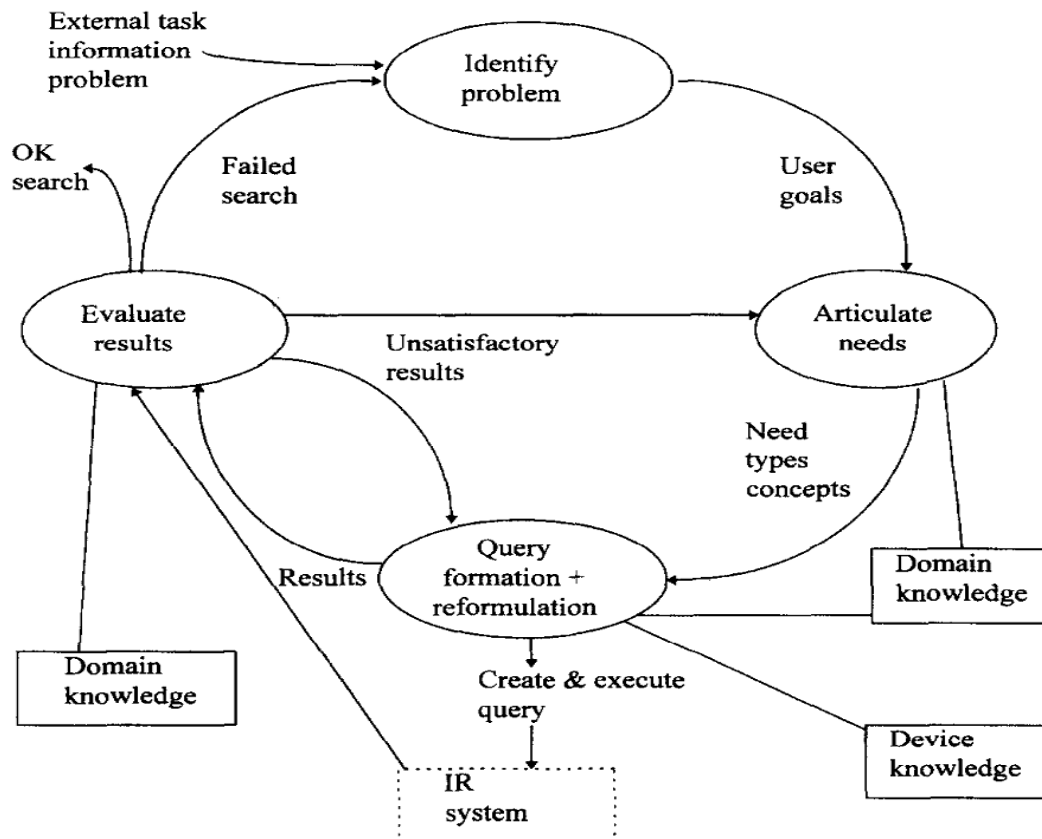


Figure 5: The process model of information searching activities and knowledge sources.  
Source: (Sutcliffe & Ennis, 1998)

#### 2.4.5. Proposed Conceptual Framework

As described above, while Longo's model mainly focuses on the HISB and the factors associated with it, the health belief model describes the variables affecting health behaviour to take actions which in turn can affect HISB. Sutcliffe and Ennis's model emphasizes the information searching process. The combination of active and passive information receipt; along with demographical, psychological, behavioural, and health-

related factors; and the process of health information searching in the online environment; can represent a far more comprehensive conceptual framework of health information seeking behaviours, as well as a framework to examine the health information searching behaviour and its challenges faced by university students from developing countries. Therefore, based on the above four models, a combined conceptual framework of this research is developed (Figure 6). The proposed conceptual framework helps to guide this research study.

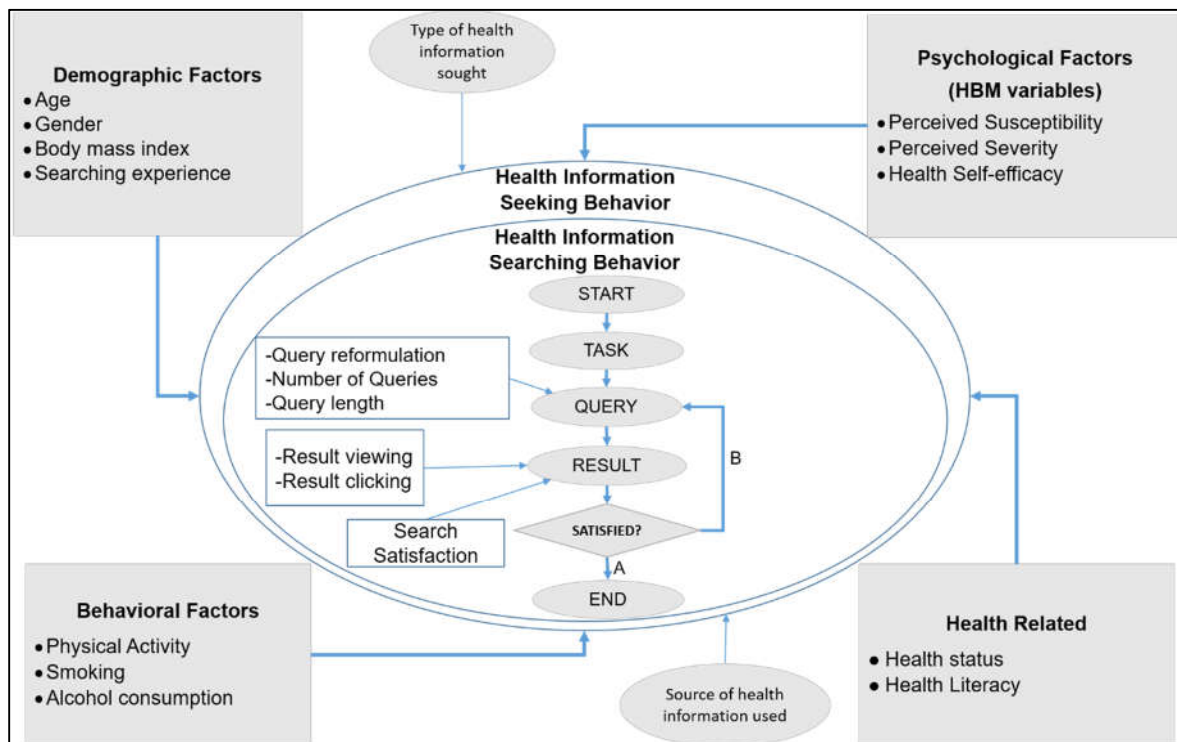


Figure 6: Conceptual Framework of Health Information Seeking Behaviour. A) Searcher ends up with a satisfactory result or unsatisfactory with a decision to abandon searching. B) Searcher reformulate a query in case of unsatisfactory results

As can be seen in Figure 6, the two big ellipses are taken from Wilson's models of information seeking and searching research areas. The demographical, behavioural and health-related factors are taken from Longo's extended models of HISB. Since Longo's model does not include the psychological factors, they are adapted from the Health Belief

Model. Finally, the figure inside the inner ellipse, which shows the process of information searching, is adapted from Sutcliffe and Ennis process models of information searching.

## **2.5. Chapter Summary**

Generally, very few studies are conducted on health information seeking behaviour in developing countries, mostly among adults and the general population (Mohd-Nor et al., 2013; Yuli et al., 2012). These studies focused mainly on the magnitude of HISB which is identified as low. Especially, the use of the Internet by people from developing countries for health information is very low (Borzekowski et al., 2006; Mohd-Nor et al., 2013; Obasola & Agunbiade, 2016) although it is a cost-effective and easy way to access information. The common sources of health information in developing countries are physicians, television, radio, interpersonal communications, magazines and newspapers (Garcia-Cosavalente et al., 2010; Gavgani et al., 2013; Mohd-Nor et al., 2013; Muhammed & D'Souza, 2014). These information sources enable people in developing countries to receive health information accidentally or unintentionally. Hence, it is possible to say that the majority of health information seekers in developing countries are passive seekers. It is also identified that people in developing countries are focused mainly on illness-related information (Gavgani et al., 2013; Muhammed & D'Souza, 2014; Prasad et al., 2013).

Very limited research studies are identified when they attempted to show relationships between HISB and its determinant factors in developing countries. These studies show inconsistent results when compared with research studies which are done in developed nations. For example, HISB does not have significant relationship with gender and age in developing countries (Yuli et al., 2012) while such relation does exist as reported in research studies which are done from developed countries (Fox, 2011b; Jung, 2014; Kim, 2015; Medlock et al., 2012; Oh et al., 2012; Pálsdóttir, 2008; Wang et al., 2013; Wong et al., 2014). A possible explanation could be people in developing countries are receiving health information passively and equally from media, such as television and radio (Garcia-Cosavalente et al., 2010; Gavgani et al., 2013; Mohd-Nor et al., 2013; Muhammed & D'Souza, 2014). Therefore, investigating and discovering the factors associated with

health information seeking behaviour of university students could help to inform healthcare planning and policy-making to improve the health condition of the university student population in developing countries.

The same is true in the development of health information retrieval systems which are generally designed based on the conceptualization and understanding of users' information searching behaviour in developed countries. People in developed countries have better knowledge and understanding of health information sources (Health Information and Quality Authority, 2011; Nutbeam, 2000). They used domain-specific search engines to locate health information. Whereas most people in developing countries mainly used general search engines (Gavgani et al., 2013; Inthiran et al., 2013; Yuli et al., 2012) which may not be helpful to explore and have a deeper understanding on health issues (Pang et al., 2014).

Challenges experienced by university students when searching for health information include difficulties in understanding health terms to search and evaluate health information (Schroeder, 2010; Yi, 2015); use of plain English sentence and slang expression to formulate queries thereby frustration and dissatisfaction with search results (Inthiran et al., 2011; Luo et al., 2008; Zhang, 2013); limited searching skills (Janeice et al., 2013), low level of health literacy, and use of short queries cause of unsatisfactory results (Spink et al., 2004; Zhang, 2013). Understanding these challenges could allow characterizing the searching behaviour of university students so as to suggest design strategies to improve health information retrieval systems.

## **CHAPTER 3: RESEARCH METHODOLOGY**

In this chapter, the following sections are covered: 1) Research methods in HISB studies, 2) Research design and setting, 3) Participants, 4) Sampling technique, 5) Data collection tools and techniques, 6) Data gathering procedure, and 7) Data management and analysis. First, the research methods in HISB studies are discussed. Under research methods in HISB, survey techniques and information retrieval experimental methodologies are discussed. The purpose of the discussion on survey techniques and information retrieval experimental methodologies is to justify and select appropriate survey technique and experimental method for this research study. Then, the research design and settings are described. Research settings are described to give an overview of the research area and the study population from which the sample is drawn. Next, information about the study participants, inclusion and exclusion criteria are provided. After a description of the study participants, the sampling techniques followed by the data collection tools and techniques are described. Under the data collection tools and techniques, information about the web-based survey, simulated situation, personal task, search domain, Morae keylogging software, observational checklist, and post-experiment survey are provided. Following the data collection tools and techniques, the data gathering procedure is discussed. Finally, methods used for data management and data analysis are explained.

### **3.1. Research Methods in Information Seeking Behaviour Studies**

In HISB studies, researchers focus on the type and amount of health information (Lambert & Loiselle, 2007), the source of health information, the factors associated with HISB, the process of health information searching and its outcome, and the interactions between users and health information retrieval systems (Kelly, 2009). The research methods or techniques used to investigate such studies involve surveys and experiments.

### **3.1.1. Survey**

A survey is an “information collection methods used to describe, compare, or explain individual and societal knowledge, feelings, values, preferences, and behaviour” (Fink, 2013). Creswell (2009) states that the survey provides a quantitative or numeric description of the sampled population on specific variables of interest. Researchers use it to determine the values and relationships of variables that can be generalized to a study population. The survey has the strength of replication and allows comparisons over different groups, times, and places.

A number of HISB research studies use the survey method to collect data. For example, Weaver et al. (2010) have used an online survey to assess the types of HISB and identify their covariates. A sample of 3140 participants have been approached through email and only 559 responses are analysed which indicates a large number of non-responses. Similarly, another study done among university students has used a web-based survey to identify active and passive health information seekers (Basic & Erdelez, 2014). An interview-based survey is applied in a study conducted on information seeking behaviour of parents and caregivers of children with mental illness in Tanzania (Lwoga & Mosha, 2013). The above research studies point out the value of surveys to gather quantitative data. National level HISB studies also use survey methods to collect data. For example, the Hong Kong family and health information trends survey use the telephone survey technique to study trends on HISB among Chinese (Wang et al., 2013). Similarly, the United States of America Health Information National Trends Survey use mail survey technique to study cancer-related information (National Cancer Institute, 2014).

Therefore, in this study, a web-based survey method is used to gather information about HISB and its associated factors. The reason to choose this method is that it has the capability to quantify and measure the values and relationships of variables that can be generalized to a study population. It has also the strength to allow comparisons over different groups, in this case, university students. A web-based survey is also a more convenient method of data collection in university campuses where an Internet connection is widely available. The web-based survey can be found in Appendix B.

### **3.1.2. Information Retrieval Experimental Methodologies**

Interactive information retrieval is “information retrieval with users” (Kelly, 2009). In interactive information retrieval, users’ behaviours are studied when they interact with information and systems. Classic information retrieval studies focus on system evaluation to improve information retrieval systems. However, interactive information retrieval studies include system evaluations along with users’ information search behaviour and their interaction with systems (Kelly, 2009). There are many possible issues to consider in order to understand how users interact with information. Some of them include, the conceptual structures of a particular domain that people work with, the processes they follow, the user interface structures they work with, and how information is communicated. To conceptualize information interaction researches, Kelly (2009) further summarize the research areas into system-focus, archetypical, and human-focus studies.

#### **3.1.2.1. System Focus Studies**

The system focus research studies focus on the evaluation and development of algorithms that can improve precision and recall of search results. Most Text Retrieval Conference<sup>2</sup> (TREC) studies are examples of system-focused researches. In such kind of studies, users’ interactions with information retrieval systems are not studied. For example, Yang et al. (1992) have used algorithms to modify and improve users query based on relevance judgements. The study uses test collection that consists of document collection, a test suit of information need which is expressed as queries, and a set of relevance judgements (relevant or non-relevant). Another example of the system-focused research study is a study by Arora et al. (2016) on the evaluation of Google and Yahoo using precision and recall. In the above two studies, users are not involved in the evaluation of information retrieval systems.

---

<sup>2</sup>Text REtrieval Conference (TREC) is a program aiming to support research within the information retrieval community by providing infrastructure for evaluation of text retrieval methodologies. Source: <http://trec.nist.gov/overview.html>

### **3.1.2.2. Archetypical Interactive Information Retrieval Studies**

The archetypical interactive information retrieval studies include “log analysis, TREC interactive studies, experimental information behaviour, and information seeking behaviour with information retrieval system” (Kelly, 2009). In interactive information retrieval, log analysis is the most common method to investigate searching behaviours, such as querying, result viewing, and result clicking behaviours. Transaction log analysis is the most frequently used method to understand users’ searching behaviour and interaction with information retrieval systems. For example, Park and Lee (2013) have used transaction logs to describe users’ searching behaviour on science and technology information in a web-based information retrieval system. Another study by Avery and Tracy (2014) have also used transaction logs to determine how students search in the context of a library instruction session.

TREC interactive studies are the other category of archetypical information interactive study, which focuses on evaluating the features of an information retrieval system in relation to users’ information seeking context. Such type of study collects data on system performance and usability, as well as feedback from users through an interview. For example, TREC 6, 7, and 8 interactive track studies aim at evaluating the features of information retrieval systems and understanding users’ information seeking context (Larson, 2001). In addition to the evaluation of the features, interviews (in TREC 7 and 8) are conducted to collect information on demographic details, search experience, about each search and the system, and users’ perceptions of the system.

The next type of archetypical information interactive study is information behaviour study which focused on users’ behaviour in the information searching process rather than a particular retrieval system. The ultimate goal of such studies is to make users’ searching experience as similar as possible in order to investigate the cause and effect of searching.

Information seeking behaviour with an information retrieval system is the last type of archetypical information interactive study listed by Kelly (2009). Such type of studies often focused on investigating users’ natural searching behaviour, searching strategies, and



relevance assessment of search results through observation techniques. These studies aim at an understanding of users' natural search behaviours in order to improve the development of better information retrieval systems. For example, Inthiran et al. (2015) use interactive information retrieval methodology and observe medical students to describe their information searching behaviour on MedlinePlus. Another study also uses MedlinePlus to study users' searching behaviour and experience on specific health-related information searching (Zhang, 2013). The study by Zhang (2013) uses an observation method to collect data from undergraduate students who do not have experience with MedlinePlus.

### **3.1.2.3. Human-Focus Studies**

The human-focus studies concern with users' information seeking behaviour in context without concern on an information retrieval system. These studies are known by their particular focus on human, their information need and searching behaviour. The common ways to collect data in such kind of studies are observations and interviews. Previous related research studies have implication for the improvement and design of information retrieval systems. A typical example could be a study done by Lopes and Ribeiro (2010) on the context effect on query formulation and subjective relevance in health information searches. Lopes and Ribeiro (2010) have used demographic characteristics, health information search experience, medical topic familiarity, task clarity, task easiness, and relevant assessment of search results to contextualize students' information need and seeking behaviour. The study has used experimental methodology with observation technique in four general search engine and three health-specific search engines to characterize searching behaviour of undergraduate students.

In this study, the human-focus method is used in order to explore university students' health information searching behaviour. The technique is chosen because it is suitable for the research objective which aims to understand university students' health information searching behaviour and their searching challenges. The technique is focused on investigating users' searching behaviour, searching strategies, and relevance assessment of search results through observation techniques. The method is useful for

research studies aiming at understanding users' search behaviours in order to improve and design better information retrieval systems. Therefore, in this study, an attempt is made to understand university students' behaviour when they search for health information so as to suggest design strategies to improve information retrieval systems.

### **3.2. Research Design and Setting**

In order to investigate health information seeking and searching behaviours of university students from developing countries, quantitative and cross-sectional study design is employed. This study is undertaken in Malaysia and Ethiopia as cases of developing countries. The two countries are selected as an example of middle and low-income countries which could represent developing countries. A report is published on the Sustainable and Economic Development Assessment (SEDA), which looks countries based on ten dimensions, such as income, economic stability, health, education, infrastructure (including technologies and the Internet), income equality, civil society, governance, and environment (The Boston Consulting Group, 2015). In the report, Malaysia and Ethiopia achieve a middle and lower level of sustainable economic development, respectively. Moreover, The World Bank classifies Malaysia and Ethiopia as middle and lower income countries, respectively (The World Bank, 2015a).

According to the Malaysian Communications and Multimedia Commission, the Internet penetration rate of the country is about 67% (Malaysian Communications and Multimedia Commission, 2015). The highest percentage of Internet users (24.2%) belongs to the age group of 20-24 years. However, the percentage distribution of the non-Internet users in the age group 20-24 years is 8.2%. Most of the Malaysian Internet users (24.1%) are from the state of Selangor. This is why university students with in the state of Selangor are included as study participants.

In Ethiopia, the 2016 Internet penetration rate is 15.4% which is extremely low (The World Bank, 2018a). This is mainly due to the limited telecommunication infrastructure, and the underdeveloped private sectors in telecommunication services (Ministry of Communication and Information Technology Ethiopia, 2016). Although the Internet penetration is low, students from Ethiopia are presumed to have Internet access on

university campuses. Students from Ethiopia use the Internet for academic, communication, and entertainment purposes (Moly, 2014). However, little is known about the students' engagement in health information seeking. This is why university students from Ethiopia are selected as study participants.

Monash University, the University of Malaya, and the Swinburne University of Technology from Malaysia and the University of Gondar from Ethiopia are selected as data collection sites. The three universities from Malaysia are taken in order to increase the diversity of the study participants considering the fact that Malaysia is demographically and culturally diversified which could lead to different HISB. For example, previous research studies indicate that demographic factors, including ethnicity, influence health information seeking (Galeshi et al., 2018; Richardson et al., 2012). At Monash University, it is observed that most of the students are Chinese by Ethnicity while most of the students at the University of Malaya are Malay. Malay and Chinese are the two most common ethnic groups constituting more than 90% of the population in Malaysia (Department of Statistics Malaysia, 2016). There is also a considerable number of students who are international and from other ethnic groups (e.g. Indian) from both universities. In order to represent students from East Malaysia, Swinburne University of Technology is included. However, only the University of Gondar from Ethiopia is included because of the ease of logistics. In addition, the principal investigator of this research study is from the University of Gondar, which makes it easy to conduct the survey. Monash University and the Swinburne University of Technology are situated in the states of Selangor and Sarawak in Malaysia, respectively, whereas the University of Malaya is located in the federal territory of Kuala Lumpur, Malaysia. The University of Gondar is located in the state of Amhara, Ethiopia.

### **3.3. Participants**

Undergraduate students from developing countries enrolled in Monash University, University of Malaya, Swinburne University of Technology Sarawak, and the University of Gondar are included as study participants. However, students from health-related discipline are excluded due to the fact that these students are most likely to be well-versed in health knowledge which may have the potential to affect the result of this study.

In this study, two groups of participants are included. The first group is for the web-based survey whereas the second group is for the experiment. In the web-based survey, undergraduate students from Monash University, University of Malaya, Swinburne University of Technology Sarawak, and the University of Gondar are included as study participants. However, in the experiment, undergraduate students only from Monash University and the University of Malaya are included as study participants. Monash University and the University of Malaya are selected for the experiment because of the ease of logistics. Monash University is the university where the principal investigator is enrolled in for a PhD program. The University of Malaya is nearby to Monash University that data collection can be carried out easily. All participants who are participated in the experiment are included in the web-based survey as well. More detail information about the web-based survey and experiment are provided in section 3.5.

### 3.4. Sampling Technique

In order to recruit participants for this research study, two sample groups, one for the survey and another for the experiment, are considered. For the survey, a total sample size of 768 (384 for Malaysia and 384 for Ethiopia) is calculated. The sample size is computed using the single population proportion formula:  $sample\ size = \frac{z^2 \times p(1-p)}{d^2}$ . The OpenEpi web tool is used to calculate the sample size, considering 50% of proportion (p), 95% confidence level (z), 5% absolute precision (d), a design effect of 2, and a non-response rate of 20%. Using the formula,  $the\ total\ sample\ size = \frac{1.96^2 \times 0.5(1-0.5)}{0.05^2} \times 2 = 768 + 20\% \times 768 = 922$ . However, only 901 undergraduate students (354 from Malaysia and 547 from Ethiopia) participated in the survey. Convenience sampling technique is used to select sample subjects in Monash University, University of Malaya, Swinburne University of Technology Sarawak, and the University of Gondar. Convenience sampling allows for easy access to participants in and outside classrooms. In this case, participants are approached and invited to take part in the survey until the required sample size is fulfilled.

For the experiment, a total of 176 undergraduate students from Monash University and the University of Malaya who already involved in the survey have participated. Bujang et

al. (2017) recommend that small sample size is sufficient to undergo regression analysis in experimental studies because an experimental study is commonly a well-planned study which attempts to control confounding factors during the design stage. In this study, potential confounders are controlled during the experiments. For example, 1) cookies in the computers where the experiments are carried out are cleared before each experiment starts, 2) participants are told to use computers for only experimental tasks so that the participants will not be distracted by other activities, 3) the experimental tasks are performed in computer labs, and 4) the experimental tasks are provided randomly to participants so as to minimize task order effect. Furthermore, Bujang et al. (2017) suggest a sample size of 156 considering a level of significance of 0.05, a power of 80%, and 10 independent variables. In this study, a sample size of 176 with a level of significance of 0.05, a power of 80%, and independent variables of 14 is considered for the experiment.

### **3.5. Data Collection Tools and Techniques**

This research study has two parts, a survey and an experiment. In the survey, a web-based survey is used as a data collection tool to gather information on HISB and its associated factors. The experiment involves users using a computer to search on three simulated situation and one personal task. In the experiment, keylogging software is used to gather information about the searching behaviour of participants from Monash University and the University of Malaya. The experiments are carried out in any search domain participants preferred. Moreover, a post-experiment survey, which is applicable only to the experiment, is used to gather search outcomes and challenges of participants after the experiment is carried out. The detail information about the data collection techniques and tools are presented in the following sub-sections below.

#### **3.5.1. Web-based Survey**

An online survey is used to obtain information on HISB, socio-demographic, health-related, behavioural, and psychological factors. The questions related to HISB, the source of health information, the type of health information, the frequency of health information, and factors associated with HISB are adapted from the 2014 US Health Information National Trends Survey (HINTS) (National Cancer Institute, 2014). HINTS is a nationally

representative survey administered by the US National Cancer Institute in a targeted population of aged eighteen and above.

HISB is measured by using a question from the United States Health Information National Trend Survey (HINTS). The Consumer Health Informatics Research Resource (CHIRr) describes the question as the most common approach to measure health information seeking (Hensel et al., 2013). The question is: "Have you ever looked for information about health or medical topics from any source?" (National Cancer Institute, 2014). Following this question, source and type of health information are asked.

Physical activity, health self-efficacy, perceived susceptibility to a health problem, and perceived severity of a health problem are measured using questions from the United States HINTS (National Cancer Institute, 2014). Two item questions are used to measure physical activity: 1) "In a typical week, how many days do you do any physical activity or exercise of at least moderate intensity, such as brisk walking, bicycling at a regular pace, and swimming at a regular pace?" 2) "On the days that you do any physical activity or exercise of at least moderate intensity, how long do you typically do these activities?" In order to measure health self-efficacy, a single item question is used: "Overall, how confident are you about your ability to take good care of your health?" Similarly, perceived susceptibility to a health problem is measured using a single item question: "How concerned are you about getting any health problem?" (National Cancer Institute, 2014). Perceived severity of a health problem is also measured using a single item question: "How getting a health problem would affect your life?" (National Cancer Institute, 2014).

Health literacy is measured using a tool from the European Health Literacy Survey (HLS-EU) (Sorensen et al., 2013). The Asian Health Literacy Association adapted the tool to be used by researchers in Asian countries (Asian Health Literacy Association (AHLA), 2013). Further, a study is undertaken to assess the validity and internal reliability of the tool in the Malaysian context (Mohamad et al., 2014). This study found high reliability; and concluded that the tool can be utilized in the Malaysian context. The HLS-EU questionnaire is also applied in Ghana to determine the health literacy level of street youths (Amoah et al., 2017). Therefore, the 47 item HLS-EU questionnaire is used to determine the health literacy of the participants under study. The 47 items are assessed

using a 4-point self-report scale (from 1 = very difficult to 4 = very easy). The detail information on how variables are defined and measured is provided in Appendix A.

The survey tool is designed using Google Forms which is a tool used to create and analyse surveys. Google Forms (<https://www.google.com/forms/about/>) is a free survey designer tool that allows setting skip and validation logic. Besides, it allows exporting data using Google Spreadsheet which makes it possible to transfer the data to statistical packages, such as SPSS or STATA, for further analysis. In order to identify each response in the survey, a unique identification number is assigned. The questions for the web-based survey can be found in Appendix B.

### **3.5.2. Experiment**

The experiment involves users using a computer to search on three simulated situation and one personal task. The experiment includes three data gathering techniques: search interactions using simulated and personal tasks, observation checklist, and post-experiment survey. In this section, first, simulated situations and simulated tasks are described. Then, information about the personal task is provided. This is followed by a discussion about search domains. Next, the keylogging software, which is used to gather search behaviour of participants, is described. After a description of the keylogging software, information about the observation checklist is provided. Finally, the post-experiment survey is described.

#### **3.5.2.1. Simulated Situations**

A common way to undertake experimentation to understand users' search behaviour is to use simulated situation (Kelly, 2009). A simulated situation consists of a short 'cover-story' that elaborates a situation that motivates to an individual requiring to search for information from an information retrieval system (Borlund, 2000). It helps to describe the search task to study participants about the source of the information need, and the environment of the situation, the problem that needs to be solved. It also enables study participants to make sense about the objective of the search (Borlund, 2000). The goal of performing a search in a simulated situation is not about to find the correct answers but to provide a snapshot of the search behaviour of participants. In addition, one can

compare and contrast search behaviour of participants because simulated situations create a standardized scenario that brings up a common information need among participants. Commonly, a simulated situation consists of a simulated task situation and an indicative request (Kelly, 2009). The task situation represents the information need of targeted users whereas the indicative request provides information to users on what to search.

In this research study, simulated situations are employed in order to describe the searching behaviour of students from Monash University and the University of Malaya when they search for health information on the web. The reason behind to apply a simulated situation for this research over experimentation using artificial tasks (e.g. tasks created from users' query logs and TREC task collection) is to make a context for the study participants that might motivate them to execute the tasks. Simulated situations can also help to conduct this research rapidly with a large number of participants (Kelly, 2009).

In this research study, participants are provided with four simulated tasks to perform in a computer lab. Four simulated tasks are used because real search behaviour might not be obtained with only one or two simulated situations (Kelly, 2009). Kelly (2009) suggests that simulated situations should not last over an hour so as to reduce participants' fatigue which could affect the result of a study. From the pilot study, which is conducted with 15 participants for both the web-based survey and experiment over a period of five days, the average data collection time is identified as 45 minutes. Previous studies use three (Zhang et al., 2012) to four (Puspitasari et al., 2015) simulated tasks to study users' health information searching behaviour. Hence, based on the timing from the pilot study and the experience from previous research studies, four simulated tasks and one personal task are used in this research study.

In order to develop scenarios for the simulated situation, health information needs of participants are gathered from the pilot study. However, a clear pattern of health information needs could not be identified amongst the participants because their information needs were evenly distributed amongst different health issues (Table 5). Hence, the simulated tasks are developed based on the health information needs of the



participants and the health problems of the young population in developing countries. The simulated situations can be found in Appendix C.

Table 5: Health information needs of 15 participants from the pilot study

Health Information needs	Number of participants
Diet or Nutrition	3
Exercise	3
Mental Health	2
Skin health	2
Sexual health	2
Cough	2
Drug use	1

The health problems selected are dengue fever, smoking, obesity, and physical activity. For example, dengue fever is a major health problem causing an economic burden in developing countries, particularly in Southeast Asian countries (Guzman & Istúriz, 2010). The disease is among the most serious health problem in Malaysia, affecting social function, wellbeing, and quality of life of predominantly children and young adults (Shepard et al., 2013). The prevalence of the disease is higher among people in the age group of 20-29 years in Malaysia (Chew et al., 2012). As of October 2015, over 98,500 dengue cases with 267 deaths are reported in Malaysia (World Health Organization, 2015b).

Smoking is the major risk factor for non-communicable diseases which contribute to 80% of premature deaths in developing countries (World Health Organization, 2015c). In Malaysia, non-communicable diseases are the leading cause of premature deaths (Ministry of Health Malaysia, 2011) contributing to 73% of total deaths (World Health Organization, 2014b). According to the 2012 Global Adult Tobacco Survey Report, over 10,000 Malaysians die as a result of smoking-related illnesses per annum. The daily use of tobacco among people in the age group 15-24 years is 32% in Malaysia (World Health Organization, 2012). Obesity and physical inactivity are other major risk factors contributing to cardiovascular diseases which are responsible for  $\frac{3}{4}$  of deaths in developing countries (World Health Organization, 2015a).

Dengue, smoking, obesity, and physical activities could be the health information needs of undergraduate students in Malaysia because these health problems are found to be the health problems of the young population in Malaysia. A review article supports that consumers health information needs are context dependent (Ramsey et al., 2017). Health information needs of an individual may depend on current physical, mental, and social wellbeing of the individual and his/her surroundings (Ramsey et al., 2017).

#### **3.5.2.2. Personal Task**

In this research study, participants are requested to provide a brief description of their personal task, which has to be developed based on their health information need. The purpose of the personal task is to compare and contrast search behaviour between simulated task and personal task. Participants are told to prepare a health-related personal task and bring it along with them. The personal task could be a task searched before or a new one. In order to balance task order effect and minimize threats to internal validity, the order of the simulated tasks and the personal task are randomly rotated. Sample brief descriptions of participants' personal task are provided below.

- *My eye power is high in terms of myopia and astigmatism. I want to know how I could reduce/maintain eye power besides seeking laser treatment. Besides, I want to know how to reduce the strain in my eyes because of the long hours using the computer to finish my assignment.*
- *How to lower cholesterol using medication?*
- *How to cut down body fats and increase muscle mass?*
- *Anxiety & Depression: causes and how to control them without the need for drugs*
- *Why tears form in my eye under light especially sunlight?*

Therefore, in this research study, four simulated tasks using a simulated situation technique and one personal task are used to study health information searching behaviour of students from Monash University and the University of Malaya, Malaysia.

#### **3.5.2.3. Search Domain**

The literature review indicates that people from developing countries mainly use general search engines to search for health information (Gavgani et al., 2013; Inthiran et al., 2013;

Yuli et al., 2012). Hence, participants are given the freedom to use any search domain they prefer to perform all search tasks.

#### **3.5.2.4. Keylogging software: Morae**

In order to gather data about querying and result viewing behaviours, Morae key logger software is used. Morae is useful to record users' interactions with information retrieval systems. The Morae software captures audio, video, on-screen activity, and keyboard and mouse inputs during a search session (TechSmith Corporation, 2015). However, in this research study, only on-screen activity, and keyboard and mouse inputs are captured. A search session starts when a participant key in the first query of a simulated task to a search engine and ends when the participant starts to fill up the post-experiment survey. A file name with a similar unique identification number which is given in the web-based survey (described in section 3.5.1) is used to save the screen recordings. This enables us to link the data from the screen recording with the data from the web-based survey.

#### **3.5.2.5. Observation Checklist**

An observation technique is used to collect data about search activities. Extraction of data about search activities using observation technique is preferred over log-based extraction because observation technique includes the context of the search. The limitation in using observation technique is it is prone to selection and researcher biases (Kelly, 2009). In order to minimize these biases during observation, an observation checklist is prepared to guide the process and to structure the observations. The observation checklist is used to extract data about search activities that are recorded using Morae keylogging software. The observation checklist contains querying activities and search result interaction activities.

The querying activities includes type of query (keyword, run-together words and question), presence of health terms in a query, presence of stop words in a query, presence of operators in a query, presence of misspelling in a query, number of queries, query length, query reformulation that changes the form of query (reform to question, change term sequence and correct spelling), and query reformulation pattern (specification, generalization, parallel movement and switch to new concepts). A term in

a query is said to be a health term when the term used is found in the MeSH (Medical Subject Headings) system. The MeSH system is a comprehensive controlled vocabulary designed for the purpose of indexing journal articles and books in the life sciences (U.S. National Library of Medicine, 2018). It serves as a thesaurus that facilitates searching. During result clicking and viewing stage, the activities observed are the search engine used (general purpose or domain specific search engine), result clicking (clicked or not clicked), viewing results beyond the first page, and the number of specific websites viewed.

The above-listed querying and result interaction activities are selected because they are important and basic elements in the process of information searching. For example, the querying activities reflect tactical and conceptual changes that are experienced by users during searching (Belkin et al., 2003; Broussard & Zhang, 2013; Lopes & Ribeiro, 2015; Zhang, 2013; Zhang et al., 2012). In addition, several users' studies have studied the querying and result interaction activities to improve information retrieval systems (Belkin et al., 2003; Broussard & Zhang, 2013; Lopes & Ribeiro, 2015; Zhang, 2013; Zhang et al., 2012). If participants are observed with other search activities when they perform all the search tasks, the search activities are recorded in the 'remark' section of the observation checklist. The observation checklist can be found in Appendix D.

#### **3.5.2.6. Post-experiment Survey**

The post-experiment survey is conducted immediately after the simulated and personal tasks are performed. The purpose of the post-experiment survey is to find out search challenges and search satisfaction of participants. The post-experiment survey contains both closed and open-ended questions. While the close-ended questions are intended to ask participants to rate their satisfaction with the search results, the open-ended questions are asked to understand the challenges faced and the search strategies used by the study participants. Participants are encouraged to explain why they are dissatisfied with the returned results. In addition, the post-experiment survey contains questions related to participants' perception of task difficulty using options, such as very difficult, difficult, neutral, easy and very easy.

The post-experiment survey is designed in Google Forms together with the survey so that each data will have a similar unique identification number. This enables to link the data from the web-based survey, screen recording, and post-experiment survey. The post-experiment survey can be found in Appendix E.

### **3.6. Data Gathering Procedure**

A web-based survey among participants from Malaysia and Ethiopia and experimentation using simulated situation among participants from Monash University and the University of Malaya are employed to collect data. Through the web-based survey, data on HISB and its associated factors, such as demographic, behavioural, psychological, and health-related factors are gathered. Following the web-based survey, the experimentation is carried out to collect information on searching behaviours of the participants. After the experiment, a post-experimental survey is administered to gather information on search outcomes and search challenges of the study participants.

Students are approached and invited to participate in the web-based study through personal communications made at the end of their class. Semester classes that are delivered during data collection period are selected to recruit participants. During personal communication, a short introduction to the study and its purpose are given. After they agreed to participate in the study, participants are requested to prepare a personal task which has to be set based on their health information need. Then, an appointment is made. Upon arrival, the participants are provided with the ethics approval letter and explanatory statement (Appendix F) of the study. Then, the participants are requested to review and sign on the informed consent form (Appendix G). After consent, they are provided with instruction letter (Appendix H) about the web-based survey, experiment and post-experiment survey. Then participants are requested to complete the web-based survey. Following the web-based survey, four simulated tasks and a personal task in a random order (to balance task order effect) are given at once to the participants. The participants are told to perform each task on any website or information retrieval system they prefer. They are also told to take as much time as they want to complete the tasks. A search session is said to be completed when the participants start to fill up the post-experiment survey regardless of whether the search results are found or not. After

completing all the tasks, the participants are requested to fill up the post-experiment survey which is about their satisfaction with the search results and the challenges they faced while searching on the web. The visual representation of the data gathering procedure is shown in Figure 7.

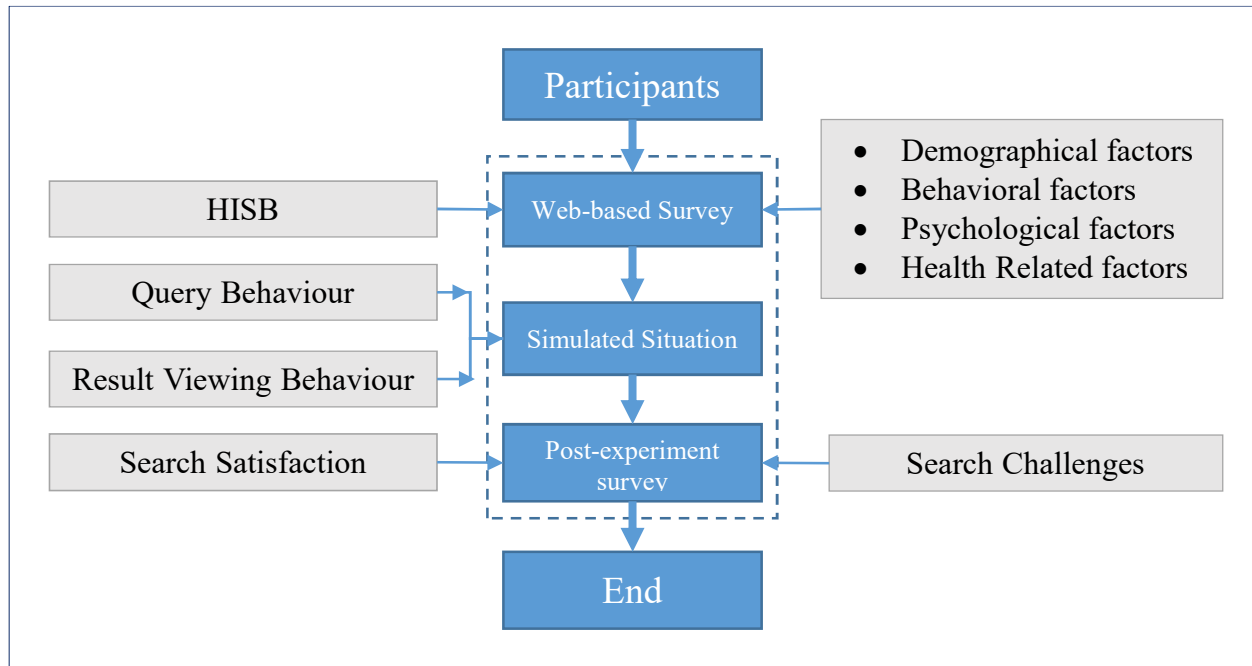


Figure 7: Diagrammatic representation of data collection methods and procedure

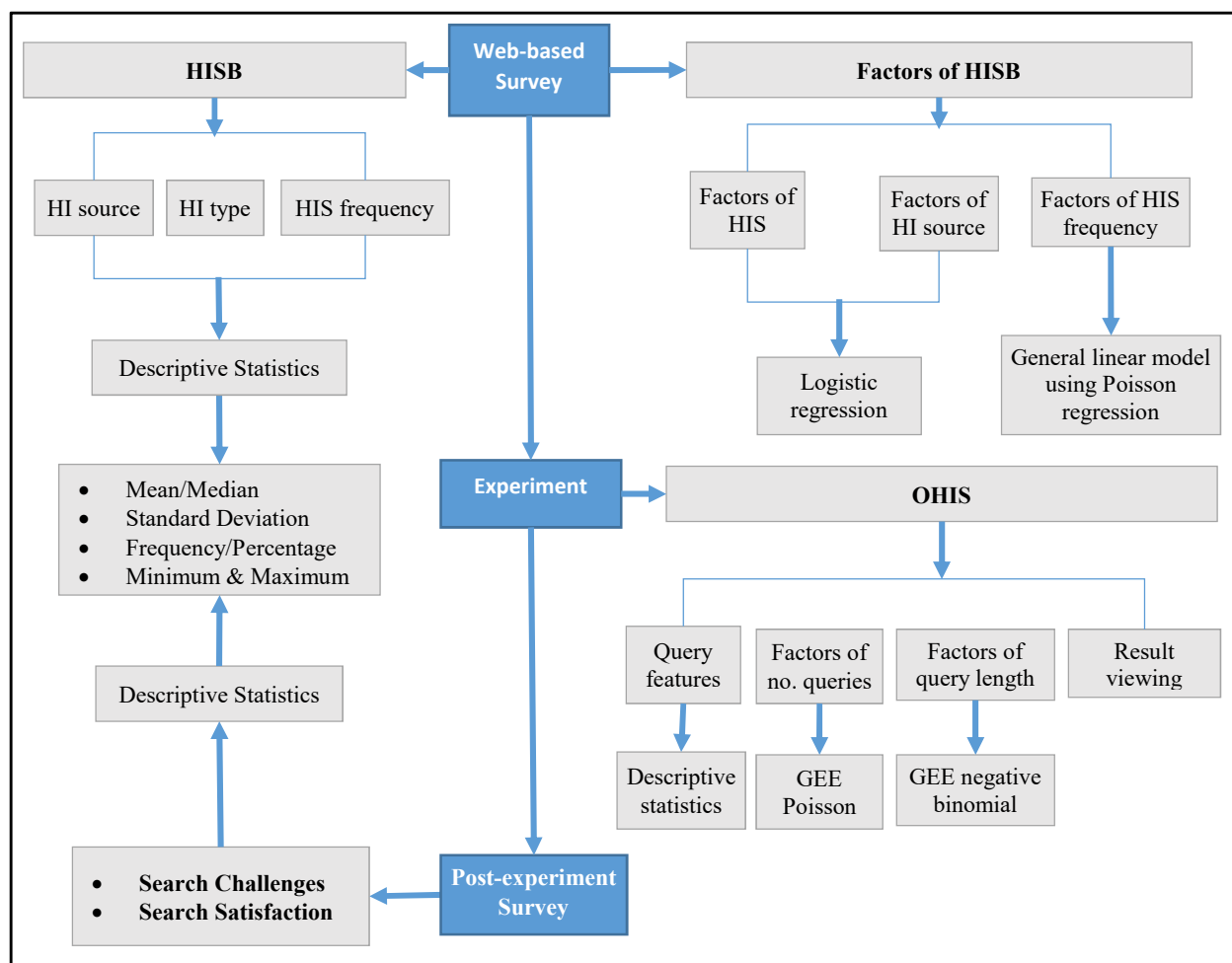
### 3.7. Data Management and Analysis

The data from the web-based survey is downloaded from Google Forms in the form of an excel sheet and then transferred to SPSS version 20 for analysis. In this research study, descriptive statistics are used to describe health information seeking and socio-demographical characteristics of the respondents. In addition, binary logistic regression method is used to identify factors associated with health information seeking and its sources. Binary logistic regression is used because it is an appropriate analysis method when the dependent variable is dichotomous or binary (Harrell, 2015). In this research study, the dependent variable (health information seeking) has two responses (yes/no). Stepwise logistic regression method is selected as it controls a confounding effect which occurs when a variable correlates with both dependent and independent variables which affect a statistical outcome (Sterne & Kirkwood, 2003). Odds ratio with 95% confidence

interval is computed to determine the strength of association between health information seeking and its predictors. All assumptions and model fitness are tested. A p-value of 0.05 is used as a cut-off value for statistically significant findings. Moreover, a general linear model using Poisson regression analysis method is used to identify factors associated with the frequency of health information seeking. Poisson is selected because the outcome variable (frequency of health information seeking) is count data.

In the experimental study, data from key logged software are transcribed for analysis. Transcribed data are entered using a pre-designed observation checklist in Google Form. The observational checklist is designed based on search activities (querying and result viewing behaviours). The pattern of query reformulations (specification, generalization, replacement, and parallel movement) developed by Rieh and Xie (2006) is used to explain changes in query reformulation process.

In order to identify the factors associated with the number of queries and query length in online health information searching, General Estimating Equation (GEE) using 'Poisson log-linear' and 'negative binomial with log link' models are applied, respectively. Poisson log-linear model is selected to identify factors associated with the number of queries because the mean is identical to the variance of the number of queries. However, since the variance is greater than the mean, which indicates over dispersion, negative binomial log link model is used to identify factors associated with query length. Figure 8 shows the diagrammatic representation of the data analysis procedure.



HI: health information, HIS = health information seeking, HISB = health information seeking behaviour, OHISB = online health information searching behaviour, GEE = General Estimating Equation

Figure 8: Schematic diagram of the data analysis procedure and methods

### 3.8. Validity and Reliability

In order to ensure the validity of the data, the web-based survey, simulated tasks, observation checklist and the post experiment survey are piloted with 15 students before the actual data collection begins. The questions for the web-based survey are validated and standardized after the following necessary corrections are made during the pilot study: 1) a new item is added to the survey to ask the reason for participants for not seeking health information, and 2) two questions are merged together as they are asking similar questions about challenges of health information seeking. Specifically, one question was asking about the challenges of health information seeking in any source



while another one was asking the challenges in online environments. The two questions were merged as one by accommodating challenges in all sources of health information. Furthermore, the validity of the web-based survey is assured by including the already validated questions from previous research studies. For example, the health literacy tool is taken from European Health Literacy Survey. The tool is already checked for its validity and reliability in the Asian (Asian Health Literacy Association (AHLA), 2013), Malaysian (Mohamad et al., 2014), and African (Amoah et al., 2017) contexts.

In order to reduce unanticipated disruptions, the experiments are carried out in computer laboratories so that internal validity of the experimental results can be increased. Furthermore, the experimental tasks are limited to four simulated tasks so that the variability of the experimental results as a result of fatigue can be reduced. If more simulated tasks were provided, there is possibility for participants to develop fatigue which could influence experimental results. Kelly (2009) recommends that simulated tasks should not last over an hour so as to reduce participants' fatigue. In this study, the average data collection time is identified as 45 minutes.

### **3.9. Chapter Summary**

In this research, two data collection methods are applied: Web-based survey and simulated situation using simulated tasks. The web-based survey is utilized to answer three research questions: 1) 'What are the sources of health information that university students from developing countries using?', 2) 'What type of health information that university students from developing countries are searching?', and 3) 'What are the factors affecting university students' health information seeking behaviour from developing countries?'. To answer the first two research questions, descriptive statistics are used. However, logistic regression analysis is applied to identify factors associated with HISB and its sources. Moreover, a general linear model using Poisson regression analysis method is used to identify factors associated with the frequency of health information seeking.

In the experimental study, simulated tasks are used to answer the fourth research question: 4) 'What searching behaviour do university students from developing countries

exhibit when they search for health information?’. In this case, General Estimating Equation (GEE) using ‘Poisson log-linear’ and ‘negative binomial with log link’ models are applied to identify the factors associated with the number of queries and query length in online health information searching, respectively. Finally, descriptive statistics are used to answer the fifth research question: 5) ‘What challenges do university students from developing countries face while searching for health information?’. Summary of the research methodology used is presented in Table 6.

Table 6: Summary of the Research Methodology

<b>Research Questions</b>	<b>Variables (Question number)</b>	<b>Data Collection Method</b>	<b>Data Analysis Method</b>
1.What sources of health information do university students from developing countries use?	<ul style="list-style-type: none"> <li>Source of health information (Part II: Q7)</li> </ul>	Web-based Survey	Descriptive statistics
2.What type of health information do university students from developing countries search for?	<ul style="list-style-type: none"> <li>Type of health information (Part II: Q12)</li> </ul>	Web-based Survey	Descriptive statistics
3.What are the factors affecting university students' health information seeking behaviour from developing countries?	<ul style="list-style-type: none"> <li>Demographic variables (Part I: Q1-10)</li> <li>Behavioural variables (Part IV: Q1-6)</li> <li>Psychological variables (Part V: Q1-3)</li> <li>Health-related variables (Part VI: Q1-4)</li> </ul>	Web-based Survey	Logistic Regression and general linear model using Poisson regression
4.What searching behaviour do university students from developing countries exhibit when they search for health information?	<ul style="list-style-type: none"> <li>Query behaviour</li> <li>Result viewing behaviour</li> <li>Search satisfaction</li> </ul>	Observation and post-experiment survey	Descriptive statistics and GEE using ‘Poisson log-linear’ and ‘negative binomial
5.What challenges do university students from developing countries face while searching for health information?	<ul style="list-style-type: none"> <li>Challenge of health information searching</li> </ul>	Post-experiment Survey	Descriptive statistics

## CHAPTER 4: RESULTS

In this section, results of this research study are presented according to the research objectives of this research study. First, the demographic information of the participants is provided. Then, results about the first research objective, which is ‘to examine the HISB of university students’, are presented in terms of the source and type of health information seeking. Next, findings of the factors associated with health information seeking are described, followed by the findings of the online health information searching behaviour of the participants. Finally, findings in relation to challenges faced by participants when searching for health information are described.

### 4.1. Demographic Information of Participants

A total of 901 undergraduate students (354 from Malaysia and 547 from Ethiopia) participated in the web-based survey. The percentage of male participants in both Malaysia (53.4%) and Ethiopia (52.3%) are slightly greater than female participants. The mean age of the participants from Malaysia in years is  $21.23 \pm 1.86$  SD with a minimum and maximum age of 17 and 34 years, respectively. Similarly, the mean age of the participants from Ethiopia in years is  $21.14 + 2.09$  SD with a minimum and maximum age of 18 and 37 years, respectively. Most participants in Malaysia are Malaysian (85.6%), and they are Chinese (49.7%) followed by Malay (31.1%) by ethnicity. Other nationalities from Malaysia include Indonesian, Sri Lankan, Bangladesh, Pakistani, Mauritian, Maldivian, Chinese, Omani, and Sudanese. However, all of the participants in Ethiopia are Ethiopian and the majorities are Amhara (77.1%) by ethnicity. Among Malaysian participants, the majorities are second-year students (41%) followed by first-year students (35%). Among Ethiopian participants, the majorities are first-year students (34.9%) followed by second-year students (24.5%). In Malaysia, the duration to complete Engineering course is 4 years while it is 3 years for Art and social science, Business and Information Technology. In Ethiopia, the duration to complete Engineering and Law is 5 years while it is 4 years for other fields of study. The majority of the participants from Malaysia and Ethiopia are currently living with their family (e.g. parents, sisters, and brothers). The demographic information of the participants is presented in Table 7.

Table 7: Socio-demographic characteristics of participants from Malaysia and Ethiopia

Variables	Malaysia (n = 354)			Ethiopia (n = 547)		
	Category	Freq.	%	Category	Freq.	%
Gender	Male	189	53.4	Male	286	52.3
	Female	165	46.6	Female	261	47.7
Age mean+SD(min,max) = 21.23 + 1.86, (17, 34)				21.14 + 2.09, (18, 37)		
BMI	Underweight (<18.5)	57	16.1	Underweight (<18.5)	132	24.1
	Normal (18.5-24.99)	227	64.1	Normal (18.5-24.99)	385	70.4
	Overweight(25-29.9)	50	14.1	Overweight(25-29.9)	25	4.6
	Obese ( $\geq 30$ )	20	5.7	Obese ( $\geq 30$ )	5	0.9
Year of study	1 <sup>st</sup> Year	124	35.0	1 <sup>st</sup> Year	191	34.9
	2 <sup>nd</sup> Year	145	41.0	2 <sup>nd</sup> Year	134	24.5
	3 <sup>rd</sup> Year	54	15.2	3 <sup>rd</sup> Year	120	21.9
	4 <sup>th</sup> Year	31	8.8	4 <sup>th</sup> Year	61	11.2
	5 <sup>th</sup> Year	NA	NA	5 <sup>th</sup> Year	41	7.5
Nationality	Malaysian	303	85.6	Ethiopian	547	100
	Indonesian	9	2.5			
	Sri Lankan	9	2.5			
	Bangladeshi	7	2.0			
	*Others	26	7.3			
Ethnicity	Chinese	176	49.7	Amhara	422	77.1
	Malay	110	31.1	Oromo	62	11.3
	Indian	23	6.5	Tigray	33	6.0
	**Others	45	12.7	*Others	30	5.5
Mother tongue	Chinese	133	37.6	Amharic	422	77.1
	Bahasa Malaysia	111	31.4	Oromiffa	62	11.3
	English	48	13.6	Tigrigna	33	6.0
	Tamil	12	3.4	**Others	30	5.5
	***Others	50	14.1			
Field of study	Art & Social Science	27	7.6	Social Science & Humanity	226	41.3
	Business	53	15.0	Business & Economics	32	5.9
	Engineering	72	20.3	Engineering	81	14.8
	IT	162	45.8	Education	9	1.6
	Science	34	9.6	Natural & computational science	196	35.8
	****Others	6	1.7	Law	3	0.5
Currently living with	Family (e.g. parents, sisters and brothers)	194	54.8	Family	336	61.4
	Relatives (e.g. aunt, uncle and cousin)	10	2.8	Relatives	30	5.5
	Friends	97	27.4	Friends	140	25.6
	Alone	53	15.0	Alone	41	7.5

\* Other nationalities in Malaysia include Bangladesh, Mauritian, Maldivian, Pakistani, Chinese, Omani and Sudanese

\*\* Other ethnicities in Malaysia include Indonesian, Javanese, Maldivian, Memon, Oman, Punjabi, Bengali, African Arab

\*\*\* Other mother tongue in Malaysia include Arabic, Bahasa Indonesia, Bengali, Cantonese, Creole, Dhivehi, Memon, Punjabi, Tamil

\*\*\*\* Other field of studies in Malaysia include Agriculture, Islamic studies, and Geology

+ Other ethnicities in Ethiopia include Afar, Gamo, Sidamic, Kembata, Wolayta, Gambella, Anuak, Konuso, Sheka, Gurage, Hammer

++ Other mother tongues in Ethiopia include Afarigna, Gamogna, Sidamegna, Kembategna, Wolaytegna, Gambella, Anuak, Konusogna, Sheka, Guragegna, Hammer

## 4.2. Health Information Seeking Behaviour

In this research study, the percentage of health information seeking among participants from Malaysia is 77.1% while it is 61.8% among participants from Ethiopia. Although the percentage of health information seeking is lower among participants from Ethiopia, health information seekers from Ethiopia seek health information more frequently than health information seekers from Malaysia. The average frequency of health information seeking among health information seekers from Ethiopia is twice per month, whereas that of health information seekers from Malaysia is once per month. About 23% and 38% of participants from Malaysia and Ethiopia have not sought health information, respectively. The main reason for not seeking health information are 1) uncertainty about where to look for information, 2) health information seeking takes too much time and effort to get relevant information, and 3) lack of time to attend television or radio health programs. Figure 9 presents the participants' reasons for not seeking health information.

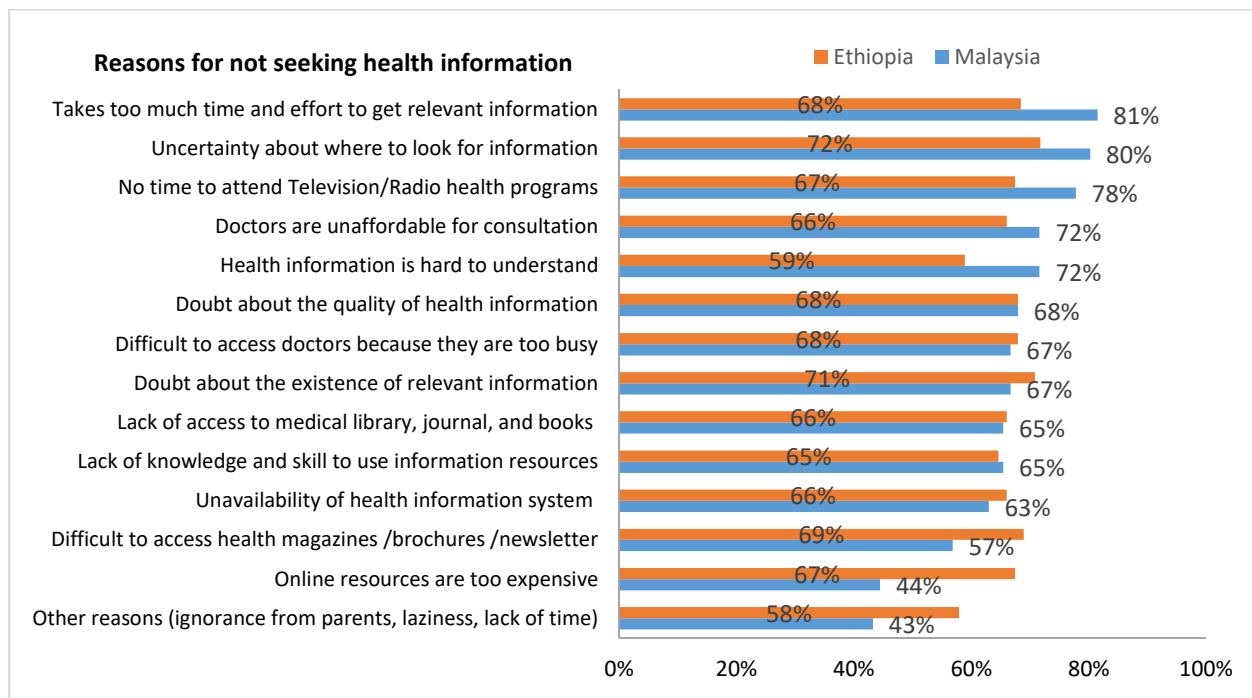


Figure 9: Participants' reasons for not seeking health information

#### 4.2.1. Source of Health Information

The Internet is the first choice for participants from Malaysia to look for health information. Amongst 273 health information seekers located in Malaysia, 82.8% prefer the Internet as their primary source of health information. In addition to the Internet, 15.8% and 13.9% of participants located in Malaysia choose doctors and family as primary sources of health information, respectively. Among their family members, participants located in Malaysia prefer to consult their mother for health information more than their father. Other sources of health information used by participants located in Malaysia include friends, books, brochures newspaper, magazines, and television or radio.

The majority of the participants from Ethiopia (41.4%) prefer television or radio as their primary source of health information. In addition, 36.4% and 34.6% of health information seekers located in Ethiopia prefer the Internet and their family or relatives as primary sources of health information, respectively. Participants located in Ethiopia also prefer doctors (23.4%), books (20.7%), and their friends (18.6%) as primary sources of health information, respectively. Similar to participants from Malaysia, participants located in Ethiopia prefer to consult their mother for health information more than their father. Figure 10 shows the preferred primary sources of health information among participants from Malaysia and Ethiopia.

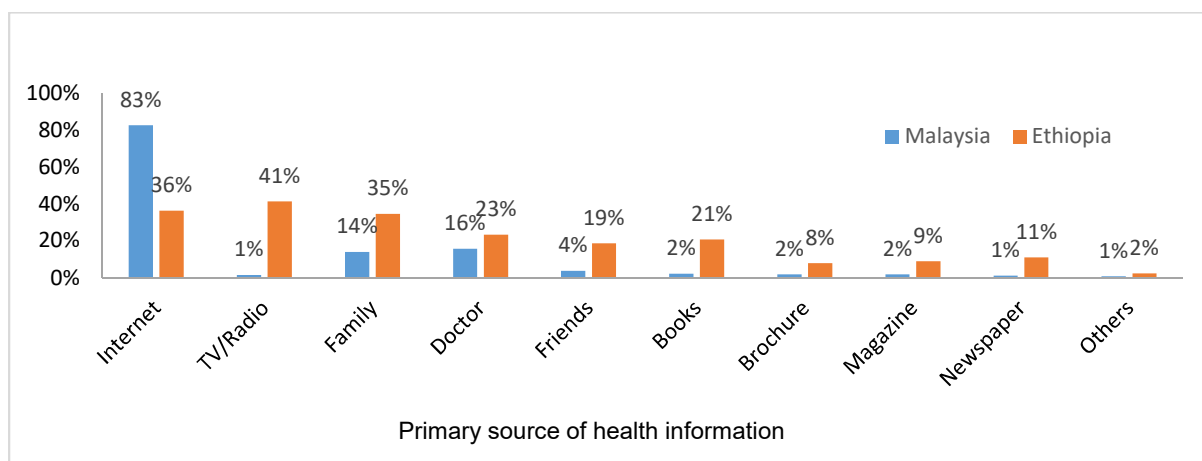
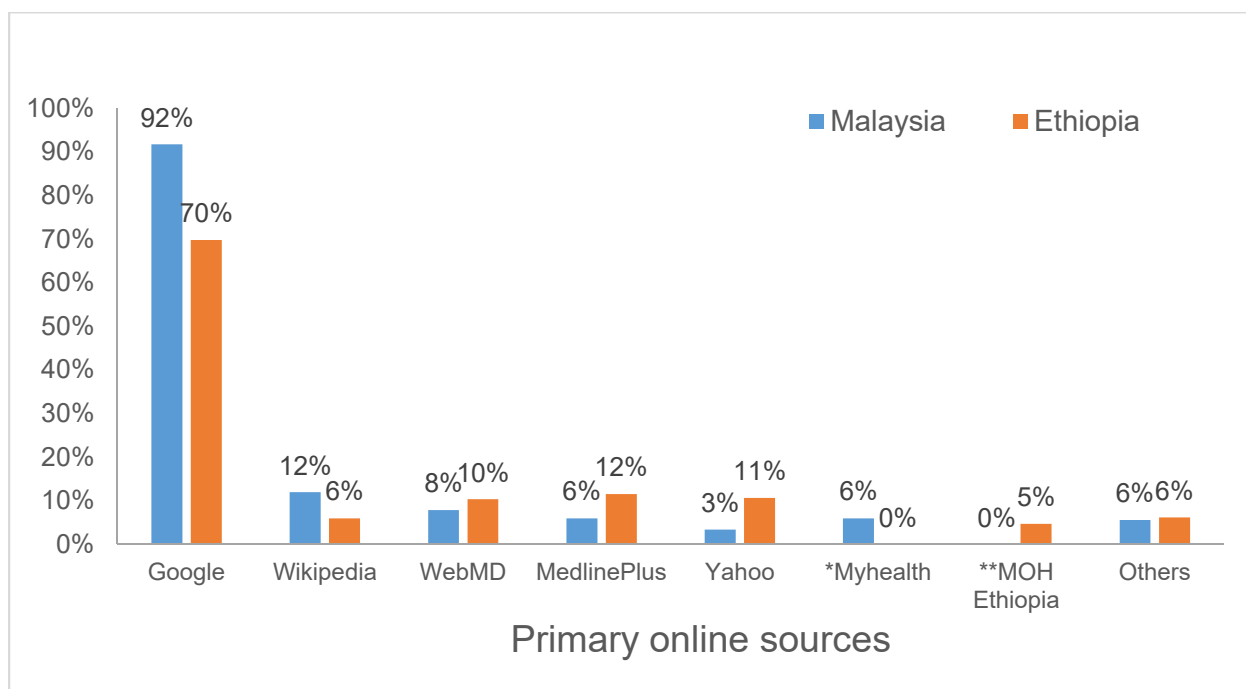


Figure 10: Primary source of health information amongst participants from Malaysia and Ethiopia

#### 4.2.1.1. Online Searching Tools Used

Google is found to be the most preferred search engine used by participants from Malaysia and Ethiopia to locate health information. Moreover, participants located in Malaysia use Wikipedia, whereas participants located in Ethiopia use MedlinePlus as their second preferred search engine to search for health information. Local sources such as Malaysian and Ethiopian health information portals are rarely used. The primary searching sources that participants located in Malaysia and Ethiopia use in their most recent time are presented in Figure 11.



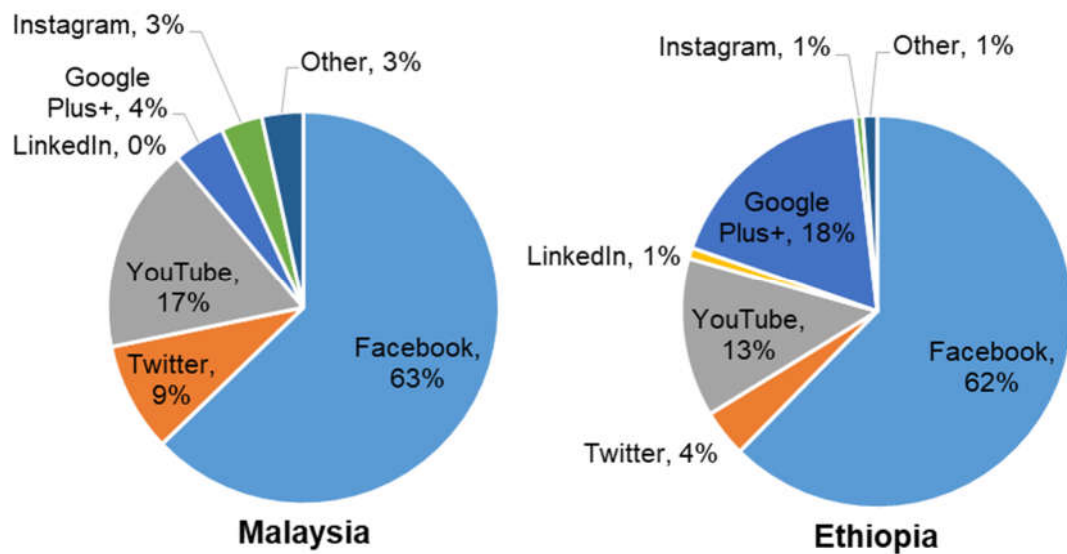
\* Myhealth is a web-based health information service controlled by the Ministry of Health Malaysia to provide health information  
\*\* MOH Ethiopia is a health information service owned by the Ministry of Health Ethiopia to provide health information services

Figure 11: Primary online searching source of health information among participants from Malaysia and Ethiopia

#### 4.2.1.2. Use of Social Media for Health Information

In this research, 58.2% of participants located in Malaysia and 47.5% of participants located in Ethiopia use social media to look for health information. Facebook (Figure 12) is the most frequent social networking site used amongst participants from both Malaysia (63%) and Ethiopia (62%). YouTube (17%) is the second popular social networking site

used to look for health information amongst participants located in Malaysia. However, Google Plus+ is found to be the second popular social networking site used to look for health information amongst participants located in Ethiopia. Other social networking sites used to look for health information amongst participants located in Malaysia and Ethiopia are Twitter, Google Plus+, Instagram, LinkedIn, WhatsApp, Tumblr, and Pinterest. Figure 12 presents the percentage of social networking sites usage to look for health information amongst participants from Malaysia (left) and Ethiopia (right). The activities performed by the participants on social media include sharing health information, watching health-related videos, participating in online forums about health information, and writing health information on online blogs. Figure 13 shows the activities of participants in social media.



\*Other social media include WhatsApp, Tumblr, and Pinterest

Figure 12: Use of social media to look for health information amongst participants located in Malaysia (left) and Ethiopia (right)



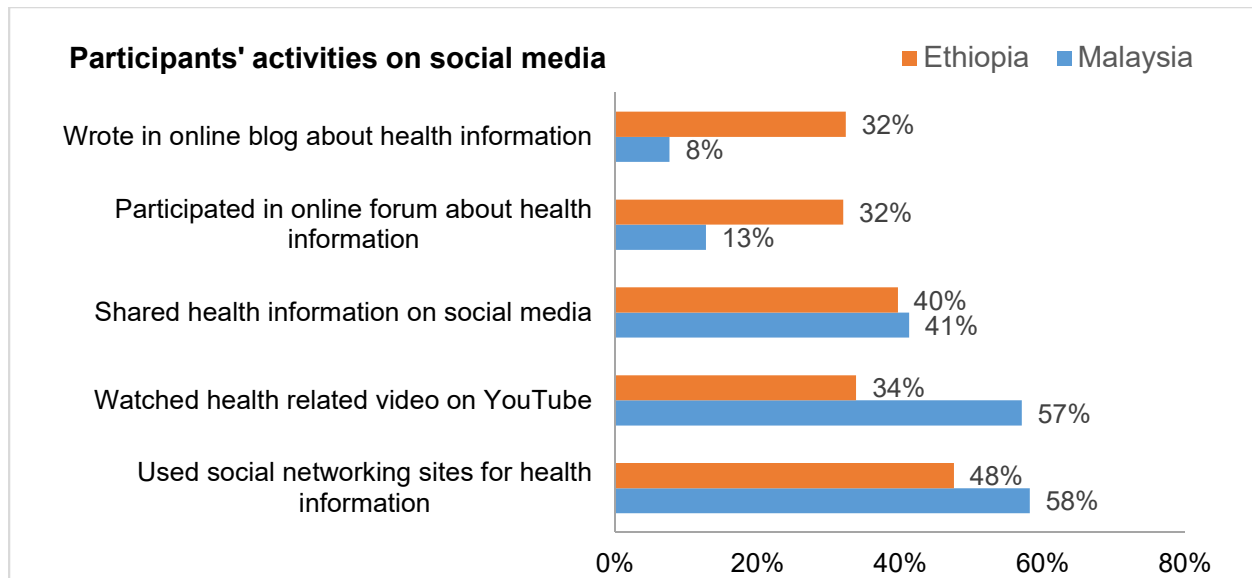
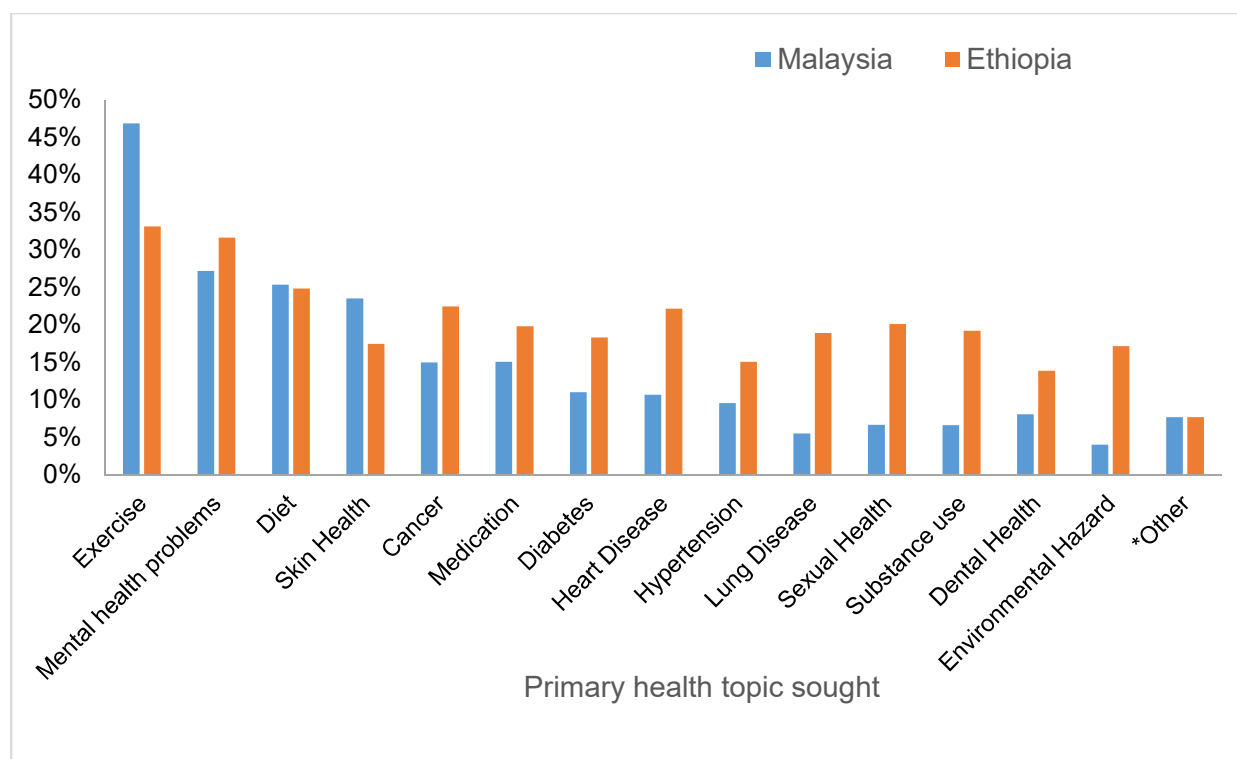


Figure 13: Participants' activities on social media

#### 4.2.2. Type of Health Information

In this study, the main reasons for participants located in Malaysia and Ethiopia to seek health information in most recent time are to seek information about fitness, diet, mental health problems, and skin health. Participants from both Malaysia and Ethiopia have sought wellness and illness-related information. A higher proportion of participants located in Malaysia have sought wellness-related information (e.g. exercise, skin health and diet) when compared with participants located in Ethiopia. However, the proportion is higher among participants located in Ethiopia with respect to illness-related information seeking (e.g. cancer, diabetes, heart disease) when compared with participants located in Malaysia. The list of health information that participants from Malaysia and Ethiopia sought is presented in Figure 14.



\* Other health topics include pregnancy and healthy lifestyle

Figure 14: Type of health information sought among participants from Malaysia and Ethiopia

### 4.3. Factors Associated with Health Information Seeking

Demographic characteristics (gender, age, body mass index, and mother tongue), behavioural (smoking, alcohol consumption, and physical activity), psychological (perceived susceptibility to and perceived severity of health problems, and health self-efficacy), and health-related (health status and health literacy) variables are used as independent variables to fit into the logistic regression model. Health literacy, perceived susceptibility to health problems, and alcohol consumption are found to have a significant association with health information seeking among participants located in Malaysia. With regard to health literacy, the odds of health information seeking are 10.6 and 2.7 times higher among participants from Malaysia who have excellent and sufficient health literacy level than those who have inadequate health literacy level with [AOR = 10.6; 95% CI (2.3, 48.8)] and [AOR = 2.5; 95% CI (1.0, 7.6)], respectively. Similarly, alcohol consumption is found to be associated with health information seeking among participants from Malaysia.

Participants from Malaysia who currently consume alcohol are 83% more likely to seek health information than those who do not currently consume alcohol [AOR = 1.83; 95% CI (1.05, 3.19)].

The third variable that has a significant association with health information seeking among participants from Malaysia is perceived susceptibility to a health problem. Participants from Malaysia who are not concerned at all about getting health problems are 69% [AOR = 0.31; 95% CI (0.13, 0.74)] less likely to seek health information than those who are very concerned. Similarly, participants from Malaysia who are concerned about getting health problems are 66% [AOR = 0.34; 95% CI (0.19, 0.62)] less likely to seek health information than those who are very concerned about getting health problems. Details of the multivariate logistic regression are presented in Table 8.

Table 8: Multivariate logistic regression showing factors affecting health information seeking among participants located in Malaysia

Variables	Category	Health Information Seeking		AOR (95% CI)	P-value
		Yes (%)	No (%)		
Gender	Male	140(51.5)	49(60.5)	1	
	Female	132(48.5)	32(39.5)	1.61(0.93, 2.78)	0.087
Health literacy	Inadequate	5(1.8)	5(6.2)	1	0.002
	Problematic	61(22.4)	31(38.3)	1.38(0.48, 3.98)	0.56
	Sufficient	162(59.6)	42(51.9)	2.74(1.0, 7.6)	0.05
	Excellent	44(16.2)	3(3.7)	10.63(2.31, 48.79)	0.002
Alcohol consumption	Currently consumer	110(40.4)	27(33.3)	1.83(1.05, 3.19)	0.034
	Currently not consumer	162(59.6)	54(66.7)	1	
Perceived susceptibility to a health problem	Very concerned	134(49.3)	22(27.2)	1	0.001
	Concerned	116(42.6)	47(58.0)	0.34(0.19, 0.62)	<0.0001
	Not concerned at all	22(8.1)	12(14.8)	0.31(0.13, 0.74)	0.008

A similar analysis is performed to identify factors associated with health information seeking among participants located in Ethiopia. Results indicate that none of the variables affects health information seeking. However, the dataset of Malaysia and Ethiopia have been merged to see whether there is a difference in health information seeking between participants from Malaysia and Ethiopia. Results indicate that there is a significant

difference between the participants located in Malaysia and Ethiopia. The odds of health information seeking among participants located in Malaysia is 1.73 times higher than participants located in Ethiopia [AOR = 1.73, 95% CI (1.25, 2.39)]. Other variables, such as alcohol consumption and health literacy level are also found to affect health information seeking in the merged dataset. The odds of health information seeking among participants from Malaysia and Ethiopia who currently consume alcohol are 1.33 times higher when compared to those who do not currently consume alcohol [AOR = 1.33, 95% CI (1.01, 1.84)]. In addition, participants from Malaysia and Ethiopia with inadequate [AOR = 0.23, 95% CI (0.09, 0.06)], problematic [AOR = 0.38, 95% CI (0.18, 0.81)], and sufficient [AOR = 0.45, 95% CI (0.22, 0.95)] health literacy level are 77%, 62%, and 55% less likely to look for health information when compared to those who have excellent health literacy level.

#### **4.3.1. Factors Associated with Source Of Health Information**

In this section, factors associated with specific sources of health information, such as the Internet, family, doctors, and television/radio, are identified. The four sources are selected because they are prevalent primary sources identified in this research study.

##### **4.3.1.1. Factors Associated with the Use of the Internet as Primary Source of Health Information**

Health literacy and perceived severity of health problems among participants located in Malaysia, and perceived health status among participants located in Ethiopia are found to affect the use of the Internet as a primary source of health information. Participants from Malaysia with inadequate, problematic, and sufficient health literacy level are 75%, 69%, and 60% less likely to use the Internet as a primary source of health information than participants with excellent health literacy level, respectively. Similarly, the odds of using the Internet as a primary source of health information are twice higher among participants from Malaysia who perceive their health problem would make them sick than those who perceive their health problem would not be a big deal [AOR = 2, 95% CI (1.04, 3.86)].

Perceived health status is the only variable found to affect the use of the Internet as a primary source of health information among participants located in Ethiopia. The odds of using the Internet as a primary source of health information are twice higher among participants from Ethiopia who perceive their health status as 'not good' than those who perceive their health status as good [AOR = 2.03, 95% CI (1.08, 3.81)].

#### **4.3.1.2. Factors Associated with the Use of Family as Primary Source of Health Information**

Variables affecting the use of family as a primary source of health information are gender among participants located in Malaysia and perceived susceptibility to health problems among participants located in Ethiopia. The odds of consulting family as a primary source of health information are 2.7 times higher among female from Malaysia than males [AOR = 2.7, 95% CI (1.26, 5.76)]. Likewise, participants located in Ethiopia who are concerned and very concerned to a health problem are 50% [AOR = 0.5, 95% CI (0.29, 0.89)] and 40% [AOR = 0.6, 95% CI (0.36, 0.99)] less likely to consult their family as primary source of health information than participants who are not concerned at all about a health problem, respectively.

#### **4.3.1.3. Factors Associated with the Use of Doctors as Primary Source of Health Information**

Body mass index (BMI), health literacy, and field of study are the variables identified to affect the use of doctors as a primary source of health information among participants located in Malaysia. However, perceived health status is found on the borderline to be a statistically significant factor affecting the use of doctors as a primary source of health information among participants located in Malaysia with a p-value of 0.06. However, none of the variables is found to affect the use of doctor for health information among participants located in Ethiopia.

As for BMI (defined in Appendix A), the odds of using doctors as a primary source of health information are 2.32 times [AOR = 2.32, 95% CI (1.05, 5.14)] higher among overweight (BMI > 29.9) participants from Malaysia than those whose BMI is normal (18.5 - 24.99). Similarly, the odds of using doctors as a primary source of health information are about three times higher among participants from Malaysia with adequate health literacy level than those with limited health literacy level [AOR = 2.79, 95% CI (1.12, 6.96)]. Field of study is the last variable identified as a factor affecting the use of doctors as a primary source of health information among participants located in Malaysia. Information technology and engineering students from Malaysia are 57% less likely to use the doctor as a primary source of health information [AOR = 0.43, 95% CI (0.22, 0.85)].

#### **4.3.1.4. Factors Associated with the Use of Television or Radio as Primary Source of Health Information**

Health literacy is the only variable identified as a factor affecting the use of television or radio as a primary source of health information among participants located in Ethiopia. Participants with limited health literacy level are 58% times more likely to use television or radio as a primary source of health information in Ethiopia [AOR = 1.58, 95% CI (1.06, 2.4)]. Table 9 shows the factors associated with the use of the Internet, family, doctors and television or radio as primary sources of health information.

Table 9: Multivariate logistic regression showing factors affecting the source of health information amongst participants from Malaysia and Ethiopia

Health information source	Variables	Category	AOR (95% CI), P-value	
			Malaysia	Ethiopia
Internet	Health literacy	Inadequate	0.25(0.06, 1.06), 0.061	
		Problematic	0.31(0.13, 0.71), 0.006	
		Sufficient	0.4(0.18, 0.88), 0.022	
		Excellent	1	
	Perceived severity of a health problem	Make me sick	2(1.04, 3.86), 0.039	
		Not a big deal	1	
	Perceived Health status	Feel healthy		1
		Less healthy		2.03(1.08, 3.81), 0.027
Family	Gender	Male	1	
		Female	2.7(1.26, 5.76), 0.011	
	Perceived susceptibility to a health problem	Very concerned		0.5(0.29, 0.89), 0.017
		Concerned		0.6(0.36, 1), 0.057
		Not concerned at all		1
Doctors	BMI	Normal (18.5-24.99)	1	
		Underweight (<18.5)	2.22(0.93, 5.33), 0.075	
		Overweight (>29.9)	2.32(1.05, 5.14), 0.038	
	Health literacy	Limited	1	
		Adequate	2.79(1.12, 6.96), 0.028	
	Perceived Health status	Feel healthy	2.59(0.96, 6.96), 0.06	
		Less healthy	1	
	Field of study	IT and Engineering	0.43(0.22, 0.85), 0.014	
		Others	1	
Television/ radio	Health literacy	Limited		1.58(1.06, 2.4), 0.034
		Adequate		1

1 = reference category

#### 4.3.2. Factors Associated with Frequency of Health Information Seeking

General linear model using Poisson regression analysis method is used to identify factors affecting the frequency of health information seeking amongst participants from Malaysia and Ethiopia. Results indicate that gender, smoking, and health literacy level are found to affect the frequency of health information seeking among participants from both Malaysia and Ethiopia. In addition, mother tongue, perceived severity of health problems, and perceived health status are found to affect the frequency of health information seeking amongst participants from Malaysia, whereas age and year of study among participants from Ethiopia.

Male participants from Malaysia and Ethiopia are 32% and 22% less likely to frequently seek health information than female participants, respectively. Regarding mother tongue, the odds of seeking health information frequently are 1.42 times higher among participants from Malaysia whose mother tongue is English than those whose mother tongue is not English [IRR = 1.42, 95% CI (1.11, 1.82)]. However, none of the participants whose mother tongue is English is found in Ethiopia. With regard to smoking, similar findings are obtained amongst participants from Malaysia and Ethiopia. The odds of seeking health information frequently are 1.3 times higher among current smokers from both Malaysia and Ethiopia than non-smokers.

Health literacy level is the other variable found to affect the frequency of health information seeking among participants from both Malaysia and Ethiopia. Participants from Malaysian with inadequate, problematic, and sufficient health literacy levels are 60%, 41%, and 33% less likely to frequently seek health information than participants with excellent health literacy level. However, participants from Ethiopia with inadequate health literacy level are 40% less likely to frequently seek health information when compared to participants with excellent health literacy level.

The odds of seeking health information among participants from Malaysia who perceive their health problems would make them sick and those who perceive their health status as healthy are 1.4 and 1.3 times higher than those who perceive their health problem as 'not a big deal' and those who perceive their health status as not healthy, respectively. Among participants from Ethiopia, the odds of seeking health information frequently are 1.04 times higher as the age of participants increases by a year [IRR = 1.04, 95% CI (1.01, 1.08)]. Similarly, the odds of seeking health information frequently are 2.08, 1.86, 1.85, and 1.55 times higher among participants from Ethiopia who are 1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup>, and 4<sup>th</sup>-year students, respectively, than participants from Ethiopia who are 5<sup>th</sup>-year students. Details of the Poisson regression analysis is presented in Table 10.



Table 10: Poisson regression analysis showing the factors affecting the frequency of health information seeking amongst participants from Malaysia and Ethiopia

Predictors	Category	Frequency of health information seeking					
		Malaysia			Ethiopia		
		Coeff	IRR (95% CI)	P-value	Coeff.	IRR (95% CI)	P-value
Gender	Male	-0.39	0.68(0.56, 0.81)	<b>&lt;0.0001</b>	-0.25	0.78(0.68, 0.89)	<b>&lt;0.0001</b>
	Female	1					
Mother Tongue	English	0.35	1.42(1.11, 1.82)	<b>0.005</b>	English = 0		
	*Other	1					
Smoking	Smoker	0.28	1.32(1.04, 1.68)	<b>0.021</b>	0.28	1.33(1.07, 1.64)	<b>0.009</b>
	Non-smoker	1			1		
Age		0.03	1.03(0.98, 1.08)	0.2	0.04	1.04(1.01, 1.08)	<b>0.026</b>
Year of study	Year 1	0.1	1.11(0.77, 1.6)	0.58	0.73	2.08(1.46, 2.96)	<b>&lt;0.0001</b>
	Year 2	0.17	1.19(0.83, 1.69)	0.34	0.62	1.86(1.32, 2.62)	<b>&lt;0.0001</b>
	Year 3	0.39	1.48(1.02, 2.16)	<b>0.042</b>	0.62	1.85(1.31, 2.61)	<b>&lt;0.0001</b>
	Year 4	1		0.079	0.44	1.55(1.06, 2.25)	<b>0.022</b>
	Year 5	-			1		<b>0.001</b>
Health literacy	Inadequate	-0.91	0.4(0.2, 0.8)	<b>0.009</b>	-0.51	0.6(0.39, 0.92)	<b>0.018</b>
	Problematic	-0.53	0.59(0.45, 0.77)	<b>&lt;0.0001</b>	-0.23	0.8(0.56, 1.15)	0.223
	Sufficient	-0.40	0.67(0.54, 0.83)	<b>&lt;0.0001</b>	-0.32	0.73(0.5, 1.05)	0.089
	Excellent	1		<b>&lt;0.0001</b>	1		<b>0.046</b>
Perceived severity of health problem	Make me sick	0.34	1.41(1.04, 1.9)	<b>0.024</b>	-0.04	0.96(0.82, 1.13)	0.63
	Not a big deal	1			1		
Perceived health status	Healthy	0.23	1.25(1.01, 1.55)	<b>0.039</b>	0.04	1.04(0.82, 1.31)	0.76
	Not healthy	1			1		

\* Other mother tongues include Arabic, Bahasa Indonesia, Bengali, Cantonese, Dhivehi, Hakka, Hokkien, Memon, Punjabi, Tamil  
1 Reference category  
P-values in bold show statistically significance (p-value < 0.05)

#### 4.4. Online Health Information Searching Behaviour

In this section, findings on online health information searching behaviour of participants from Malaysia are described. Only participants from Malaysia are included because of the ease of logistics. Hence, in this section, the demographic information is provided first. Then, results of query behaviour are presented. Finally, findings on result viewing behaviour are described.

##### 4.4.1. Demographic Information

A total of 176 participants from Malaysia participated in the experimental study. Each of the participants performed four simulated tasks and one personal task. Eighty-eight males and 88 females took part in the experiment (Table 11). The age of the participants ranges from 17 to 34 years, mean = 21.2, standard deviation = 2.2. The majority of the

participants are first year students (39.8%), and they are from engineering school (32.4%). From the total, 78.4% of the participants ever used the Internet to look for health information whereas 21.6% never used. The average year of experience of the participants in online health information searching is  $5 \pm 2.6$  SD. Participants are asked about their experience in searching the experimental tasks. From the total, about 74%, 61%, 72%, and 75% of the participants have responded that they search for dengue, smoking, obesity, and physical activity before, respectively.

Table 11: Socio-demographic characteristics of participants from Malaysia

Variables	Category	Frequency	Percentage
Gender	Male	88	50
	Female	88	50
	Total	176	100
Age	Min = 17, Max = 34, Mean = $21.2 \pm 2.21$ SD		
Field of study	Art & Social Science	22	12.5
	Business	47	26.7
	Engineering	57	32.4
	IT	21	11.9
	Science	29	16.5
	Total	176	100
Year of study	1 <sup>st</sup> Year	70	39.8
	2 <sup>nd</sup> Year	50	28.4
	3 <sup>rd</sup> Year	34	19.3
	4 <sup>th</sup> Year	22	12.5
	Total	176	100
Health information seeking from the Internet	Yes	138	78.4
	No	38	21.6
	Total	176	100
Years of experience in health information seeking from the Internet	Mean = 4.72, Median = 5, SD = 2.57		

Based on the analysis of the computer screen recordings, the following aspects of health information searching behaviour are examined: query behaviour (query terms, type of query, number of query, query length, and conceptual factors of query formulation), and result viewing behaviour (result clicked and source opened). These aspects of searching behaviours are used because they are important and basic elements in the process of

information searching, as they reflect tactical and conceptual changes that are experienced by users during searching (Belkin et al., 2003; Broussard & Zhang, 2013; Lopes & Ribeiro, 2015; Zhang, 2013; Zhang et al., 2012). Moreover, many users' studies have studied the querying and result viewing behaviours to improve information retrieval systems (Belkin et al., 2003; Broussard & Zhang, 2013; Lopes & Ribeiro, 2015; Zhang, 2013; Zhang et al., 2012). Investigating the aspects of searching behaviours mentioned above could lead to understanding users' interaction with health information retrieval system, which has the contribution to improve or design an effective, usable, and user-friendly system.

#### **4.4.2. Query Behaviour**

The analysis of query behaviour for the four simulated tasks and one personal task examined five aspects: query terms, type of queries, number of queries, query length, and factors associated with the number of queries and query length. The presentation of the results is for the overall search tasks (dengue, smoking, obesity, physical activity and personal task). This is done to obtain the overall picture instead of task-specific viewpoint which is likely to have variance and bias across search tasks. However, task-specific results are provided in tables.

##### **4.4.2.1. Query Terms**

The analysis of query terms identifies three types of query terms: 1) Keywords which signify a topic, and are related strongly to a topic (Abhishek & Hosanagar, 2007), 2) stop words which do not help to characterize a document, and 3) operators which are words or symbols that help to connect query terms. Participants use mostly keywords to formulate queries which represent participants' understanding of each search tasks. The majority (88.9%) of these terms are from the task descriptions. However, sometimes, participants use alternative words instead of using terms from task description. For example, a participant formulates a query of 'guidelines or recommendations on physical activity'. In this case, the keywords ('guidelines', 'recommendations', and 'physical activity') are taken from the task description of the physical activity. Another participant

formulates a query of 'moderate intensity aerobic exercise'. In this case, the keywords ('moderate', 'intensity', 'aerobic' and 'exercise') are not taken from the physical activity task description.

The use of stop words shows the tendency of participants to use natural language during searching. The participants also use Boolean operators, specifically, and, AND, and or, to connect query terms. Particularly, the term 'and' appears more frequently than AND, but as it is lowercase, it is arguable that most of them are intended to be a natural language expression rather than a Boolean operator. Although participants are provided with task description, some participants misspell the query terms, for example, dangue, symtoms, and simtoms. It is found that 4% of the queries issued have a spelling error. However, participants use the auto-correction method suggested by search engines for 28.2% of the queries. The use of health terms to formulate queries is found to be high across all tasks. Table 12 shows the basic query features stratified by tasks.

#### **4.4.2.2. Type of Query**

In this research study, three types of queries are identified: a question-based, a word-based, and a run-together words-based. A question-based query is a query that contains question words, such as what, how, when, where, which, and why (Jadhav, Andrews, et al., 2014; Jadhav, Sheth, et al., 2014). A word-based query contains a single word whereas that of a run together query contains words more than one. About 68% of the queries issued are run-together words whereas a question-based constitutes 22.4%. Examples of each type of query that participants issued are listed below.

1. Run-together words-based queries: healthy body weight, dengue fever Malaysia, cause and effect of dengue, treatment of smoking, major cause of cancer
2. Word-based queries: obesity, dengue, smoking, cigarettes
3. Question-based queries: what is dengue fever, how to have a healthy bodyweight, how to quit smoking, how to be physically attractive

Table 12: Basic query features across tasks

Query feature	Task 1 (Dengue)	Task 2 (Smoking)	Task 3 (Obesity)	Task 4 (Physical activity)	Task 5 (Personal Task)	All Tasks
<b>Use of stop words:</b> Yes	138	190	238	177	101	844
No	181	106	180	106	158	731
<b>Use of operators:</b> Yes	49	36	44	34	19	182
No	270	260	374	249	240	1393
<b>Use of health terms:</b> Yes	314	278	387	256	222	1457
No	5	18	31	27	37	118
<b>Source of query:</b> Task description	308	277	345	231	240	1401
Outside task description	11	19	73	52	19	174
<b>Spelling error:</b> Yes	16	12	19	8	10	65
No	301	283	398	274	247	1503
<b>Query type:</b> Question-based	43	93	75	81	61	353
Word-based	18	31	59	10	38	156
Run-together words	258	172	284	192	160	1066
<b>Total</b>	<b>319</b>	<b>296</b>	<b>418</b>	<b>283</b>	<b>259</b>	<b>1575</b>

#### 4.4.2.3. Number of Queries and Query Length

Participants issued a total of 1575 queries to perform five tasks, including their personal task. However, some participants skipped performing some of the simulated tasks and their personal task during the experiment as shown in Table 13. For example, 30 participants skipped performing their personal task during the experimentation. These are considered as missing during analysis.

The number of queries issued ranged from 1 to 10 across the tasks whereas the average number of queries is 1.92 for the simulated tasks and 1.77 for the personal task. The average query length identified in this study is 4.4 terms for the simulated tasks while it is 3.7 for the personal task. Table 13 shows the basic query features stratified by tasks.

Table 13: Number of queries and query length

Query feature	Task 1 (Dengue)	Task 2 (Smoking)	Task 3 (Obesity)	Task 4 (Physical activity)	Task 5 (Personal Task)	All Tasks
No. of participants	172	171	171	171	146	176
Missing	4	5	5	5	30	
No. of query: Mean	1.85	1.73	2.44	1.65	1.77	1.9
Median	1	1	2	1	1	1
SD	1.55	1.09	1.81	1.14	1.1	1.4
Min	1	1	1	1	1	1
Max	10	6	10	8	6	10
Query length: Mean	3.94	4.31	4.37	5.2	3.71	4.31
Median	3	4	4	5	3	4
SD	2.44	3.38	3.8	3.85	2.42	3.32
Min	1	1	1	1	1	1
Max	21	30	25	48	21	48

A paired-sample t-test is computed to compare the number of queries and query length between the simulated tasks and personal task. It is found that there is a significant difference in the number of queries between simulated tasks (mean = 1.96, SD = 0.99) and personal task (mean = 1.77, SD = 1.1);  $t(145) = 2.38$ ,  $p = 0.019$ . The number of queries is found to be higher in simulated tasks than a personal task. Similarly, a significant difference in query length is found between simulated tasks (mean = 4.4, SD = 2.5) and personal tasks (mean = 3.8, SD = 2.6);  $t(145) = 2.45$ ,  $p = 0.016$ , implying longer query length in simulated tasks than personal task. These could mean query behaviour demonstrated in the simulated task is different from the query behaviour demonstrated in the personal tasks.

#### 4.4.2.4. Factors Associated with the Number of Queries and Query Length

Generalized Estimating Equation (GEE) is used to examine the effect of demographic (gender, age, mother tongue, BMI, search experience, and Internet use experience), behavioural (physical activity, smoking, alcohol consumption), psychological (health self-efficacy and perceived susceptibility to and perceived severity of health problems), and health-related (health status and health literacy) factors on the number of queries and query length. The variables are selected because they have the potential influence on query formulation. For example, gender, age, health status (Kim, 2015; Wang et al.,

2013), health literacy (Ellis et al., 2012; Gutierrez et al., 2014), mother tongue (Matsveru, 2013), and Internet use experience (Escoffery et al., 2005; Janeice et al., 2013) are found to influence health information searching.

GEE is selected because it allows analysis for repeated measurements (Zeger & Liang, 1986). In this research study, the outcome variables (query length and number of queries) are measured repeatedly over the four simulated tasks and one personal task. Moreover, GEE can accommodate missing observations. In the GEE analysis, 'Poisson log-linear' model is used for the number of queries as its mean (1.9) and the variance (1.97) are almost identical whereas 'negative binomial with log link' model is used for the query length because its variance (11.01) is greater than the mean (4.31). Both Poisson log-linear and negative binomial regression models are appropriate models for count data (Hoffmann, 2016).

Out of all the variables entered to the regression models, mother tongue, physical activity, perceived task difficulty, and frequency of health information seeking are found to affect the number of queries across all the tasks. As shown in Table 14, the number of queries is 1.34 times higher among participants whose mother tongue is English than participants whose mother tongue is not English [AOR = 1.34, 95% CI (1.04, 1.72)]. Similarly, the number of queries is 1.2 times higher among participants who are physically active than physically inactive participants [AOR = 1.18, 95% CI (1.01, 1.38)]. Perceived task difficulty is the other variable found to affect the number of queries. The number of queries is 1.4 times higher among participants who perceive the tasks as easy than those who perceive the tasks not easy [IRR = 1.4, 95% CI (1.06, 1.85)]. The third variable found to affect the number of query is the frequency of health information seeking. It is found that a unit increase in the frequency of health information seeking will increase the number of queries by 0.03.

On the other hand, perceived susceptibility to a health problem, previous health information seeking, and health literacy are the factors found to be associated with the query length across all the tasks. Those participants who are concerned about health problems and those who have never sought health information have issued 1.29 and 1.23

times longer queries than their counterparts, respectively. Regarding health literacy level, participants with inadequate health literacy level are 40% less likely to issue longer query length when compared to those who have excellent health literacy level. The outputs of Poisson and negative binomial regression analysis are described in Table 14.

Table 14: Poisson and negative binomial regression analysis showing the factors associated with the number of queries and query length across all the tasks

Predictors	Category	Number of Queries			Query Length		
		Coeff.	IRR (95% CI)	p-value	Coeff.	IRR (95% CI)	p-value
Mother Tongue	English	0.29	1.34(1.04, 1.7)	<b>0.023</b>	-0.02	0.98(0.9, 1.12)	0.75
	*Other	1			1		
Physical activity	Active	0.17	1.18(1.01, 1.4)	<b>0.038</b>	0.07	1.08(0.93, 1.24)	0.32
	Inactive	1			1		
Task difficulty	Easy	0.34	1.4(1.06, 1.85)	<b>0.018</b>	0.07	1.08(0.68, 1.71)	0.76
	Not easy	1			1		
Frequency of health information seeking		0.03	1.03(1, 1.06)	<b>0.043</b>	-0.02	0.99(0.96, 1)	0.56
Tasks	Dengue	0.06	1.07(0.9, 1.23)	0.37	0.05	1.06(0.95, 1.2)	0.31
	Smoking	-0.01	0.99(0.89, 1.1)	0.86	0.15	1.16(1.02, 1.3)	<b>0.015</b>
	Obesity	0.34	1.4(1.24, 1.59)	<b>&lt;0.0001</b>	0.15	1.16(1.03, 1.3)	<b>0.015</b>
	Physical Activity	-0.05	0.95(0.84, 1.1)	<b>0.046</b>	0.31	1.36(1.22, 1.5)	<b>&lt;0.0001</b>
	Personal Task	1		<b>&lt;0.0001</b>	1		<b>&lt;0.0001</b>
Health literacy	Inadequate	0.04	1.04(0.7, 1.52)	0.084	-0.52	0.6(0.42, 0.84)	<b>0.003</b>
	Problematic	0.24	1.28(1.01, 1.6)	0.044	-0.12	0.89(0.66, 1.2)	0.424
	Sufficient	0.17	1.18(0.97, 1.4)	0.09	-0.16	0.85(0.63, 1.2)	0.297
	Excellent	1		0.214	1		<b>0.015</b>
Perceived susceptibility to health problem	Concerned	0.03	1.03(0.85, 1.3)	0.76	0.25	1.29(1.1, 1.5)	<b>0.001</b>
	Not concerned	1			1		
Health information seeking	Yes	1			1		
	No	0.1	1.11(0.89, 1.4)	0.36	0.20	1.23(1.04, 1.44)	<b>0.014</b>

\* Other mother tongues include Arabic, Bahasa Indonesia, Bengali, Cantonese, Dhivehi, Hakka, Hokkien, Memon, Punjabi, Tamil

1 Reference category

#### 4.4.3. Query Reformulation

Query reformulation is the modification of a previous query for a given search task. For example, a user may search for 'dengue fever', but modify the original query to 'sign and symptom of dengue fever' if the user is unsatisfied with the results from the initial query. In this research study, it is found that 54% of the queries are reformulated queries across all the tasks. The average query reformulation is  $2.6 \pm 1.3$  SD for the simulated tasks,  $1.8 \pm 0.99$  SD for the personal task. The maximum number of query reformulation is 9 for the



simulated tasks and 5 for the personal task. A paired-sample t-test shows that there is significant difference between the simulated tasks (mean = 2.6, SD = 1.3) and the personal task (mean = 1.8, SD = 0.99) with regard to number of query reformulations;  $t(16) = 2.8$ ,  $p = 0.012$ . The number of query reformulation is found to be higher in simulated tasks than a personal task.

This research study identifies two types of query reformulations across all the tasks: reformulation that leads to conceptual changes (86.2%) and reformulation that changes the form of query (13.8%). In the conceptual changes, four reformulation patterns are observed: 1) Specification – where participants specify the query by adding or replacing terms; 2) generalization – where participants generalize query by deleting or replacing terms; 3) parallel movement – where participants replace terms with alternative words having similar meaning; and 4) switch to new concepts – where participants reformulate query without overlapping with the previous query. The most frequent reformulation pattern is ‘switching to new concepts’ (55.3%) followed by ‘specification’ (22.3%). With respect to query reformulation that changes the form of a query, three reformulation patterns are identified across all the tasks: 1) reform to question, 2) change term sequence, and 3) spelling correction. The most frequent reformulation that changes the form of the query is ‘reform to question’ followed by ‘spelling correction’. Table 15 shows the pattern of query reformulations.

Table 15: Query reformulation pattern

Query reformulation	Number of query reformulation across tasks						Example
	Task 1	Task 2	Task 3	Task 4	Personal Task	Total	
Conceptual changes							
Specification	36	30	46	17	37	166	exercise→ exercise frequency
Generalization	10	8	9	8	4	39	traditional treatment dengue→ dengue treatment
Parallel movement	14	26	39	35	13	127	how dengue occurs→ how you get dengue
Switch to new concept	84	63	154	50	60	411	how to prevent duegue→ signs and symptoms of dengue
Change of the form of query							
Reform to question	18	24	28	13	15	98	BMI→how to calculate BMI
Change the terms sequence	0	1	1	1	2	5	how can I be considered a physically fit person→ how can be physically fit
spelling correction	2	2	8	1	2	15	weigght loss→ weight loss

#### 4.4.4. Result Viewing Behaviour

Out of the 1575 queries issued to search engines, 86.2% of them are clicked and viewed. Most of the results that are not clicked are those whose queries are with a spelling error, taken outside task description, and keyword-based queries. However, most of the results clicked are those ranked high on the result list. The analysis shows that 42% of the accessed results are the first on the result list, 20% are the second, 19% are the third, and 8% are the fourth. The remaining (11%) ranged from the fifth to the tenth. None of the participants clicked results beyond the first page.

The most frequent search engine used is Google (96.4%). The other search engine used are Bing (1.2%), MedlinePlus (0.4%), YouTube (0.3%), CDC.gov (0.3%), Reddit (0.2%), and Wikipedia (0.2%). In total, participants have visited 2449 results when performing all the search tasks. These results come from 504 unique websites. Among these websites, health-specific websites, such as webmd.com (10.5%), cdc.gov (5.9%), nhs.uk (4.7%), nhlbi.nih.gov (3.7%), who.int (3.6%), and mayoclinic.org (3.1%) are found to be the most

frequently visited sites. Wikipedia (3%), which is a general purpose site, appeared to be one the most popular site visited by the participants when performing all the tasks. The top ten most frequently visited sites are listed in Table 16.

Table 16: Top ten most websites visited by participants

	<b>Visited Websites</b>	<b>Frequency</b>	<b>Percentage</b>
1	www.webmd.com	257	10.5
2	www.cdc.gov	144	5.9
3	www.nhs.uk	115	4.7
4	www.nhlbi.nih.gov	91	3.7
5	www.who.int	88	3.6
6	www.mayoclinic.org	77	3.1
7	www.wikipedia.org	73	3.0
8	www.medicinenet.com	66	2.7
9	www.helpguide.org	56	2.3
10	www.smokefree.gov	50	2.0

#### 4.5. Challenges of Health Information Searching

Although most of the participants (79%) are happy with the overall search experience, they come across challenges and barriers when performing all the searching tasks. Nine questions are given to participants to rate (agree/disagree) for possible challenges of performing all the search tasks (Figure 15). The most frequent challenges faced by participants are: 1) time consuming to wade through volumes of data to find required information, 2) difficulty in judging the quality of health information, 3) obtaining excessive and irrelevant information, 4) low credibility of the information sources, and 5) not able to obtain information immediately after clicking the result. Other challenges and barriers that the participants faced when performing the personal and simulated tasks are difficulties to formulate search tasks to query, difficulty to understand the information, and inability to find required information. In addition, an open-ended question is given to participants to indicate other challenges of performing all the search tasks. Participants indicate that they face difficulty in judging the reliability of the information obtained, difficulty to understand health terms, and difficulty in getting specific information in relation to their information need.

Similarly, participants are asked in the online survey about the challenges and barriers of health information seeking from the Internet. About 58% of participants responded that they face difficulties in judging the quality of information when seeking health information online previously. Moreover, 46.8% of the participants expressed in the online survey that health information seeking from the Internet is time-consuming. Too much and irrelevant search results are the other barriers encountered by 46.4% of the participants when they seek health information previously. Figure 15 lists the challenges and barriers faced by participants when performing all the search tasks.

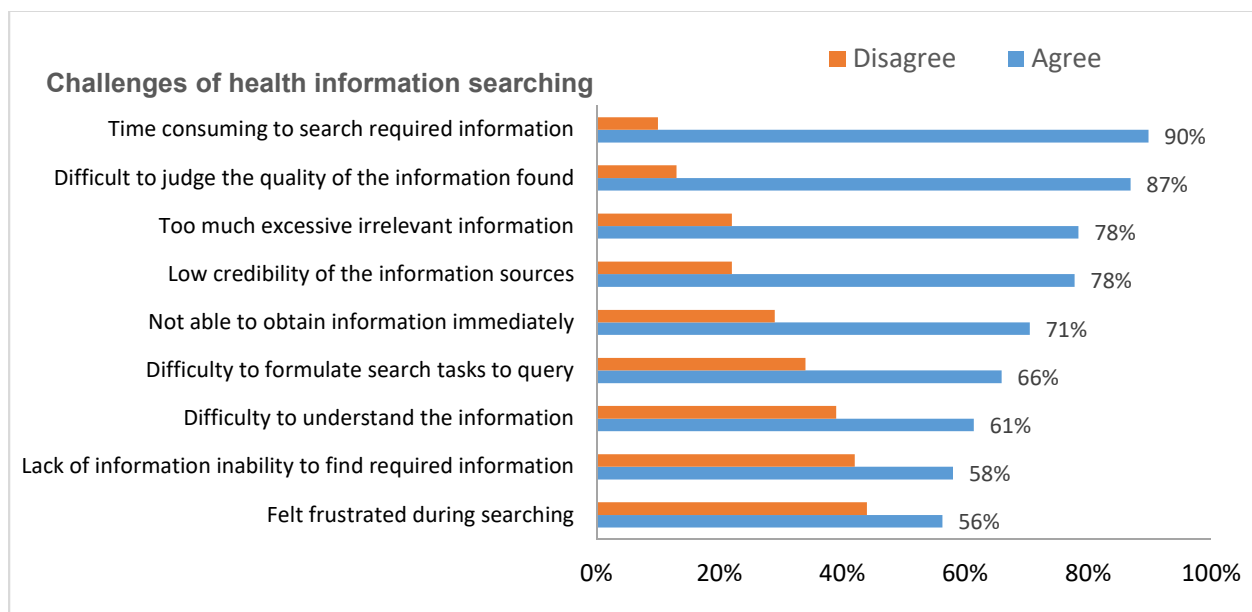


Figure 15: Challenges faced by participants when performing searching tasks

## 4.6. Chapter Summary

This section provides summary results from the individual sections of this chapter, such as health information seeking, the source of health information, the type of health information sought, factors associated with health information seeking, and online health information searching behaviour. Table 17 shows the summary results of this chapter.

Table 17: Summary results of health information seeking and searching behaviours

Result Sections	Malaysia	Ethiopia
Health information seeking	77.1%	61.8%
Primary source of health information	Internet	Television or Radio
Use of social media for health information	58.2%	47.5%
Type of health information sought (top 5)	Exercise, mental health problems, diet, skin health, and cancer	Exercise, mental health problems, diet, cancer, and heart disease
Factors associated with health information seeking	Health literacy, perceived susceptibility to health problems, and alcohol consumption	None
Factors associated with the use of the Internet as primary source of health information	Health literacy and perceived severity of health problems	Perceived health status
Factors associated with the use of the family as primary source of health information	Gender	Perceived susceptibility to health problems
Factors associated with the use of the doctors as primary source of health information	Body mass index (BMI), health literacy, and field of study	None
Factors associated with the use of the television or radio as primary source of health information	None	Health literacy
Factors associated with the frequency of health information seeking	Gender, smoking, health literacy, mother tongue, perceived severity of health problems and perceived health status	Gender, smoking, health literacy, age and year of study
Average number of queries	1.92 queries	NA
Average query length	4.3 terms per query	NA
Use of health terms	92.5% of participants used health terms	NA
Misspellings	4%	NA
Factors associated with the number of queries	Mother tongue, physical activity, perceived task difficulty, and frequency of health information seeking	NA
Factors associated with the query length	perceived susceptibility to a health problem, previous health information seeking, and health literacy	NA
Query reformulation that leads to conceptual changes	Switch to new concepts (55.3) Specification (22.3%) Parallel movement (17.1%) Generalization (5.3%)	NA
The most frequently visited sites	1. webmd.com (10.5%) 2. cdc.gov (5.9%) 3. nhs.uk (4.7) 4. nhlbi.nih.gov (3.7%) 5. who.int (3.6%) 6. mayoclinic.org (3.1%) 7. Wikipedia (3%)	NA

The most frequent challenges (Top 5)	<ol style="list-style-type: none"> <li>1. time-consuming to wade through volumes of data to find required information</li> <li>2. difficulties in judging the quality of health information</li> <li>3. excessive and irrelevant information</li> <li>4. low credibility of the information sources</li> <li>5. not able to obtain information immediately</li> </ol>	NA
---	---	----

In Table 17, the word 'None' is to indicate that none of the variables is identified to be the factors affecting health information seeking among Ethiopian participants, the use of the doctors as primary source of health information among Ethiopian participants, and the use of the television or radio as primary source of health information among Malaysian participants. Moreover, 'NA' indicates that the results are not available in Ethiopia because the experiment study is only conducted among Malaysian participants. The percentages in the first row shows the proportion of university students from Malaysia (77.1%) and Ethiopia (61.8%) that have sought health information. Similarly, the percentages in the third row indicates that 58.2% and 47.5% of university students from Malaysia and Ethiopia have used social media for health information, respectively.

## **CHAPTER 5: DISCUSSION AND RESEARCH CONTRIBUTION**

The purpose of this chapter is to summarize and discuss the empirical findings reported in Chapter 4. First, findings from the online survey are discussed in relation to health information seeking behaviour and its associated factors. Then, findings from the experiment with regard to the online health information searching behaviour are discussed. Finally, a discussion of theoretical and practical contributions is given.

### **5.1. Health Information Seeking Behaviour**

Limited research studies from developing countries have investigated health information seeking behaviour. These studies have reported the magnitude of health information seeking as low, for example, 46% in Malaysia (Mohd-Nor et al., 2013) and 59% in China (Yuli et al., 2012). However, this research study identifies a relatively higher proportion of health information seekers both from Malaysia (77.1%) and Ethiopia (61.8%) when compared with previous studies in Malaysia and China. The discrepancy could be due to differences in the study period and participants. The above studies in Malaysia and China were conducted in 2012 and 2013, respectively when the Internet penetration rates were lower (42.3% in China and 65.8% in Malaysia) as compared to the Internet penetration rate in 2016 (53.2% in China and 78.8% in Malaysia) (The World Bank, 2018a). Better Internet penetration rate could increase interest to search for health information as it gives ease and prompt access (Cutilli, 2010). Moreover, the participants in the research studies in Malaysia and China are a rural community and the general population, respectively who are different from university students. Due to their university education, university students are expected to have good Internet skills. Their Internet skills and the availability of the Internet in university campuses can motivate university students to look for health information when compared to rural community and the general population (Janeice et al., 2013).

This research study finds more health information seekers among Malaysian participants (77.1%) than Ethiopian participants (61.8%). The higher percentage of health information seeking in Malaysia could be due to the higher literacy rate (The World Bank, 2018b), which could have the potential to increase health awareness (Beiser & Stewart, 2005).

Moreover, the Internet penetration rate in Malaysia (78.8% in 2016) is higher than in Ethiopia (15.4% in 2016) (The World Bank, 2018a), which has a significant contribution to increasing access to health information (Anker et al., 2011; Tonsaker et al., 2014). However, there are still participants from Malaysia (23%) and Ethiopia (38%) who have not sought health information. The main reason is that the participants are uncertain about where to look for health information. Hence, there is a need to create awareness and promote national and international health-specific information retrieval systems for the university population.

In terms of the frequency of health information seeking, Ethiopian participants seek two times more frequently than Malaysian participants. The possible explanation for Ethiopian seekers to look for health information more frequently than Malaysian seekers could be due to the higher morbidity and mortality rates (World Health Organization, 2015d) which could motivate Ethiopians to look for more frequently.

#### **5.1.1. Source of Health Information**

Previous research studies from developing countries have reported traditional media such as television and radio, and personal communication such as communication with doctors, family, and friends as the most common sources of health information (Garcia-Cosavalente et al., 2010; Gavgani et al., 2013; Mohd-Nor et al., 2013; Muhammed & D'Souza, 2014; Prasad et al., 2013). In this study, television and radio remain the first choice of the majority of Ethiopian participants (41.4%). The reason for Ethiopians to prefer television or radio as the main source of health information could be because it is easily accessed. Furthermore, Ethiopians may have limited access to the Internet outside the university campus as an alternative source of health information. The Internet penetration rate in Ethiopia is 15.4% in 2016 which is very low (The World Bank, 2018a). However, the Internet is found to be the primary source of health information among Malaysian participants (82.8%). This could be due to the higher Internet penetration rate (78.8% in 2016) in Malaysia (The World Bank, 2018a), which enables easy and prompt access to health information (Cutilli, 2010). In addition, 35% of Ethiopians and 14% of Malaysians prefer their family as the primary source of health information. Among their family members, both Malaysians and Ethiopians prefer to consult their mother for health



information more than their father. This is because the mother seems to have a special attachment to young adults (Markiewicz et al., 2006).

In summary, the Internet remains the preferred source of health information as a significant proportion of university students from both Malaysia (82.8%) and Ethiopia (36%) are using it although the majority of Ethiopians (41%) prefer television or radio. The good Internet skill which is cultivated through university education (Janeice et al., 2013) seems to motivate university students to prefer the Internet as a preferred source of health information.

#### **5.1.1.1. Use of Social Media for Health Information**

Social media as a source of health information has a great implication for health communication as it provides a platform for people to communicate and share information (Parvanta et al., 2011). Limited research studies from developing countries show low usage of social media for health information. For example, a study in Iran reports that only 8.2% of public library users use social media for health information (Gavgani et al., 2013). However, in this study, relatively higher proportions of university students from Malaysia (58.2%) and Ethiopia (47.5%) have used social media for health information. The difference could be due to the age difference of the participants in the two studies. In this research study, the mean age of Malaysian and Ethiopian university students is 21 years which is the most active age group to use the Internet and social media (Malaysian Communications and Multimedia Commission, 2015). The higher proportion of social media use for health information indicates that social media are serving as an important platform to spread health messages. Hence, health promoters could use social media as a potential tool for health promotion activities.

#### **5.1.2. Type of Health Information**

Weaver et al. (2010) have classified the types of health information into two: wellness (preventive purposes) and illness (curative purposes) information. The wellness information can include personal and public hygiene, environmental protection, proper diet, and adequate exercise, whereas the illness information includes information about

disease or health problem control and treatment. The classification of wellness and illness information can help to determine distinct health motives of health information seeking (Weaver et al., 2010).

Previous research studies report that people from developing countries are mostly illness-related information seekers (Mohd-Nor et al., 2013; Muhammed & D'Souza, 2014; Prasad et al., 2013; Sharif et al., 2015). This could be due to high morbidity and mortality rates (World Health Organization, 2015d) that could motivate people in developing countries to look for illness-related information. However, this research study identifies university students from both Malaysia and Ethiopia as both illness- (e.g. cancer, diabetes, and heart disease) and wellness-related (e.g. exercise, diet, and environmental hazard) information seekers. The reason for university students to pay attention to wellness-related information in addition to illness-related information could be due to their high level of health awareness (Ansari & Stock, 2010; Mehta & Bhat, 2015) which could be cultivated through their formal education in university campuses.

## **5.2. Factors Associated with Health Information Seeking**

In this section, the underlying factors associated with health information seeking, the source of health information, and the frequency of health information seeking are discussed.

Among all the variables entered to the logistic regression model, health literacy, perceived susceptibility to health problems, and alcohol consumption are found to have a significant association with health information seeking among Malaysian participants. Higher health literacy level, alcohol consumption, and concerning about getting health problems contribute to health information seeking. Previous research studies support that an adequate level of health literacy engages individuals in health information seeking (Ellis et al., 2012; Eriksson-Backa et al., 2012; Gutierrez et al., 2014). Similarly, people tend to look for health information when they consider themselves more susceptible to health problems (Ahadzadeh et al., 2015; Mills & Todorova, 2016). With regard to alcohol consumption, the reason why alcohol consumers are more likely to look for health information could be due to their higher perceived susceptibility (90.4%) to and severity

(87.9%) of health problems, which are identified in this study. A research study identifies a similar finding that alcohol consumers are more likely to seek health information when compared to those who never consume alcohol (Wang et al., 2013).

#### **5.2.1. Factors Associated with the Source of Health Information**

Among all the sources, the factors are identified for the choice of the Internet, family, doctors, and television or radio for health information. In this research, the four sources are selected because they are identified as primary sources of health information.

Higher health literacy level and higher perceived severity of health problems among Malaysian participants and poor perceived health status among Ethiopian participants have a positive effect on the choice of the Internet as a primary source of health information. Previous research studies have reported that people with low health literacy level show little or no engagement in health information seeking (Eriksson-Backa et al., 2012; Gutierrez et al., 2014) because low levels of health literacy is associated with difficulties in expressing health problems (Manganello, 2008) which could be an obstacle to look for health information. Their higher health literacy level could help participants to express their health information on the Internet easily as it allows easy and prompt access to health information (Bratucu et al., 2014; Cutilli, 2010). Hence, it is important to increase the health literacy level of university students through health education in university campus so that the students could use the Internet for health information.

The odds of using the Internet as a primary source of health information are twice higher among Malaysian participants who perceive their health problems would make them sick when compared to those who perceive their health problem would not be a big deal. This finding is supported by the result of another study which is done in Malaysia (Ahadzadeh et al., 2015) which indicates that perceived severity of health problem has a positive effect on Internet use for health information. Similarly, the odds of using the Internet as a primary source of health information are twice higher among Ethiopian participants who perceive their health status as 'not good' than those who perceive their health status as good. Their health problem seems to motivate participants to look for health information. Previous

research studies indicate that having health problems are associated with frequent health information seeking (Kim, 2015; Wang et al., 2013; Weaver et al., 2010).

Gender among Malaysian participants and 'perceived susceptibility to health problems' among Ethiopian participants are found to affect the use of family as a primary source of health information. Being female can lead to the use of family as a primary source of health information in Malaysia. The special attachment females have to their parents more than males (David & Lyons-Ruth, 2005; Samuolis et al., 2001) could motivate females to choose their family as a primary source of health information. Likewise, Ethiopian participants who are concerned and very concerned to a health problem are 50% and 40% less likely to consult their family as a primary source of health information when compared to Ethiopian participants who are not concerned at all to a health problem, respectively. This could be because people are likely to primarily consult health providers when they are concerned about health problems as health providers believed to have a better solution for their health problems. Research studies from developing countries indicate doctors as a preferred source of health information (Garcia-Cosavalente et al., 2010; Muhammed & D'Souza, 2014) due to their trustworthiness for health information (Hesse et al., 2005; Oh et al., 2012).

Factors linked to using doctors as a primary source of health information among Malaysian participants are body mass index (BMI), health literacy, and field of study. None of the variables is found to affect the use of doctor for health information among Ethiopian participants. The reason could be the low percentage (14%) of Ethiopian participants to use doctors as the primary source of health information. In this research, overweight Malaysians are likely to consult doctors for health information due to their higher healthcare utilization when compared to Malaysians whose BMI classification is normal. Research studies have identified higher healthcare utilization among overweight individuals (Peytremann-Bridevaux & Santos-Eggimann, 2007; Suehs et al., 2017). Similarly, participants with adequate health literacy level are more likely to use doctors as the primary source of health information. This could be potentially because adequate health literacy level enables people to get motivated to access, understand, and utilize health information (World Health Organization, 2009b). Furthermore, adequate health

literacy level enables people to easily and comfortably communicate with doctors (Altin & Stock, 2016; Lin et al., 2014).

With regard to the field of study, information technology and engineering students in Malaysia are 57% less likely to use the doctor as a primary source of health information. This could be because the field of studies, such as IT and engineering are presumed to have less engagement in health-related subjects and hence students in these field may not pay attention to health information. However, further research could be needed to answer why IT and engineering students are less likely to use doctors for health information when compared to students from other field of studies, such as business and social sciences.

The only factor associated with the use of television or radio as the primary source of health information among Ethiopian participants is health literacy. Ethiopian participants with limited health literacy level are 58% more likely to use television or radio as a primary source of health information. People with low health literacy level show little or no engagement in health information seeking (Ellis et al., 2012) but likely to receive health information from traditional media, such as television and radio. However, people with adequate health literacy level exhibit the most engagement in health information seeking and seek health information from the Internet actively (Ellis et al., 2012).

#### **5.2.2. Factors Associated with the Frequency of Health Information Seeking**

Results of the general linear model using Poisson regression analysis identified that gender, smoking, and health literacy among both Malaysian and Ethiopian participants; mother tongue, perceived severity of health problems, and perceived health status among Malaysian participants; and age and year of study among Ethiopian participants are found to affect the frequency of health information seeking.

Female participants from Malaysia and Ethiopia are 68% and 78% more likely to frequently seek health information than male participants, respectively. Previous research studies support this finding that females are generally more popular health information seekers when compared to males (Escoffery et al., 2005; Kim, 2015; Manierre, 2015;

Medlock et al., 2012; Wang et al., 2013). The reason why females look for health information more than males could be because they are more concerned about their health (Mokhtar et al., 2009). Furthermore, morbidity is higher among females (Hosseinpour et al., 2012) which could motivate them to look for health information more than males. On the other hand, English as mother tongue contributes to more frequent health information seeking when compared to other mother tongue. This could be because English as a mother tongue could help participants to express their ideas better than those whose mother tongue is not English.

Although health information seekers have healthier lifestyles than avoiders (Beaudoin & Hong, 2011), the frequency of health information seeking is 1.3 times higher among current smokers from both Malaysia and Ethiopia than non-smokers. The reason why currently smokers seek health information more frequently could be because they may have greater needs for health information to restore healthy behaviour. Furthermore, current smokers have high perceived susceptibility to and severity of health problems (82%) which is identified in this study.

As for health literacy, adequate health literacy level contributes to frequent health information seeking among participants from Malaysia and Ethiopia. This is because an adequate level of health literacy engages people in health information seeking (Ellis et al., 2012; Eriksson-Backa et al., 2012; Gutierrez et al., 2014). Similarly, the frequency of health information seeking is 1.4 and 1.3 times higher among participants from Malaysia who perceive their health problems would make them sick and those who perceive their health status as healthy when compared to participants from Malaysia who perceive their health problem as 'not a big deal' and those who perceive their health status as not healthy, respectively. The higher perceived severity of health problems and their healthy lifestyle seem to motivate the participants to frequently seek health information.

In this research study, the older age group of university students from Ethiopia are found to frequently seek health information. This is because most participants from Ethiopia use traditional media such as television and radio as the primary source of health information, which is the characteristics of the older age group. A previous research study indicates that older age groups are more likely to use traditional media for health information

seeking frequently while younger age groups use the Internet (Tennant et al., 2015). This could be due to the fact that younger age groups are more close to digital technologies and the Internet than older groups (Fox & Duggan, 2013). With regard to the year of study, freshman students from Ethiopia are more likely to frequently seek health information when compared to 5<sup>th</sup>-year students. This could be because freshman year is a critical period where factors such as homesickness, sleeping problem and academic load act on students influencing their health. Hence, their health problems could motivate freshman students to frequently seek health information.

In general, the factors identified in this research study are related to the factors described in the proposed conceptual framework in section 2.4.5 (Figure 6). In the proposed conceptual framework, demographic, behavioural, psychological, and health-related factors are indicated as the factors associated with health information seeking behaviour. The finding of this research study also show that most demographic, behavioural, psychological, and health-related factors affect health information seeking behaviour. That is, demographic factors such as gender and age are found to influence frequency of health information seeking. Behavioural factors such as alcohol and smoking are also found to influence health information seeking and frequency of health information seeking, respectively. With regard to psychological factors, perceived susceptibility to and severity of health problems are found to affect health information seeking and frequency of health information seeking, respectively. Health literacy and perceived health status are the health-related variables found to influence frequency of health information seeking. However, variables such as physical activity and health self-efficacy which are listed as factors in the proposed conceptual framework (Figure 6), are not found to influence health information seeking behaviour.

### **5.3. Online Health Information Searching**

In this section, findings of the online health information searching behaviour in relation to query and result viewing behaviours are discussed.

#### **5.3.1. Querying Formulation**

This research study identifies three types of queries, word-based (9.6%), question-based (22.4%), and run together words-based (68%) queries. Most of the participants use run-together words-based queries, showing that the participants seem to look for specific answers rather than exploring their information needs. Analysis of query reformulation pattern in previous research studies shows that users issue a single-word query when they want to broaden their search and explore more information to have a deeper understanding of their information needs (Broussard & Zhang, 2013; Rieh & Xie, 2006). On the other hand, users add terms to their original query to find specific information to fulfil their information needs (Broussard & Zhang, 2013; Rieh & Xie, 2006). Formulating queries using questions suggests a need to include an automated question-answering technique in health information retrieval systems so that the systems can receive users' natural language questions and return precise answers to searchers (Ramprasath & Hariharan, 2014).

The average number of queries issued by the participants is identified to be 1.9 for the simulated tasks, which indicates less reformulation effort, implying succinct query issuance by the participants. Spink et al. (2004) point out that health-related searches lack query reformulation which could be attributed to succinct query issuance. On the other hand, the average query length issued to search engines is 4.4 terms per query which is longer than the average query length (2.3 terms per query) identified in a research study done by Zhang (2013) on searching behaviour of students when looking for specific health-related Information in MedlinePlus. The long queries in this research could be due to the higher health literacy level of the participants. In this study, it is identified as those participants with adequate health literacy level are more likely to issue long queries (result section 4.4.2.4, Table 14). A previous study shows that people with adequate health literacy level are less likely to face difficulties in articulating queries (Zarcadoolas et al., 2003). Hence, people with adequate health literacy level could easily and comfortably express their queries to search engines. The use of long query length could mean that participants express their information needs well to search engines to retrieve required information to satisfy their need because long query length is associated with increased user satisfaction (Belkin et al., 2003; Lopes & Ribeiro, 2010). When users issue long queries that characterize their information need, their search effectiveness



increases, resulting in better information retrieval. In this study, the level of satisfaction with search results is 79%.

A paired-sample t-test identifies a significant difference in the number of queries between simulated tasks and personal task. The number of queries is higher in simulated tasks than a personal task. This could be because of the complexity of the simulated tasks. For example, in task 1 (about dengue), participants are requested to find the causes, signs and symptoms, and treatment of dengue. In this case, participants are observed formulating queries three times (cause of dengue, sign and symptoms of dengue, and treatment of dengue) which makes the number of queries three. Similarly, a significant difference in query length is found between simulated tasks and personal tasks. The result implies that the query length is longer in simulated tasks than a personal task. This could be because participants can easily take query terms from simulated task description and easily express their information needs to search engines. In this study, 88.2% of the queries are formulated using query terms from task descriptions.

### **5.3.2. Contextual Factors of Query Formulation**

Context is essential to understand information searching behaviour, especially in the domain of health as it is rich in contexts. In this section, the contextual factors associated with the number of queries and query length are discussed.

Out of all the variables entered to the Poisson regression model, mother tongue, physical activity, perceived task difficulty, and frequency of health information seeking are found to affect the number of queries. English as mother tongue helps to issue more number of queries as compared to the non-English mother tongue. This could be because participants whose mother tongue is English could formulate queries easily and issue more number of queries to acquire detailed information about the search tasks. Similarly, the number of queries are found to be 1.2 times higher among participants who are physically active than physically inactive participants. This is because healthier people are more likely to actively and frequently search for health information to increase their health knowledge and improve their health status (Pálsdóttir, 2008), implying increased experience and skills in health information searching (Belkin et al., 2003; Lopes & Ribeiro,

2010). Physically active individuals could issue more queries to increase their search effectiveness.

With regard to perceived task difficulty, the number of queries is 1.4 times higher among those participants who perceive the tasks as easy when compared to those who perceive the tasks not easy. Their understanding of the searching tasks could enable those participants who perceive the search tasks as easy to issue more number of queries to get the required information. Similarly, more number of queries is observed among frequent health information seekers. This finding is in line with a study done by Aula (2003) which indicates that frequent web searchers have the tendency to formulate more number of queries to have a deeper understanding of search tasks.

On the other hand, perceived susceptibility to health problems, previous health information seeking, and health literacy are found to be the factors associated with the query length. Queries are 1.29 times longer among participants who are concerned about health problems than those who are not concerned about health problems. This could be because those participants who are concerned about health problems are more likely to look for health information, which is identified in this study (Section 4.3). Their experience in health information seeking could help the participants who are concerned about health problems to issue longer queries to retrieve the required information. Previous studies (Aula, 2003; Lopes & Ribeiro, 2010) reveal that web searchers have the tendency to issue longer queries. On the contrary, queries are found to be 1.23 times longer among participants who have never sought health information. This could be because participants who have never sought health information use too many stop words during query formulation. About 60% of the participants who have never sought health information use stop words in their queries. Since they have no experience in health information seeking previously, participants who have never sought health information may not be aware that stop words have insignificant use to characterize documents.

Regarding health literacy, participants with inadequate health literacy level are less likely to issue longer query when compared to those with adequate health literacy level. This could be because they may have difficulties in expressing health queries. People face

difficulties in articulating health queries mainly due to limited health literacy which is characterized as having necessary skills and competencies needed to find, understand, evaluate, and use health information (Zarcadoolas et al., 2003).

### **5.3.3. Query Reformulation**

Query reformulation is the modification of the previous search query to retrieve better results. In this research study, it is observed that about 86% of the query reformulation are conceptual changes. Among the conceptual changes, 'switch to new concepts' is the most prevalent reformulation pattern followed by 'specification'. The popularity of 'switch to new concepts' and 'specification' might indicate a lack of defining search strategies and understanding the level of specificity of the search tasks. These could lead participants to formulate queries by trial and error. Providing synonymic and related health terminologies in information retrieval systems could help users in selecting suitable terms.

With respect to query reformulation that changes the form of a query, the most frequent reformulation pattern is 'reform to question' followed by 'spelling correction'. The reformulation of queries to questions suggests a need to include automated question-answering techniques in information retrieval systems so that the system can receive users' natural language questions and return precise answers (Ramprasath & Hariharan, 2014). Providing an auto-correction technique for misspelled terms could help to minimize users' reformulation efforts as a result of misspellings.

### **5.3.4. Search Domain**

Google is identified as the most frequent search engine used by participants (96.4%). This could be due to lack of proper knowledge and awareness of domain-specific sources among people from developing countries (Raj et al., 2015). Google is useful to support searchers engaged in basic search tasks, but it may not be helpful for searchers who need a deeper understanding of a health topic with multifaceted search tasks that need multiple search sessions and continuous interactions (Pang et al., 2014). Hence, it is necessary to promote domain-specific search engines or health information retrieval systems, such as WebMed, MedlinePlus, and national health information portals (e.g.

Myhealth), which are rarely used by participants in this study. Results of other research studies from developing countries, such as in Iran (Gavgani et al., 2013), Southeast Asia (Inthiran & Soyiri, 2015), Malaysia (Inthiran et al., 2013), China (Yuli et al., 2012) have supported our finding that Google is the most preferred search engine to search for health information. Very few other search engines, such as Bing (1.2%), MedlinePlus (0.4%), YouTube (0.3), Reddit (0.2%), and Wikipedia (0.2%) are utilized by participants.

Participants have visited a total of 504 websites. The most frequently visited websites are webmd.com (10.5%), cdc.gov (5.9%), nhs.uk (4.7%), nhlbi.nih.gov (3.7%), who.int (3.6%), and mayoclinic.org (3.1%). Wikipedia, which is a general purpose site, appeared to be one the popular site visited by the participants when performing all the search tasks. The results from the online survey also indicate Wikipedia as the second main source of health information. The reason for participants to use Wikipedia frequently could be because it is ranked high in search engines. Furthermore, the contents could be easier for participants to understand. In the future, it is worthwhile to study the role of general purpose websites, such as Wikipedia, as it is found as one of the most frequently visited websites. Another aspects of the websites visited is the national health portals, such as moh.gov.my and myhealth.gov, which are rarely used in this study. Promoting such websites among university students could help to get country-specific health information.

#### **5.4. Challenges of Health Information Searching**

Overall, most of the participants are happy with their search experience. However, they have challenges and barriers when performing all the searching tasks. Most of the participants indicate that it is time-consuming to wade through volumes of data to find the required information. This could be due to a lack of understanding the level of specificity of the search tasks, which can be explained by their query reformulation effort. When users are not satisfied with their initial search results, they reformulate the initial queries and spend a long time in the searching session. Providing related terms from the original query could help users to specify their information need to information retrieval system so as to get the required information. Participants have also faced difficulties in judging the quality of health information. Judging the quality of health information by looking at a piece

of information could be challenging for users. However, health information retrieval systems could be designed in a way to provide information on how other previous searchers rate a piece of information. Furthermore, providing information on how to evaluate health information in information retrieval systems could help users in judging the quality of health information. Excessive and irrelevant information is the other challenge the participants faced when performing the search tasks. The excessive and irrelevant information might indicate that participants were not specific in their query formulation. Providing specific terms or queries after users issue their first query to an information retrieval system could help users to be specific in their query formulation.

## **5.5. Research Contribution**

This section presents the theoretical and practical contribution from the findings of this research study. First, the theoretical contributions are presented followed by the practical contributions.

### **5.5.1. Theoretical Contribution**

The research findings obtained from the survey and experiment provide several research contributions in relation to health information seeking and searching behaviours. To begin with, HISB of university students from middle and low-income countries, and their difference in health information seeking between the countries are known. For example, a significant proportion of university students from developing countries have sought health information. The percentage of health information seeking from a case in a middle-income country is 77.1% while it is 61.8% in a low-income country. The results could help for health promotion activities. Although there are a significant proportion of health information seekers from developing countries, there is still a need to promote health information seeking among university students so as to increase the percentage of health information seeking and improve the health status of the students. Educating university students about the importance of health information seeking through university wellness and safety programs, websites and social media pages, could motivate and engage students in health information seeking.

#### **5.5.1.1. Identification of the Sources of Health Information**

University students from a middle-income country are found to use the Internet as the main source of health information, whereas students from a low-income country prefer mostly television and radio as the main source of health information. However, a significant proportion of university students from a low-income country also use the Internet as a preferred source of health information. Furthermore, social media is identified as useful platforms to spread health messages among university students from both middle and low-income countries. Facebook is identified as the most frequent social networking sites used by university students from developing countries. The results could help health promoters to pass their health messages via the source that university students prefer. For example, university students from Ethiopia prefer television and radio as their main source of health information. Health educators in Ethiopia could design a television or radio program to educate individuals on how to prevent and control prevalent diseases, such as cardiovascular disease, diabetes, cancer, and chronic respiratory disease (Misganaw et al., 2017). This can be done through doctors' discussion with community in a live television and radio programs or through advertisements. Similarly, the government of Malaysia could provide information on how to prevent and control non-communicable diseases, which are the leading cause of premature deaths in Malaysia (Ministry of Health Malaysia, 2011), through its health information portal ([www.myhealth.gov.my](http://www.myhealth.gov.my)) as the majority of the participants from Malaysia prefer the Internet as the main source of health information. In addition, health information can be passed through governmental social media pages.

#### **5.5.1.2. Identification of the Type of Health Information**

Unlike previous research studies which identify people from developing countries as illness-related information seekers (Mohd-Nor et al., 2013; Muhammed & D'Souza, 2014; Prasad et al., 2013; Sharif et al., 2015), this research study identifies university students from both middle- and low-income countries as both illness- and wellness-related seekers. The specific information that university students sought most is about fitness, diet, mental health problems, and skin health. Health educators and promoters could use this finding to provide useful information about fitness, diet, mental health problems, and skin health

to university students on university wellness and safety programs, websites and social media pages.

#### **5.5.1.3. Identification of the Factors Associated with Health Information Seeking**

The factors associated with health information seeking, the sources of health information and the frequency of health information seeking are identified. Identifying the factors could help to promote health and improve the health status of university students in developing countries. Health literacy, perceived susceptibility to a health problem, and alcohol consumption are the factors identified to affect health information seeking among university students from a middle-income country. Adequate health literacy level, alcohol consumption, and concerning about getting health problems contribute to health information seeking.

#### **5.5.1.4. Identification of the Factors Associated with the Sources of Health Information**

Health literacy and perceived severity of health problems among university students from a middle-income country, and perceived health status among university students from a low-income country are found to affect the use of the Internet as a primary source of health information. Higher health literacy level and higher perceived severity of health problems among university students in a middle-income country and poor perceived health status among university students in a low-income country have a positive effect on the choice of the Internet as a primary source of health information.

Gender among university students from a middle-income country and 'perceived susceptibility to health problems' among university students from a low-income country are the factors found to affect the use of family as a primary source of health information. Female students and students who are not concerned with health problems are found to use family as a primary source of health information. With regard to doctors, the factors identified are body mass index, health literacy, and field of study among university students from a middle-income country. Overweight and adequate health literacy level individuals are more likely to consult doctors for health information. Furthermore, IT and

engineering students are less likely to consult doctors for health information when compared to students from other field of studies, such as business and social sciences. Health literacy is also found to affect the use of television or radio as the primary source of health information among university students from a low-income country. University students with limited health literacy level are more likely to use television or radio as a primary source of health information.

The identification of the factors associated with the source of health information could help to identify individual groups with a preferred source of health information so that health educators and promoters could pass health information through medium individuals are convenient with. For example, it is good to provide health information through the Internet to university students with adequate health literacy level as these students from a middle-income country are more likely to use the Internet as a primary source of health information. Similarly, health educators and promoters could provide health information to university students with limited health literacy level through television and radio programs as the media are convenient sources of health information to students from a low-income country.

#### **5.5.1.5. Identification of Factors Associated with the Frequency of Health Information Seeking**

This research study has identified the factors associated with the frequency of health information seeking. Identifying the factors could help to promote health and improve the health status of university students in developing countries. Gender, smoking, and health literacy among university students from both middle-and low-income countries; mother tongue, perceived severity of health problems, and perceived health status among university students from a middle-income country; and age and year of study among university students from a low-income country are found to affect the frequency of health information seeking. Being female, smoking, adequate health literacy level, higher perceived severity of health problems, and English as mother tongue contribute to frequent health information seeking.



#### **5.5.1.6. Identification of the Contextual Factors of Query Formulation**

This research study has made an effort to explore factors that contextualize users health information searching. Specifically, the factors associated with query formulation are identified. Identifying the factors could help information retrieval systems to learn and predict the information needs of a user, relate one information with another, display required information in a suitable manner to users, relates tasks performed by a searcher and inform other searchers to use it (Bierig & Göker, 2006; Ingwersen & Jarvelin, 2005). In this research study, mother tongue, physical activity, perceived task difficulties, and frequency of health information seeking are found to affect the number of queries while perceive susceptibility to health problems, previous health information seeking, and health literacy are identified as factors associated with the query length. English as mother tongue, physical activity, perceiving the search tasks as easy contribute to more number of queries. On the other hand, being concerned about health problems, having experience in health information seeking, and adequate health literacy level lead to formulating longer queries.

#### **5.5.1.7. Query Reformulation Patterns**

This research has examined query reformulation activities involved when participants search for health information. The results suggest that, in the search process, participants have focused on switching to new concepts, changing part of the concepts in the previous query, and making a query more specific. These changes indicate that when searching for health information participants often begin with general concepts then go to specific aspects of the search tasks. Furthermore, the efforts to make a query more specific could mean a lack in understanding the level of specificity of the search tasks. Providing synonymic and related terminologies in information retrieval systems could help users in specifying their query. Participants are also observed when they attempt to change the form of query. Reform to questions and spelling correction are identified as the most frequent changes made by participants. Providing auto-spelling correction technique in information retrieval systems could help users to minimize their reformulation effort.

#### **5.5.1.8. Search Domain**

This research contributes to the understanding of search engines used and websites that participants visited when performing the search tasks. The most frequently used search engine is Google (96.4%), which is a general search engine, which might not support searchers who need a deeper understanding about a health topic with multifaceted search tasks having multiple search sessions. Efforts should be made to promote health-specific search engines such as WebMed, MedlinePlus, and national health information portals (e.g. Myhealth) which are rarely used in this study. Such search engines could help to get country-specific and health-specific information. In addition, health-specific search engines are likely to give more relevant results specific to the domain of health (Trivedi, 2009).

The websites visited by the participants show a long distribution, with 32.5% of the websites accounting for 86.1% of the visits that participants made to websites. Although the majority of the sites visited by the participants are health-specific websites, such as webmd.com, cdc.gov, nhs.uk, nhlbi.nih.gov, who.int, and mayoclinic.org, Wikipedia, which is a general purpose site, appeared to be one the most popular site visited by the participants when performing all the search tasks. Another contribution is the national health portals, such as moh.gov.my and myhealth.gov, which are rarely used by the participants, accounting for only 0.4% of the total number of visits. Efforts should be made to promote the use of such sites. Such websites could help to provide country-specific health information.

#### **5.5.1.9. Identification of the Challenges in Health Information Searching**

Generally, university students are believed to have good Internet skills due to their university education which might provide them with the skills to use the Internet for academic purposes. However, they still experience difficulties when searching for health information on the Internet. The most frequent challenge identified when participants perform the search tasks are: 1) time consuming to wade through volumes of data to find required information, 2) difficulties in judging the quality of health information, 3) excessive and irrelevant information, 4) low credibility of the information sources, and 5) not able to obtain information immediately. Improving the design of information retrieval

systems could help to minimize the challenges. For example, providing related queries may help users to get specific information related to their information need so that irrelevant information could be avoided. Providing information on how to evaluate health information in information retrieval systems could help users in judging the quality of health information as well.

### **5.5.2. Practical Contribution**

The practical contributions include suggestions for future research studies and for the improvement of information retrieval systems in the domain of health. The contributions are specifically in relation to health information seeking and searching behaviours.

#### **5.5.2.1. Comprehensive Theoretical Framework**

From the literature review, a new comprehensive theoretical framework that could help to guide future research studies on HISB and OHISB is developed from four basic theoretical models, such as Wilsons' 1999 nested model of information seeking and information searching research areas (Wilson, 1999), Longo's 2010 expanded model of health information seeking behaviour (Longo et al., 2010), the Health Belief Model (Rosenstock, 1974), and Sutcliffe and Ennis' Process Model of Information Searching (Sutcliffe & Ennis, 1998). Wilson's model enables to view OHISB as a subset of HISB and helps to define and limit the scope of HISB and OHISB. Longo's model guides to investigate factors associated with HISB by taking the contextual and personal factors into account. Longo's model fails to include psychological factors, such as perceived susceptibility to health problems, perceived severity of health problems, and health self-efficacy. Hence, the Health Belief Model is used to include the psychological factors which are considered as the driving forces behind in health information seeking (Rosenstock, 1974). The Sutcliffe and Enns' model shows the standard information searching process enabling researchers to easily understand users' querying and result viewing behaviours. The type and source of health information seeking along with the demographical, psychological, behavioural, and health-related factors, and the process of online health information searching could represent a far more comprehensive conceptual model to study HISB and OHISB. The newly developed conceptual framework could help to guide

future research studies on information seeking and searching behaviours in general and health information seeking and searching behaviours in particular.

#### **5.5.2.2. The Need for Question and Answering Technique**

The question-based query is found to be a popular type of query identified in this study. Out of all queries submitted to search engines, 22.4% of them are found to be question-based queries. This indicates participants prefer to express their query or information need in a more natural way using natural language questions. Wildemuth et al. (1994). State that the most natural way for people to seek information is to ask questions. The popularity of question-based queries could indicate the participants' preference to find specific answers to their information needs. Furthermore, analysis of query reformulation reveals that the most frequent reformulation pattern that changes the form of the query is 'reform to a question' which shows the need to find specific answers for information needs in a natural way. Including a question-answering technique in health information retrieval systems could help users to express their information need in natural language questions and retrieve specific answers for their questions. A question and answering system can automatically analyse documents and return answers in response to users' questions (Dodiya & Jain, 2013). Information extraction and natural language processing techniques could be employed to provide relevant answers to questions posed by users. For example, if a user issue a natural language question, the question and answering system will submit the question to a search engine. Then, the system extracts likely answers from the search results and returns the most plausible answers to the user.

#### **5.5.2.3. The Need for Better Query Refining Assistant**

This research study identifies that most queries are run together-based queries, implying that the participants have attempted to find specific answers for their information needs. Furthermore, the result of query reformulation indicates that 'specification' is the popular query reformulation pattern identified in this study, which could indicate a lack in defining search strategies and understanding the level of specificity of the search tasks. As a result, participants have faced challenges of getting excessive and irrelevant information from the result page. Hence, providing query refining feature to health information

retrieval systems could help users to refine their search and retrieve specific answers for their information needs. Ways to provide query refining features could be refining by keywords and health topics. Refining by keyword could help to access results related to the query, whereas refining by health topics could help to access results based on specific health topics. Furthermore, query suggestion could be provided for users to perform related searches that allows either to formulate a new query or to add terms to an existing query. For example, if a user perform a search using a query 'how to stay healthy', the information retrieval system needs to provide an option that allows the user to refine the search by health topics or keywords such as exercise, physical activity, or diet. The system shall also suggest new queries related to the original query, for example, 'healthy living' and 'how to stay fit'.

## **5.6. Chapter Summary**

This research study aims to examine health information seeking behaviour and its associated factors among university students from developing countries. The results reveal that a higher proportion of university students from Malaysia (77.1%) and Ethiopia (61.8%) have sought health information mainly from the Internet. Their university education, which could help to increase their Internet skills, coupled with the availability of the Internet in university campuses seem to motivate university students to seek health information more than the general population from previous study in China (59%) (Yuli et al., 2012) and rural community in Malaysia (46%) (Mohd-Nor et al., 2013).

In line with previous studies from developing countries (Garcia-Cosavalente et al., 2010; Gavvani et al., 2013; Mohd-Nor et al., 2013; Muhammed & D'Souza, 2014; Prasad et al., 2013), television or radio are found to be the primary source of health information among Ethiopian participants. However, similar to previous studies carried out on developed countries (Feng & Xie, 2015; Hackman & Pember, 2016; Medlock et al., 2015; Weaver et al., 2010; Yuli et al., 2012), the Internet is identified to be the primary source of health information among Malaysian participants. Due to the limited access to the Internet outside university campuses (The World Bank, 2018a), as an alternative source of health information together with the ease of access to television or radio programs might have caused Ethiopians to prefer television or radio as the primary source of health information.

However, Malaysian prefer the Internet for health information because of the higher Internet penetration rate (The World Bank, 2018a), which enables easy and prompt access to health information (Cutilli, 2010).

With regard to the factors associated with health information seeking, participants with adequate health literacy level and higher perceived susceptibility to health problems and alcohol consumers in Malaysia are more likely to look for health information than their counterparts. The findings are similar to other studies in relation to health literacy (Ellis et al., 2012; Gutierrez et al., 2014), perceived susceptibility to health problems (Ahadzadeh et al., 2015; Mills & Todorova, 2016), and alcohol consumption (Wang et al., 2013).

This research also aims to explore the OHISB of university students from Malaysia. The results indicate that the participants have demonstrated behaviours of long queries, less query reformulation effort, misspellings, not viewing results beyond the first page, and clicking search results at the top of the result list. Long queries could mean well expression of queries to search engine because long query length is associated with increased user satisfaction (Belkin et al., 2003; Lopes & Ribeiro, 2010). The long queries could be due to the higher health literacy level of the participants as those participants with adequate health level are more likely to issue long queries which are identified in this study (result section 4.4.2.4, Table 14). People with adequate health literacy level are less likely to face difficulties in articulating queries (Zarcadoolas et al., 2003) and hence, they can easily and comfortably express their queries.

With regard to the contextual factor of query formulation, those participants whose mother tongue is English, physically active participants, those who perceive the search tasks as easy, and frequent health information seekers are found to issue more number of queries. English as a mother tongue is found to help users issue more number of queries as their English likely help them to easily express their ideas so that they can issue more number of queries. Similarly, if users perceive the search tasks as easy, then they are likely to easily express their information need and issue more number of queries. A research study by Aula (2003) supports our finding that frequent health information seekers are more likely to explore information by issuing more number of queries.

This study has examined the website visited by the participants. The majority of the websites visited are health-specific sites, however, Wikipedia, which is a general purpose site, is frequently used. Future studies may investigate the role of general purpose website for health information searching. Governmental health portals are rarely used by participants. Promoting such sites could help users to obtain and understand country-specific health information.

## CHAPTER 6: CONCLUSION

Health information seeking is a key element of health communication for better health outcome. To inform health promoters, educators, and search systems to better support health information searching, this research study is conducted. This research project has found significant proportions of university students from middle and low-income countries seeking health information. This study has also discovered a statistically significant difference between middle and low-income countries in health information seeking, more number of seekers being observed among university students in a middle-income country. Although not statistically significant, university students from a low-income country have more frequently sought health information than students from a middle-income country.

In terms of the sources of health information, university students from a middle-income country are found to use the Internet as the main source of health information, whereas students from a low-income country prefer television or radio to look for health information. However, a significant proportion of students from a low-income country also use the Internet as a source of health information. Social media are found to be playing a major role for health communications, indicating it as a potential tool for health promotion activities.

Unlike previous research studies which identify people from developing countries as illness-related information seekers, this study identifies university students from developing countries as both illness- and wellness-related information seekers. The main reasons for university students to look for health information are to seek information about fitness, diet, mental health problems, and skin health.

Overall, this research identifies that demographic, psychological, behavioural and health related factors which are listed in Figure 6 of the proposed conceptual framework in section 2.4.5 can affect health information seeking behaviour among university students from developing countries. Health literacy, perceived susceptibility to a health problem, and alcohol consumption are the factors identified to affect health information seeking among university students from the middle-income country. Furthermore, health literacy



and perceived severity of health problems among students from a middle-income country, and perceived health status among students from a low-income country are found to affect the use of the Internet as a primary source of health information. Regarding the frequency of health information seeking, gender, smoking, and health literacy level are found to affect the frequency of health information seeking among university students from both middle and low-income countries. In addition, mother tongue, perceived severity of health problems, and perceived health status are found to affect the frequency of health information seeking among students from a middle-income country, whereas age and year of study among students from a low-income country.

In the experimental study, it is found that participants have demonstrated search behaviours of long queries (4.3 terms per query), less query reformulation effort, misspellings, and clicking search results at the top of the result list. Furthermore, analysis of the number of queries using GEE analysis method identifies mother tongue, physical activity, perceived task difficulty, and frequency of health information seeking as factors associated with the number of queries. GEE analysis also identifies perceived susceptibility to health problems, previous health information seeking, and health literacy as factors affecting the query length.

Analysis of query reformulation indicates reformulation that leads to conceptual changes and reformulation that changes the form of a query as the two types of query reformulation identified in this study. In the conceptual changes, 'switching to new concepts' is identified as the most frequent reformulation pattern followed by 'specification'. Likewise, the most frequent query reformulation that changes the form of the query is 'reform to question' followed by 'spelling correction'.

The examination of search results identifies Google as the most frequently used search engine used by participants although it may not support searchers that need a deeper understanding of search tasks with multiple interactions with the search engine. Domain-specific search engines, for example, MedlinePlus and WebMD, are rarely used in this study. The majority of the websites viewed are found to be health-specific (e.g. webmd.com, cdc.gov, nhs.uk, nhlbi.nih.gov, who.int, and mayoclinic.org). However, Wikipedia, which is a general purpose site, is found to be a popular website visited by

participants. Governmental health information portals, such as moh.gov.my and myhealth.gov are rarely used by participants.

Although most participants are happy with the overall search experience, they have encountered several challenges and barriers when performing all the search tasks. The most prevalent search challenges are 1) time consuming to wade through volumes of data to find required information, 2) difficulties in judging the quality of health information, 3) excessive and irrelevant information, 4) low credibility of the information sources, and 5) not able to obtain information immediately. Understanding these challenges could lead to more usable and suitable health information systems.

This research study provides interesting findings. However, it has limitations. First, the simulated situations may not reflect participants' real health information needs that could give true search behaviours, although a personal task is utilized which accounts for real information needs to some extent. In addition, the experimental setting may not reflect true search behaviour as the participants know they are being observed and keystrokes are being logged. In future, researchers will need to conduct study with real information needs in a naturalistic settings so that true search behaviours can be observed.

Only one university from Ethiopia is considered to recruit participants for the web-based survey, which is a limitation to generalize the findings of the study because most of the participants in Ethiopia are from Amhara region where the university is located. The same is true for the participants in Malaysia although effort was made to include participants from east and west Malaysia. Future research studies should consider participants from diverse ethnic groups so as to generalize the findings.

This research study employed undergraduate students only from two developing countries, which may not be sufficient enough to represent all undergraduate students from developing countries. Future research studies will need to recruit participants from more developing countries and ensure adequate samples are taken so that the findings will be representative to all developing countries.

In order to recruit participants who are conveniently available to participate, convenience sampling technique is used in this research study. Although convenience sampling

technique is easy, cheap, and has the advantage of collecting data in short duration of time, it is prone to selection bias and sampling errors. Hence, future research studies could use other sampling techniques, for example simple random sampling technique, to recruit study subjects so that selection bias and sampling error can be minimized.

In addition, this research only employed undergraduate students from Malaysia for the experimental study. Undergraduate students from Ethiopia are not invited to participate in the experimental study due to resource constraints, which could show different OHISB. In the future, similar research study will be undertaken to explore the OHISB of undergraduate students from Ethiopia.

## REFERENCE

- Abhishek, V., & Hosanagar, K. (2007). *Keyword Generation for Search Engine Advertising Using Semantic Similarity Between Terms*. Paper presented at the Proceedings of the Ninth International Conference on Electronic Commerce, Minneapolis, MN, USA.
- Ahadzadeh, A. S., Pahlevan Sharif, S., Ong, F. S., & Khong, K. W. (2015). Integrating health belief model and technology acceptance model: An investigation of health-related internet use. *Journal of Medical Internet Research*, 17(2), e45. doi:10.2196/jmir.3564
- Alamrew, Z., Bedimo, M., & Azage, M. (2013). Risky Sexual Practices and Associated Factors for HIV/AIDS Infection among Private College Students in Bahir Dar City, Northwest Ethiopia. *International Scholarly Research Notice Public Health*, 2013, 9. doi:10.1155/2013/763051
- Allen, S., Geiger, B. F., Howard, V. J., Ivankova, N. V., Martin, M. Y., O'Neal, M., & Safford, M. M. (2013). *Development and validation of a survey instrument to assess health information-seeking behaviors among African American young professionals*. (PhD PhD Dissertation), University of Alabama at Birmingham, Birmingham. Retrieved from <http://contentdm.mhsl.uab.edu/cdm/ref/collection/etd/id/1479> (3561249)
- Altin, S. V., & Stock, S. (2016). The impact of health literacy, patient-centered communication and shared decision-making on patients' satisfaction with care received in German primary care practices. *BioMed Central Health Services Research*, 16(1), 450. doi:10.1186/s12913-016-1693-y
- Amnesty International. (2016). *Ethiopia Offline: Evidence of social media blocking and internet censorship in Ethiopia*. Retrieved from London, UK: <https://www.amnesty.org/download/Documents>
- Amoah, P. A., Gyasi, R. M., Koduah, A. O., & Edusei, J. (2017). Health literacy and self-perceived health status among street youth in Kumasi, Ghana. *Cogent Medicine*, 4(1275091), 1-13.
- Anker, A. E., Reinhart, A. M., & Feeley, T. H. (2011). Health information seeking: A review of measures and methods. *Patient Educ Couns*, 82(3), 346-354. doi:10.1016/j.pec.2010.12.008
- Ansari, W. E., & Stock, C. (2010). Is the Health and Wellbeing of University Students Associated with their Academic Performance? Cross Sectional Findings from the United Kingdom. *International Journal of Environmental Research and Public Health*, 7(2), 509-527. doi:10.3390/ijerph7020509
- Arora, M., Kanjilal, U., & Varshney, D. (2016). Evaluation of information retrieval: Precision and recall. *International Journal of Indian Culture and Business Management*, 12(2), 224-236. doi:10.1504/IJICBM.2016.074482
- Asian Health Literacy Association (AHLA). (2013). Using HLS-EU-Asia-Q Questionnaire. Retrieved from [http://www.ahla-asia.org/main.php?fid=100&com\\_id=106](http://www.ahla-asia.org/main.php?fid=100&com_id=106)
- Aula, A. (2003, 5-8 November 2003). *Query Formulation in Web Information Search*. Paper presented at the Proceedings of IADIS International Conference WWW/Internet, Algarve, Portugal.
- Avery, S., & Tracy, D. G. (2014). Using transaction log analysis to assess student search behavior in the library instruction classroom. *Reference Services Review*, 42(2), 320-335. doi:doi:10.1108/RSR-08-2013-0044
- Banas, J. (2008). A Tailored Approach to Identifying and Addressing College Students' Online Health Information Literacy. *American Journal of Health Education*, 39(4), 228-236. doi:10.1080/19325037.2008.10599043
- Bargh, J. A., & McKenna, K. Y. (2004). The internet and social life. *Annual Review Psychology*, 55, 573-590. doi:10.1146/annurev.psych.55.090902.141922
- Basic, J., & Erdelez, S. (2014). Active and Passive Acquisition of Health-related Information on the Web by College Students. *Proceedings of the American Society for Information Science and Technology*, 51(1), 1-5. doi:10.1002/meet.2014.14505101149

- Beaudoin, C. E., & Hong, T. (2011). Health information seeking, diet and physical activity: An empirical assessment by medium and critical demographics. *International Journal of Medical Informatics*, 80(8), 586-595. doi:10.1016/j.ijmedinf.2011.04.003
- Beiser, M., & Stewart, M. (2005). Reducing Health Disparities. *Canadian Journal of Public Health*, 96(2), S4-S5.
- Beiter, R., Nash, R., McCrady, M., Rhoades, D., Linscomb, M., Clarahan, M., & Sammut, S. (2015). The prevalence and correlates of depression, anxiety, and stress in a sample of college students. *Journal of Affective Disorders*, 173(0), 90-96. doi:<http://dx.doi.org/10.1016/j.jad.2014.10.054>
- Belkin, N. J., Kelly, D., Kim, G., Kim, J.-Y., Lee, H.-J., Muresan, G., . . . Cool, C. (2003, July 28 to August 1, 2003). *Query Length in Interactive Information Retrieval*. Paper presented at the Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Informaion Retrieval, Toronto, Canada.
- Bierig, R., & Göker, A. (2006, October 18 - 20, 2006). *Time, Location and Interest: An Empirical and User-centred Study*. Paper presented at the Proceedings of the 1st International Conference on Information Interaction in Context, Copenhagen, Denmark.
- Biswas, A., Oh, P. I., Faulkner, G. E., Bajaj, R. R., Silver, M. A., Mitchell, M. S., & Alter, D. A. (2015). Sedentary time and its association with risk for disease incidence, mortality, and hospitalization in adults: a systematic review and meta-analysis. *Ann Intern Med*, 162(2), 123-132. doi:10.7326/m14-1651
- Boden, C. (2014). Overcoming the linguistic divide: A barrier to consumer health information. *Journal of the Canadian Health Libraries Association*, 30(3), 75-80. doi:10.5596/c09-020
- Borlund, P. (2000). Experimental components for the evaluation of interactive information retrieval systems. *Journal of Documentation*, 56(1), 71-90. doi:doi:10.1108/EUM0000000007110
- Borzekowski, D. L., Fobil, J. N., & Asante, K. O. (2006). Online access by adolescents in Accra: Ghanaian teens' use of the internet for health information. *Dev Psychol*, 42(3), 450-4588. doi:10.1037/0012-1649.42.3.450
- Bratucu, R., Gheorghe, I. R., Purcarea, R. M., Gheorghe, C. M., Popa Velea, O., & Purcarea, V. L. (2014). Cause and effect: The linkage between the health information seeking behavior and the online environment- a review. *Journal of Medicine and Life*, 7(3), 310-316.
- Broussard, R., & Zhang, Y. (2013). *Seeking Treatment Options: Consumers' Search Behaviors and Cognitive Activities*. Paper presented at the Proceedings of the 76th ASIS&T Annual Meeting: Beyond the Cloud: Rethinking Information Boundaries, Montreal, Quebec, Canada.
- Brown, K. J. (2013). Southeast Asia Digital Future in Focus. In comScore (Ed.), (pp. 33). Virginia, USA: comScore.
- Bujang, M. A., Sa'at, N., Ikhwan, T. M., & Sidik, T. A. B. (2017). Determination of Minimum Sample Size Requirement for Multiple Linear Regression and Analysis of Covariance Based on Experimental and Non-experimental Studies. *Epidemiology Biostatistics and Public Health*, 14(3), 1-9.
- Case, D. O. (2012). *Looking for information : A survey of research on information seeking, needs, and behavior* (3rd ed. ed.). Bingley, West Yorkshire: Bingley, West Yorkshire : Emerald Group Publishing.
- Chen, Y., & Feeley, T. H. (2014). Numeracy, information seeking, and self-efficacy in managing health: An analysis using the 2007 Health Information National Trends Survey (HINTS). *Health Commun*, 29(9), 843-853. doi:10.1080/10410236.2013.807904
- Chew, M. H., Rahman, M. M., & Salleh, S. A. (2012). Dengue in Malaysia: An epidemiological perspective study. *Pakistan Journal of Medical Sciences*, 28(4), 643-647.
- Cotten, S. R., & Gupta, S. S. (2004). Characteristics of online and offline health information seekers and factors that discriminate between them. *Social Science and Medicine* 59(9), 1795-1806. doi:10.1016/j.socscimed.2004.02.020

- Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed. ed.). Thousand Oaks, Calif.: Thousand Oaks, Calif. : Sage Publications.
- Cutilli, C. C. (2010). Seeking health information: What sources do your patients use? *Orthopaedic Nursing*, 29(3), 214-219. doi:10.1097/NOR.0b013e3181db5471
- Dachew, B. A., Azale Bisetegn, T., & Berhe Gebremariam, R. (2015). Prevalence of Mental Distress and Associated Factors among Undergraduate Students of University of Gondar, Northwest Ethiopia: A Cross-Sectional Institutional Based Study. *PLoS ONE*, 10(3), e0119464. doi:10.1371/journal.pone.0119464
- David, D. H., & Lyons-Ruth, K. (2005). Differential attachment responses of male and female infants to frightening maternal behavior: Tend or befriend versus fight or flight? *Infant mental health journal*, 21(1), 1-18. doi:10.1002/imhj.20033
- Deliens, T., Deforche, B., De Bourdeaudhuij, I., & Clarys, P. (2015). Determinants of physical activity and sedentary behaviour in university students: A qualitative study using focus group discussions. *BioMed Central Public Health*, 15:201. doi:10.1186/s12889-015-1553-4
- Department of Statistics Malaysia. (2016). *Current Population Estimates*. Retrieved from Putrajaya, Malaysia: [https://www.dosm.gov.my/v1/index.php?r=column/pdfPrev&id=OWlxdEVoYlJCS0hUZzJyRUcvZ\\_EYxZz09](https://www.dosm.gov.my/v1/index.php?r=column/pdfPrev&id=OWlxdEVoYlJCS0hUZzJyRUcvZ_EYxZz09)
- Dickinson, D., & Raynor, D. K. T. (2003). Ask the patients—they may want to know more than you think. *British Medical Journal*, 327(7419), 861-861.
- Dodiya, T., & Jain, S. (2013). *Comparison of Question Answering Systems*. Paper presented at the Intelligent Informatics. Advances in Intelligent Systems and Computing, Berlin, Heidelberg.
- Edwards, C., Fox, R., Gillard, S., Gourlay, S., Guven, P., Jackson, C., . . . Drennan, V. (2013). *Explaining health managers' information-seeking behaviour and use*. Retrieved from London: [http://www.netscc.ac.uk/hsdr/files/project/SDO\\_FR\\_08-1808-243\\_V10.pdf](http://www.netscc.ac.uk/hsdr/files/project/SDO_FR_08-1808-243_V10.pdf)
- Eheman, C., Berkowitz, Z., Lee, J. W., Mohile, S. G., Purnell, J. Q., Roscoe, J. A., . . . Morrow, G. R. (2009). Information-seeking styles among cancer patients before and after treatment by demographics and use of information sources. *Journal of Health Communication*, 14(5), 487-502. doi:10.1080/10810730903032945
- Ekström, S., Kull, I., Nilsson, S., & Bergström, A. (2015). Web-Based Self-Reported Height, Weight, and Body Mass Index Among Swedish Adolescents: A Validation Study. *Journal of Medical Internet Research*, 17(3), e73. doi:10.2196/jmir.3947
- El-Attar, T., Gray, J., Nair, S. N., Ownby, R., & Czaja, S. J. (2005). *Older Adults and Internet Health Information Seeking*. Paper presented at the HFES 49th Annual Meeting: The Human Factors and Ergonomics Society Annual Meeting, Orlando.
- Ellis, J., Mullan, J., Worsley, A., & Pai, N. (2012). The role of health literacy and social networks in arthritis patients' health information-seeking behavior: A qualitative study. *Int J Family Med*, 2012, 6. doi:10.1155/2012/397039
- Eriksson-Backa, K., Ek, S., Niemela, R., & Huotari, M. L. (2012). Health information literacy in everyday life: A study of Finns aged 65-79 years. *Health Informatics J*, 18(2), 83-94. doi:10.1177/1460458212445797
- Escoffery, C., Miner, K. R., Adame, D. D., Butler, S., McCormick, L., & Mendell, E. (2005). Internet use for health information among college students. *J Am Coll Health*, 53(4), 183-188. doi:10.3200/jach.53.4.183-188
- European Commission. (2014, 28/11/2014). Europeans becoming enthusiastic users of online health information. Retrieved from <https://ec.europa.eu/digital-agenda/en/news/europeans-becoming-enthusiastic-users-online-health-information>

- Feinberg, I., Greenberg, D., & Frijters, J. (2015). *Understanding Health Information Seeking Behaviors of Adults with Low Literacy, Numeracy, and Problem Solving Skills: Results from the 2012 US PIAAC Study*. Retrieved from Atlanta, GA: [https://piaac.squarespace.com/s/Feinberg\\_Greenberg\\_Frijters\\_PIAAC.pdf](https://piaac.squarespace.com/s/Feinberg_Greenberg_Frijters_PIAAC.pdf)
- Feng, Y., & Xie, W. (2015). Digital Divide 2.0: The Role of Social Networking Sites in Seeking Health Information Online From a Longitudinal Perspective. *Journal of Health Communication*, 20(1), 60-68. doi:10.1080/10810730.2014.906522
- Fiksdal, S. A., Kumbamu, A., Jadhav, S. A., Cocos, C., Nelsen, A. L., Pathak, J., & McCormick, B. J. (2014). Evaluating the Process of Online Health Information Searching: A Qualitative Approach to Exploring Consumer Perspectives. *Journal Medical Internet Research*, 16(10), e224. doi:10.2196/jmir.3341
- Fink, A. (2013). *How to conduct surveys : A step-by-step guide* (5th ed. ed.). Los Angeles: Los Angeles : SAGE.
- Fox, S. (2011a). *Health Topics: 80% of internet users look for health information online*. Retrieved from Washington, D.C.:
- Fox, S. (2011b). *The Social Life of Health Information, 2011*. Retrieved from Washington, D.C.: <http://www.pewinternet.org/2011/05/12/the-social-life-of-health-information-2011/>
- Fox, S., & Duggan, M. (2013). *Health Online 2013*. Retrieved from Washington, D.C.: <http://www.pewinternet.org/2013/01/15/health-online-2013/>
- Galeshi, R., Sharman, J., & Cai, J. (2018). Influence of ethnicity, gender, and immigration status on millennials' behavior related to seeking health information. *Equality, Diversity and Inclusion*, 37(6), 621-631.
- Garcia-Cosavalente, H. P., Wood, L. E., & Obregon, R. (2010). Health information seeking behavior among rural and urban Peruvians: Variations in information resource access and preferences. *Information Development*, 26(1), 37-45. doi:10.1177/0266666909358640
- Gavvani, V. Z. (2010). *Health Information Need and Seeking Behavior of Patients in Developing Countries' Context: An Iranian Experience*. Paper presented at the Proceedings of the 1st ACM International Health Informatics Symposium, Arlington, Virginia, USA.
- Gavvani, V. Z., Qeisari, E., & Jafarabadi, M. A. (2013). Health information seeking behavior: A study of a developing country. *Library Philosophy and Practice*, 2013:902, 1-29.
- Gebreslassie, M., Feleke, A., & Melese, T. (2013). Psychoactive substances use and associated factors among Axum university students, Axum Town, North Ethiopia. *BioMed Central Public Health*, 13(1), 1-9. doi:10.1186/1471-2458-13-693
- Gholami, M., Fallahi Khoshknab, M., Maddah, S. S. B., Ahmadi, F., & Khankeh, H. (2014). Barriers to health information seeking in Iranian patients with cardiovascular disease: A qualitative study. *Heart & Lung: The Journal of Acute and Critical Care*, 43(3), 183-191. doi:<http://dx.doi.org/10.1016/j.hrtlng.2014.01.010>
- Goeuriot, L., Jones, G. J. F., Kelly, L., Müller, H., & Zobel, J. (2016). Medical information retrieval: introduction to the special issue. *Information Retrieval Journal*, 19(2), 1-5.
- Goje, M., Salmiah, M. S., Ahmad Azuhairi, A., & Jusoff, K. (2014). Physical Inactivity and Its Associated Factors among University Students. *IOSR Journal of Dental and Medical Sciences*, 13(10), 119-130. doi:10.6084/M9.FIGSHARE.1263367
- Graham, C., Richardson, A., King, S., Chiera, B., & Olds, T. (2014). An initial exploration of the association between psychological distress and sedentary behaviour in first year undergraduates. A Practice Report *The International Journal of the First Year in Higher Education*, 5(2), 95-101.
- Griffin, R. M. (2014). Stress Management Health Center [Press release]. Retrieved from <http://www.webmd.com/balance/stress-management/features/10-fixable-stress-related-health-problems>



- Gutierrez, N., Kindratt, T. B., Pagels, P., Foster, B., & Gimpel, N. E. (2014). Health literacy, health information seeking behaviors and internet use among patients attending a private and public clinic in the same geographic area. *J Community Health*, 39(1), 83-89. doi:10.1007/s10900-013-9742-5
- Guzman, A., & Istúriz, R. E. (2010). Update on the global spread of dengue. *International Journal of Antimicrobial Agents*, 36(1), S40-S42. doi:<http://dx.doi.org/10.1016/j.ijantimicag.2010.06.018>
- Hackman, C. L., & Pember, S. E. (2016). Qualitative investigation of health information seeking behavior utilizing social media in a college population. *American Journal of Health Studies*, 31(1), 11.
- Hamzah, M. R., Mohammad, E. M. W., Abdullah, M. Y., & Ayub, S. H. (2015). Scenario on Health Information Seeking in Malaysia: A Systematic Review. *Journal of Human Development and Communication*, 4(2015), 7-20.
- Harrell, F. E. a. (2015). *Regression modeling strategies: With applications to linear models, logistic and ordinal regression, and survival analysis* (SpringerLink Ed. Second edition. ed.): Cham : Springer.
- Health Information and Quality Authority. (2011). *International Review of National Health Information Sources* Retrieved from Dublin, Ireland: <http://lenus.ie/hse/handle/10147/221845>
- Hensel, B. K., Leshner, G., & Logan, R. A. (2013). Consumer Health Informatics Research Resource. Retrieved from <http://chirr.nlm.nih.gov/>
- Hesse, B. W., Nelson, D. E., Kreps, G. L., Croyle, R. T., Arora, N. K., Rimer, B. K., & Viswanath, K. (2005). Trust and sources of health information: the impact of the Internet and its implications for health care providers: Findings from the first Health Information National Trends Survey. *Arch Intern Med*, 165(22), 2618-2624. doi:10.1001/archinte.165.22.2618
- Hoffmann, J. P. a. (2016). *Regression models for categorical, count, and related variables : An applied approach* (ProQuest Ed.): Oakland, California : University of California Press.
- Hosseinpour, A. R., Stewart Williams, J., Amin, A., Araujo de Carvalho, I., Beard, J., Boerma, T., . . . Chatterji, S. (2012). Social Determinants of Self-Reported Health in Women and Men: Understanding the Role of Gender in Population Health. *PLoS ONE*, 7(4), e34799. doi:10.1371/journal.pone.0034799
- Hu, R., Lu, K., & Joo, S. (2013). Effects of topic familiarity and search skills on query reformulation behavior. *Proceedings of the American Society for Information Science and Technology*, 50(1), 1-9.
- Hunt, J., & Eisenberg, D. (2010). Mental Health Problems and Help-Seeking Behavior Among College Students. *Journal of Adolescent Health*, 46(1), 3-10. doi:10.1016/j.jadohealth.2009.08.008
- Hussain, S., & Arasad, M. (2015). Prevalence of Social Networks Addiction among Ethiopian Youths *International Journal of Science and Research*, 4(3), 834-837.
- Ingwersen, P., & Jarvelin, K. (2005). Information retrieval in context: IRIx. *Special Interest Group on Information Retrieval Forum*, 39(2), 31-39. doi:10.1145/1113343.1113351
- Inthiran, A., Alhashmi, S., & Ahmed, P. (2011). Profiling a Non-medical Professional Searcher on a Medical Domain: What Do Search Patterns and Demographic Details Reveal? In M. Salem, K. Shaalan, F. Oroumchian, A. Shakery, & H. Khelalfa (Eds.), *Information Retrieval Technology* (Vol. 7097, pp. 157-168): Springer Berlin Heidelberg.
- Inthiran, A., Alhashmi, S. M., & Ahmed, P. K. (2013). Online Consumer Health: A Malaysian Perspective. *Information Technology in Developing Countries*.
- Inthiran, A., Alhashmi, S. M., & Ahmed, P. K. (2015). A user study on the information search behaviour of medical students. *Malaysian Journal of Library & Information Science*, 20(1), 61-77.
- Inthiran, A., & Soyiri, I. (2015). Searching for Health Information Online for My Child: A Perspective from South East Asia. In B. R. Allen, J. Hunter, & L. M. Zeng (Eds.), *Digital Libraries: Providing Quality Information: 17th International Conference on Asia-Pacific Digital Libraries, ICADL 2015, Seoul, Korea, December 9-12, 2015. Proceedings* (pp. 76-81). Cham: Springer International Publishing.
- Jadhav, A., Andrews, D., Fiksdal, A., Kumbamu, A., McCormick, J. B., Misitano, A., . . . Pathak, J. (2014). Comparative Analysis of Online Health Queries Originating From Personal Computers and Smart



- Devices on a Consumer Health Information Portal. *Journal of Medical Internet Research*, 16(7), e160. doi:10.2196/jmir.3186
- Jadhav, A., Sheth, A., & Pathak, J. (2014). Analysis of Online Information Searching for Cardiovascular Diseases on a Consumer Health Information Portal. *AMIA Annual Symposium Proceedings*, 2014, 739-748.
- Janeice, A., Eileen, M., & Trauth, M. (2013). Socio-economic influences on health information searching in the USA: The case of diabetes. *Information Technology & People*, 26(4), 324-346. doi:10.1108/ITP-09-2012-0098
- Jena, S., Ch, H., & Tiwari, r. (2015). Stress and mental health problems in 1st year medical students: A survey of two medical colleges in Kanpur, India. *International Journal of Research in Medical Sciences*, 3(1), 130-134.
- Jiang, J., Awadallah, A. H., Shi, X., & White, R. W. (2015). *Understanding and Predicting Graded Search Satisfaction*. Paper presented at the Proceedings of the Eighth ACM International Conference on Web Search and Data Mining, Shanghai, China.
- Jung, M. (2014). Determinants of health information-seeking behavior: Implications for post-treatment cancer patients. *Asian Pacific Journal of Cancer Prevention*, 15(16), 6499-6504.
- Kelly, D. (2006). Measuring online information seeking context, Part 1: Background and method. *Journal of the American Society for Information Science and Technology*, 57(13), 1729-1739. doi:10.1002/asi.20483
- Kelly, D. (2009). *Methods for Evaluating Interactive Information Retrieval Systems with Users* (Vol. 3). Boston, USA: Boston : Now Publishers Inc.
- Kelly, D., & Sugimoto, C. R. (2013). A systematic review of interactive information retrieval evaluation studies, 1967–2006. *Journal of the American Society for Information Science and Technology*, 64(4), 745-770. doi:10.1002/asi.22799
- Kim. (2015). An exploratory study of inactive health information seekers. *International Journal of Medical Informatics*, 84(2), 119-133. doi:10.1016/j.ijmedinf.2014.10.003
- Kim, Park, S.-Y., & Bozeman, I. (2011). Online health information search and evaluation: observations and semi-structured interviews with college students and maternal health experts. *Health Information & Libraries Journal*, 28(3), 188-199. doi:10.1111/j.1471-1842.2011.00948.x
- Kuttschreuter, M., Rutsaert, P., Hilverda, F., Regan, Á., Barnett, J., & Verbeke, W. (2014). Seeking information about food-related risks: The contribution of social media. *Food Quality and Preference*, 37(October 2014), 10-18. doi:<http://dx.doi.org/10.1016/j.foodqual.2014.04.006>
- Lalazaryan, A., & Zare-Farashbandi, F. (2014). A Review of models and theories of health information seeking behavior. *International Journal of Health System and Disaster Management*, 2(4), 193-203. doi:10.4103/2347-9019.144371
- Lambert, S. D., & Loiselle, C. G. (2007). Health information seeking behavior. *Qual Health Res*, 17(8), 1006-1019. doi:10.1177/1049732307305199
- Larouere, B., & Winters, C. S. (2013). *Examination of Sedentary Time, Physical Activity, and Body Mass Index (BMI) in College-Aged Students*. Paper presented at the Mid-Atlantic Regional Chapter Conference Proceedings - American College of Sports Medicine, Harrisburg, USA.
- Larson, R. R. (2001). TREC interactive with Cheshire II. *Information Processing & Management*, 37(3), 485-505. doi:[http://dx.doi.org/10.1016/S0306-4573\(00\)00058-3](http://dx.doi.org/10.1016/S0306-4573(00)00058-3)
- Lee, S. Y., Hwang, H., Hawkins, R., & Pingree, S. (2008). Interplay of Negative Emotion and Health Self-Efficacy on the Use of Health Information and Its Outcomes. *Communication Research*, 35(3), 358-381.
- Li, F., Li, M., Guan, P., Ma, S., & Cui, L. (2015). Mapping publication trends and identifying hot spots of research on Internet health information seeking behavior: A quantitative and co-word biclustering analysis. *Journal of Medical Internet Research*, 17(3), e81. doi:10.2196/jmir.3326

- Lin, X., Wang, M., Zuo, Y., Li, M., Lin, X., Zhu, S., . . . Lamoureux, E. L. (2014). Health Literacy, Computer Skills and Quality of Patient-Physician Communication in Chinese Patients with Cataract. *PLoS ONE*, 9(9), e107615. doi:10.1371/journal.pone.0107615
- Longo, D. R. (2005). Understanding health information, communication, and information seeking of patients and consumers: A comprehensive and integrated model. *Health Expectations*, 8(3), 189-194. doi:10.1111/j.1369-7625.2005.00339.x
- Longo, D. R., Ge, B., Radina, M. E., Greiner, A., Williams, C. D., Longo, G. S., . . . Salas-Lopez, D. (2009). Understanding breast-cancer patients' perceptions: Health information-seeking behaviour and passive information receipt. *Journal of Communication in Healthcare*, 2(2), 184-206. doi:10.1179/cih.2009.2.2.184
- Longo, D. R., Schubert, S. L., Wright, B. A., LeMaster, J., Williams, C. D., & Clore, J. N. (2010). Health Information Seeking, Receipt, and Use in Diabetes Self-Management. *Annals of Family Medicine*, 8(4), 334-340. doi:10.1370/afm.1115
- Lopes, C. T., & Ribeiro, C. (2010). *Context Effect on Query Formulation and Subjective Relevance in Health Searches*. Paper presented at the Proceedings of the Third Symposium on Information Interaction in Context, New Brunswick, New Jersey, USA.
- Lopes, C. T., & Ribeiro, C. (2015). Effects of Terminology on Health Queries: An Analysis by User's Health Literacy and Topic Familiarity. In A. Woodsworth & W. D. Penniman (Eds.), *Current Issues in Libraries, Information Science and Related Fields* (Vol. 39, pp. 145-184). UK: Emerald Group Publishing Limited.
- Luo, G., Tang, C., Yang, H., & Wei, X. (2008). *MedSearch: A Specialized Search Engine for Medical Information Retrieval*. Paper presented at the Proceedings of the 17th ACM Conference on Information and Knowledge Management, Napa Valley, California, USA.
- Lwoga, E. T., & Mosha, N. F. (2013). Information seeking behaviour of parents and caregivers of children with mental illness in Tanzania. *Library Review*, 62(8/9), 567-584. doi:10.1108/LR-10-2012-0116
- Mackert, M., Champlin, S., & Mabry, A. (2014). *Understanding College Student Health Literacy*. Paper presented at the 2014 St. David's CHPR Annual Conference Promoting Healthy Communities: Interdisciplinary Perspectives, Austin, Texas. <https://www.utexas.edu/nursing/chpr/docs/2014/poster/Mackert%20Abstract%2027.pdf>
- Malaysian Communications and Multimedia Commission. (2015). *Internet users survey 2014*. Retrieved from Cyberjaya, Selangor: <http://www.skmm.gov.my/Resources/Statistics/Internet-users-survey.aspx>
- Manganello, J. A. (2008). Health literacy and adolescents: A framework and agenda for future research. *Health Education Research*, 23(5), 840-847.
- Manierre, M. J. (2015). Gaps in knowledge: Tracking and explaining gender differences in health information seeking. *Social Science & Medicine*, 128(March 2015), 151-158. doi:10.1016/j.socscimed.2015.01.028
- Marchionini, G. (2006). Exploratory Search: From Finding to Understanding. *Communications of the ACM*, 49(4), 41-46. doi:10.1145/1121949.1121979
- Markiewicz, D., Lawford, H., Doyle, B. A., & Haggart, N. (2006). Developmental Differences in Adolescents' and Young Adults' Use of Mothers, Fathers, Best Friends, and Romantic Partners to Fulfill Attachment Needs. *Journal of Youth & Adolescence*, 35(1), 121-134. doi:10.1007/s10964-005-9014-5
- Matsveru, D. (2013). *Information needs and Information seeking behaviour of Namibian pastors*. In (pp. 196). Retrieved from <https://www.hausarbeiten.de/document/233272>
- Mc Laughlin, G. H. (1969). SMOG Grading-a New Readability Formula. *Journal of Reading*, 12(8), 639-646.

- McMullan, M. (2006). Patients using the Internet to obtain health information: How this affects the patient–health professional relationship. *Patient Educ Couns*, 63(1–2), 24–28. doi:<http://dx.doi.org/10.1016/j.pec.2005.10.006>
- Medlock, Eslami, S., Askari, M., Arts, L. D., Sent, D., de Rooij, E. S., & Abu-Hanna, A. (2015). Health Information–Seeking Behavior of Seniors Who Use the Internet: A Survey. *Journal of Medical Internet Research*, 17(1), e10. doi:10.2196/jmir.3749
- Medlock, Eslami, S., Askari, M., Sent, D., Buurman, B., De Rooij, S., & Abu-Hanna, A. (2012). Health information seeking behavior of elderly Internet users in the Netherlands. *European Geriatric Medicine*, 3(1), S90–S91. doi:<http://dx.doi.org/10.1016/j.eurger.2012.07.195>
- Medlock, Eslami, S., Askari, M., Sent, D., de Rooij, S. E., & Abu-Hanna, A. (2013). The consequences of seniors seeking health information using the internet and other sources. *Stud Health Technol Inform*, 192, 457–460.
- Mehta, V., & Bhat, A. (2015). Health Awareness and Behavior Among Adolescent Students in a Rural School: A Cross Sectional Observational Study. *International Journal of Research in Medical Sciences*, 3(12), 3499–3502. doi:10.18203/2320-6012.ijrms20151231
- Mengistu, T. S., Melku, A. T., Bedada, N. D., & Eticha, B. T. (2013). Risks for STIs/HIV infection among Madawalabu University students, Southeast Ethiopia: A cross sectional study. *Reproductive Health*, 10(1), 1–7. doi:10.1186/1742-4755-10-38
- Miller, L. M., & Bell, R. A. (2012). Online health information seeking: The influence of age, information trustworthiness, and search challenges. *J Aging Health*, 24(3), 525–541. doi:10.1177/0898264311428167
- Mills, A., & Todorova, N. (2016). *An Integrated Perspective on Factors Influencing Online Health-Information Seeking Behaviours*. Paper presented at the Australasian Conference on Information Systems Mills & Todorova, New South Wales, Australia.
- Ministry of Communication and Information Technology Ethiopia. (2016). The National ict Policy and Strategy. Retrieved from <http://www.mcit.gov.et/web/english/the-national-ict-policy>
- Ministry of Health Malaysia. (2008). *Malaysia's Health*. Retrieved from Putrajaya, Malaysia: <http://www.moh.gov.my/images/gallery/publications/mh/Malaysia%20Health%202008-2.pdf>
- Ministry of Health Malaysia. (2011). *NCD Prevention and Control: Annual Report*. Retrieved from Putrajaya, Malaysia: <http://www.moh.gov.my/images/gallery/nspncd/AnnualReport2010NCD.pdf>
- Ministry of Health Malaysia. (2012). *Annual Report: Ministry of Health*. Retrieved from Putrajaya, Malaysia: [http://vlib.moh.gov.my/cms/content.jsp?id=com.tms.cms.section.Section\\_cc7aebb6-c0a81049-18ac8000-11671d74](http://vlib.moh.gov.my/cms/content.jsp?id=com.tms.cms.section.Section_cc7aebb6-c0a81049-18ac8000-11671d74)
- Minnie, D., & Srinivasan, S. (2011). Meta Search Engines for Information Retrieval on Multiple Domains. *International Journal of Technology And Engineering System*, 2(2), 115–118.
- Misganaw, A., Haregu, T. N., Deribe, K., Tessema, G. A., Deribew, A., Melaku, Y. A., . . . Naghavi, M. (2017). National mortality burden due to communicable, non-communicable, and other diseases in Ethiopia, 1990–2015: Findings from the Global Burden of Disease Study 2015. *Population Health Metrics*, 15:29, 1–17. doi:10.1186/s12963-017-0145-1
- Mohamad, E., Su, T. T., Majid, H. A., & Chinna, K. (2014). *Health Literacy in Malaysia: HLS-Asia questionnaire validation*. Paper presented at the The Second International Conference on Health Literacy and Health Promotion, Taipei, Taiwan.
- Mohd-Nor, R., Chapun, T. E., & Wah, C. R. J. (2013). Malaysian rural community as consumer of health information and their use of ICT. *Malaysian Journal of Communication*, 29(1), 161–178.
- Mokhtar, I. A., Goh, J.-E., Li, K. J., & Tham, C. X.-L. (2009). Medical and Health Information Seeking among Singapore Youths: An Exploratory Study. *Singapore Journal of Library and Information Management*, 38(2009), 49–76.

- Moly, T. M. (2014). Assessment of Internet Use by Under Graduate Students of Information Science in College of Computing and Informatics, Haramaya University, Ethiopia. *International Journal of Innovative and Applied Research*, 2(1), 41-50.
- Muhammed, K., & D'Souza, S. C. (2014). A Review of Preventive Health Information Seeking Behavior of Common People in Developing Countries. Paper presented at the MARCON 2014: International Marketing Conference, Knoxville Marriott, India.
- Nasrabadi, A. N., Sabzevari, S., & Bonabi, T. N. (2015). Iranian Women's Experiences of Health Information Seeking Barriers: A Qualitative Study in Kerman. *Iranian Red Crescent medical journal*, 17(2), e25156.
- National Cancer Institute. (2014). Health Information National Trends Survey. *Survey Instruments*. Retrieved from <http://hints.cancer.gov/instrument.aspx>
- Noguchi, S., Ogino, D., & Sato, H. (2015). Information Channels and Needs of Health Professionals of Clinical Trials in Japan. *Journal of Pharmaceutical Care & Health Systems*, 2(5), 1-6. doi:10.4172/2376-0419.1000144
- Norris, S. A., Anuar, H., Matzen, P., Cheah, J. C. H., Jensen, B. B., & Hanson, M. (2014). The life and health challenges of young Malaysian couples: Results from a stakeholder consensus and engagement study to support non-communicable disease prevention. *BioMed Central Public Health*, 14(2), S2-S6. doi:10.1186/1471-2458-14-S2-S6
- Nutbeam, D. (2000). Health literacy as a public health goal: A challenge for contemporary health education and communication strategies into the 21st century. *Health Promotion International*, 15(3), 259-267. doi:10.1093/heapro/15.3.259
- Obasola, O. I., & Agunbiade, O. M. (2016). Online Health Information Seeking Pattern Among Undergraduates in a Nigerian University. *SAGE Open*, 6(1), 1-9. doi:10.1177/2158244016635255
- Oh, K. M., Kreps, G. L., Jun, J., Chong, E., & Ramsey, L. (2012). Examining the health information-seeking behaviors of Korean Americans. *Journal of Health Communication*, 17(7), 779-801. doi:10.1080/10810730.2011.650830
- Oman Global School-based Student Health Survey. (2012). *Oman Global School-based Student Health Survey Report*. Retrieved from Oman: [http://www.who.int/chp/gshs/Oman\\_GSHS\\_Country\\_Report.pdf](http://www.who.int/chp/gshs/Oman_GSHS_Country_Report.pdf)
- Osei Asibey, B., Agyemang, S., & Boakye Dankwah, A. (2017). The Internet Use for Health Information Seeking among Ghanaian University Students: A Cross-Sectional Study. *International journal of telemedicine and applications*, 2017, 1-9. doi:10.1155/2017/1756473
- Pálsdóttir, Á. (2008). Information behaviour, health self-efficacy beliefs and health behaviour in Icelanders' everyday life. *Information Research*, 13(1), Paper 334.
- Pang, P. C.-I., Chang, S., Pearce, J., & Verspoor, K. (2014). *Online Health Information Seeking Behaviour: Understanding Different Search Approaches*. Paper presented at the The 18th Pacific Asia Conference on Information system: IT Ubiquitous and Collaborative Innovation, Chengdu, China.
- Park, M., & Lee, T. s. (2013). Understanding science and technology information users through transaction log analysis. *Library Hi Tech*, 31(1), 123-140. doi:doi:10.1108/07378831311303976
- Parvanta, C. F., Nelson, D. E., Parvanta, S. A., & Harner, R. N. (2011). *Essentials of public health communication* (C. F. Parvanta Ed.). Sudbury, MA: Sudbury, MA : Jones & Bartlett Learning.
- Patel, S., & Dowse, R. (2013). Understanding the medicines information-seeking behaviour and information needs of South African long-term patients with limited literacy skills. *Health Expectations*, 18(5), 1494-1507. doi:10.1111/hex.12131
- Percheski, C., & Hargittai, E. (2011). Health Information-Seeking in the Digital Age. *Journal of American College Health*, 59(5), 379-386. doi:10.1080/07448481.2010.513406
- Peytremann-Bridevaux, I., & Santos-Eggimann, B. (2007). Healthcare utilization of overweight and obese Europeans aged 50–79 years. *Journal of Public Health*, 15:377. doi:10.1007/s10389-007-0103-7

- Prasad, B. M., Srinath, S., & Chadha, S. (2013). *Health Literacy: How do Communities Get Informed About TB in India?* Paper presented at the Health Systems in Asia: Equity, Governance and Social Impact, National University of Singapore, Singapore.
- Puspitasari, I., Moriyama, K., Fukui, K., & Numao, M. (2015). Effects of individual health topic familiarity on activity patterns during health information searches. *JMIR Med Inform*, 3(1), e16. doi:10.2196/medinform.3803
- Raj, S., Sharma, V. L., Singh, A., & Goel, S. (2015). The health information seeking behaviour and needs of community health workers in Chandigarh in Northern India. *Health Info Libr J*, 32(2), 143-149. doi:10.1111/hir.12104
- Ramprasath, M., & Hariharan, S. (2014). Query refinement based question answering system using pattern analysis. *Advances in Natural and Applied Sciences*, 8(17), 8-15.
- Ramsey, I., Corsini, N., Peters, M. D. J., & Eckert, M. (2017). A rapid review of consumer health information needs and preferences. *Patient Educ Couns*, 100(9), 1634-1642. doi:10.1016/j.pec.2017.04.005
- Reang, T., & Bhattacharjya, H. (2013). A study to assess the emotional disorders with special reference to stress of medical students of agartala government medical college and govinda ballabh pant hospital. *Indian journal of community medicine*, 38(4), 207-211. doi:10.4103/0970-0218.120154
- Richardson, A., Allen, J. A., Xiao, H., & Vallone, D. (2012). Effects of race/ethnicity and socioeconomic status on health information-seeking, confidence, and trust. *J Health Care Poor Underserved*, 23(4), 1477-1493. doi:10.1353/hpu.2012.0181
- Rieh, S. Y., & Xie, H. (2006). Analysis of multiple query reformulations on the web: The interactive information retrieval context. *Information Processing & Management*, 42(3), 751-768. doi:<http://dx.doi.org/10.1016/j.ipm.2005.05.005>
- Rosenstock, I. M. (1974). Historical Origins of the Health Belief Model. *Health Education and Behaviour*, 2(4), 328-335. doi:10.1177/109019817400200403
- Rubens, M., Darrow, W. W., Sebekos, E., Gabbidon, K., Batra, A., & Tanaka, H. (2014). *Exploring the Determinants of Risky Sexual Behavior Among Ethnically Diverse University Students: The Student Behavioral Health Survey*. Paper presented at the APHA 142nd Annual Meeting and Expo, New Orleans, USA. <https://apha.confex.com/apha/142am/webprogram/Paper305680.html>
- Samuolis, J., Layburn, K., & Schiaffino, K. M. (2001). Identity Development and Attachment to Parents in College Students. *Journal of Youth and Adolescence*, 30(3), 373-384. doi:10.1023/A:1010448313516
- Schroeder, H. (2010). Creating Library Tutorials for Nursing Students. *Medical Reference Services Quarterly*, 29(2), 109-120. doi:10.1080/02763861003723135
- Sharif, S. P., Ahadzadeh, A. S., & Wei, K. K. (2015). A Moderated Mediation Model of Internet Use for Health Information. *Journal of Social Sciences*, 4(1), 611-625.
- Sharit, J., Taha, J., Berkowsky, R. W., Profita, H., & Czaja, S. J. (2015). Online Information Search Performance and Search Strategies in a Health Problem-Solving Scenario. *Journal of cognitive engineering and decision making*, 9(3), 211-228. doi:10.1177/1555343415583747
- Shepard, D. S., Lees, R., Ng, C. W., Undurraga, E. A., Halasa, Y., & Lum, L. (2013). *Burden of Dengue in Malaysia. Report from a Collaboration between Universities and the Ministry of Health of Malaysia*. Retrieved from Kuala Lumpur, Malaysia: [http://people.brandeis.edu/~shepard/Report\\_dengue\\_in\\_Malaysia\\_v50.pdf](http://people.brandeis.edu/~shepard/Report_dengue_in_Malaysia_v50.pdf)
- Sorensen, K., Van den Broucke, S., Pelikan, J., Fullam, J., Doyle, G., Slonska, Z., . . . Consortium, o. b. o. t. H.-E. (2013). Measuring health literacy in populations: Illuminating the design and development process of the European Health Literacy Survey Questionnaire (HLS-EU-Q). *BioMed Central Public Health*, 13:948, 1-10.
- Sparck Jones, K. (1981). *Information retrieval experiment*. London: Boston.



- Spink, A., Yang, Y., Jansen, J., Nykanen, P., Lorence, D. P., Ozmutlu, S., & Ozmutlu, H. C. (2004). A study of medical and health queries to web search engines. *Health Information & Libraries Journal*, 21(1), 44-51. doi:10.1111/j.1471-1842.2004.00481.x
- Sterne, J., & Kirkwood, B. (2003). *Essential Medical Statistics*. Hoboken, UNKNOWN: John Wiley & Sons, Ltd.
- Stewart-Brown, S., Evans, J., Patterson, J., Petersen, S., Doll, H., Balding, J., & Regis, D. (2000). The health of students in institutes of higher education: An important and neglected public health problem? *J Public Health Med*, 22(4), 492-499.
- Stommel, M., & Wills, C. E. (2004). *Clinical Research: Concepts and Principles for Advanced Practice Nurses*. (Brief Article)(Book Review). In Vol. 28. (pp. 465).
- Suehs, B. T., Kamble, P., Huang, J., Hammer, M., Bouchard, J., Costantino, M. E., & Renda, A. (2017). Association of obesity with healthcare utilization and costs in a Medicare population. *Current Medical Research and Opinion*, 33(12), 2173-2180. doi:10.1080/03007995.2017.1361915
- Sutcliffe, A., & Ennis, M. (1998). Towards a cognitive theory of information retrieval. *Interacting with Computers*, 10(3), 321-351. doi:10.1016/S0953-5438(98)00013-7
- TechSmith Corporation. (2015). Morae Components and Features. Retrieved from <https://www.techsmith.com/morae-features.html>
- Tennant, B., Stellefson, M., Dodd, V., Chaney, B., Chaney, D., Paige, S., & Alber, J. (2015). eHealth literacy and Web 2.0 health information seeking behaviors among baby boomers and older adults. *Journal of Medical Internet Research*, 17(3), e70. doi:10.2196/jmir.3992
- Terasaki, D. J., Gelaye, B., Berhane, Y., & Williams, M. A. (2009). Anger expression, violent behavior, and symptoms of depression among male college students in Ethiopia. *BioMed Central Public Health*, 9:13, 1-8. doi:10.1186/1471-2458-9-13
- Tesfaye, G., Derese, A., & Hambisa, M. T. (2014). Substance Use and Associated Factors among University Students in Ethiopia: A Cross-Sectional Study. *Journal of Addiction*, 2014, 1-8. doi:10.1155/2014/969837
- Thaver, A. M., & Kamal, A. (2010). Impact of information sources on the knowledge of adolescents about hepatitis B. *Journal of Pakistan Medical Association*, 60(12), 1072-1075.
- The Boston Consulting Group. (2015). *Why well-being should drive growth strategies: The 2015 Sustainable and Economic Development Assessment* Retrieved from Boston, USA:
- The World Bank. (2013). Health, Nutrition and Population Statistics. Retrieved from <http://datatopics.worldbank.org/hnp/HNPDash.aspx>
- The World Bank. (2015a). Country and Lending Groups. Retrieved from <http://data.worldbank.org/about/country-and-lending-groups>
- The World Bank. (2015b). Internet users (per 100 people). Retrieved from <http://data.worldbank.org/indicator/IT.NET.USER.P2/countries>
- The World Bank. (2018a). Individuals using the Internet (% of population). Retrieved from <https://data.worldbank.org/indicator/IT.NET.USER.ZS?locations=ET-MY>
- The World Bank. (2018b). Literacy rate, adult total (% of people ages 15 and above). Retrieved from <https://data.worldbank.org/indicator/SE.ADT.LITR.ZS>
- Tonsaker, T., Bartlett, G., & Trpkov, C. (2014). Health information on the Internet: Gold mine or minefield? *Canadian Family Physician*, 60(5), 407-408.
- Treiman, K., & Squiers, L. (2005). *The CIS research agenda: Overview of relevant research*. USA: National Institutes of Health, National Cancer Institute.
- Trivedi, M. (2009). A study of search engines for health sciences. *International Journal of Library and Information Science*, 1(5), 069-073.
- Turchik, J. A. (2012). Sexual Victimization Among Male College Students: Assault Severity, Sexual Functioning, and Health Risk Behaviors. *American Psychological Association*, 13(3), 243-255.

- U.S. National Library of Medicine. (2018). MeSH Browser Overview. Retrieved from <https://www.nlm.nih.gov/mesh/mbinfo.html>
- Wang, Viswanath, K., Lam, T. H., Wang, X., & Chan, S. S. (2013). Social Determinants of Health Information Seeking among Chinese Adults in Hong Kong. *PLoS ONE*, 8(8), e73049. doi:10.1371/journal.pone.0073049
- Weaver, J. B., Mays, D., Weaver, S. S., Hopkins, G. L., Eroğlu, D., & Bernhardt, J. M. (2010). Health Information–Seeking Behaviors, Health Indicators, and Health Risks. *American Journal of Public Health*, 100(8), 1520-1525. doi:10.2105/AJPH.2009.180521
- Wildemuth, B. M., de Bliek, R., Friedman, C. P., & Miya, T. S. (1994). Information-seeking behaviors of medical students: A classification of questions asked of librarians and physicians. *Bull Med Libr Assoc*, 82(3), 295-304.
- Willis, E., & Royne, M. B. (2017). Online Health Communities and Chronic Disease Self-Management. *Health Commun*, 32(3), 269-278. doi:10.1080/10410236.2016.1138278
- Wilson, T. D. (1999). Models in information behaviour research. *Journal of Documentation*, 55(3), 249-270. doi:10.1108/EUM00000000007145
- Wong, C., Harrison, C., Britt, H., & Henderson, J. (2014). Patient use of the internet for health information. *Australian Family Physician*, 43(12), 875-877.
- World Health Organization. (2003). *Report of a Joint WHO/FAO Expert Consultation: Diet Nutrition and the Prevention of Chronic Diseases*. Retrieved from Geneva, Switzerland: <http://www.fao.org/docrep/005/ac911e/ac911e00.HTM>
- World Health Organization. (2009a). *Global Health Risks: Mortality and burden of disease attributable to selected major risks*. Retrieved from Geneva, Switzerland: [http://www.who.int/healthinfo/global\\_burden\\_disease/GlobalHealthRisks\\_report\\_full.pdf](http://www.who.int/healthinfo/global_burden_disease/GlobalHealthRisks_report_full.pdf)
- World Health Organization. (2009b). *Track 2: Health Literacy and Health Behaviour*. Paper presented at the The 7th WHO Global Conference on Health Promotion, Nairobi, Kenya. <http://www.who.int/healthpromotion/conferences/7gchp/track2/en/>
- World Health Organization. (2010). *Global Recommendations on Physical Activity for Health*. Retrieved from Geneva, Switzerland <http://www.who.int/dietphysicalactivity/pa/en/>
- World Health Organization. (2012). Global Adult Tobacco Survey. In *Malaysia* (pp. 2). Geneva, Switzerland: WHO.
- World Health Organization. (2014a). *Global Status Report on noncommunicable diseases*. Retrieved from Geneva, Switzerland: [http://apps.who.int/iris/bitstream/10665/148114/1/9789241564854\\_eng.pdf?ua=1](http://apps.who.int/iris/bitstream/10665/148114/1/9789241564854_eng.pdf?ua=1)
- World Health Organization. (2014b). Noncommunicable Diseases (NCD) Country Profiles: Malaysia. In *Malaysia* (pp. 1). Geneva, Switzerland: WHO.
- World Health Organization. (2015a). Obesity and overweight. Retrieved from <http://www.who.int/mediacentre/factsheets/fs311/en/>
- World Health Organization. (2015b). Update on the Dengue situation in the Western Pacific Region. *Dengue Situation Update* 477, 477(3 November 2015), 6. Retrieved from [http://www.wpro.who.int/emerging\\_diseases/dengue\\_biweekly\\_20151103.pdf?ua=1](http://www.wpro.who.int/emerging_diseases/dengue_biweekly_20151103.pdf?ua=1)
- World Health Organization. (2015c). *WHO report on the global tobacco epidemic*. Retrieved from Geneva, Switzerland: [http://apps.who.int/iris/bitstream/10665/178574/1/9789240694606\\_eng.pdf?ua=1&ua=1](http://apps.who.int/iris/bitstream/10665/178574/1/9789240694606_eng.pdf?ua=1&ua=1)
- World Health Organization. (2015d). *World Health Statistics*. Retrieved from Geneva, Switzerland: [http://www.who.int/gho/publications/world\\_health\\_statistics/en/](http://www.who.int/gho/publications/world_health_statistics/en/)
- Yang, J.-J., Korfhage, R., & Rasmussen, E. M. (1992). *Query Improvement in Information Retrieval Using Genetic Algorithms - A Report on the Experiments of the TREC Project*. Paper presented at the The First Text REtrieval Conference (TREC), Gaithersburg, Maryland.

- Yeganova, L., Comeau, D. C., Kim, W., & Wilbur, W. J. (2009). How to interpret PubMed queries and why it matters. *Journal of the American Society for Information Science and Technology*, 60(2), 264-274. doi:10.1002/asi.20979
- Yi, Y. J. (2015). Health literacy and health information behavior of Florida public library users: A mixed methods study. *Journal of Librarianship and Information Science*, 47(1), 17-29. doi:10.1177/0961000614531159
- Yuli, Y., Su, Z., & Lijuan, X. (2012, 3-5 Aug. 2012). *Preventive Health Information Seeking Behavior Among General Population in China*. Paper presented at the 2012 International Symposium on Information Technologies in Medicine and Education, Hokodate, Hokkaido, Japan.
- Yun, E. K., & Park, H. A. (2010). Consumers' disease information-seeking behaviour on the Internet in Korea. *Journal of Clinical Nursing*, 19(19-20), 2860-2868. doi:10.1111/j.1365-2702.2009.03187.x
- Zarcadoolas, C., Pleasant, A., & Greer, D. S. (2003). Elaborating a definition of health literacy: A commentary. *Journal of Health Communication*, 8(1), 119-120. doi:10.1080/713851982
- Zeger, S. L., & Liang, K. Y. (1986). Longitudinal data analysis for discrete and continuous outcomes. *Biometrics*, 42(1), 121-130. doi:10.2307/2531248
- Zeng, Q. T., Kogan, S., Plovnick, R. M., Crowell, J., Lacroix, E. M., & Greenes, R. A. (2004). Positive attitudes and failed queries: An exploration of the conundrums of consumer health information retrieval. *International Journal of Medical Informatics*, 73(1), 45-55. doi:10.1016/j.ijmedinf.2003.12.015
- Zhang. (2013). Searching for specific health-related information in MedlinePlus: Behavioral patterns and user experience. *Journal of the Association for Information Science and Technology*, 65(1), 53-68. doi:10.1002/asi.22957
- Zhang, Tran, B. X., Le, H. T., Nguyen, H. D., Nguyen, C. T., Tran, T. D., . . . Ho, R. C. (2017). Perceptions of Health-Related Information on Facebook: Cross-Sectional Study Among Vietnamese Youths. *Interact J Med Res*, 6(2), e16. doi:10.2196/ijmr.8072
- Zhang, Wang, P., Heaton, A., & Winkler, H. (2012). *Health Information Searching Behavior in MedlinePlus and the Impact of Tasks*. Paper presented at the Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium, Miami, Florida, USA.



## APPENDIX

### Appendix A: Operational Definitions

**Health information seeking behaviour:** is defined as the ways in which people acquire or receive information about their health, diseases, risks to health, and health promotion. (Lambert & Loiselle, 2007). The Consumer Health Informatics Research Resource (CHIRr) group indicate that there is no single and widely used measure of health information seeking (Hensel et al., 2013). However, one of the most common approach to measure health information seeking is using a question from the United States Health Information National Trend Survey (HINTS), as suggested by CHIRr. The question is: *“Have you ever looked for information about health or medical topics from any source?”* (National Cancer Institute, 2014). Following this question, the time of seeking and frequency of seeking can be asked.

**Physical activity:** The World Health Organization (WHO) defined physical activity as “any bodily movement produced by skeletal muscles that require energy expenditure” (World Health Organization, 2010). According to the WHO recommendation, “adults of age 18–64 should do at least 150 minutes of moderate-intensity aerobic physical activities (e.g. brisk walking, bicycling at a regular pace, and swimming at a regular pace) throughout a week or do at least 75 minutes of vigorous-intensity aerobic physical activity throughout the week or an equivalent combination of moderate and vigorous-intensity activity”. In this study, physical activity is measured using two item questions which are used in the United States Health Information National Trend Survey (National Cancer Institute, 2014). The two items are 1) *“In a typical week, how many days do you do any physical activity or exercise of at least moderate intensity, such as brisk walking, bicycling at a regular pace, and swimming at a regular pace?”* 2) *“On the days that you do any physical activity or exercise of at least moderate intensity, how long do you typically do these activities?”*

**Health self-efficacy:** “Self-efficacy is the extent to which people believe they are capable of performing specific behaviours in order to attain certain goals” (Hensel et al., 2013). In the context of health, it can be measured by a single question, *“Overall, how confident are you about your ability to take good care of your health?”* which is developed by the

United States Health Information National Trend Survey (National Cancer Institute, 2014). The response consists of a five-point scale ranging from 1 (completely confidence) to 5 (not confident at all).

**Health literacy:** The World Health Organization defined Health Literacy as “the cognitive and social skills which determine the motivation and ability of individuals to gain access to, understand and use information in ways which promote and maintain good health”. One of the assessment tool used to measure health literacy in a population is the European Health Literacy Survey (HLS-EU). The tool is developed as per the HLS-EU project by its consortium including Maastricht University and eight institutes and universities in eight countries (Sorensen et al., 2013). The Asian Health Literacy Association adapted the tool to be used by researchers in Asian countries (Asian Health Literacy Association (AHLA), 2013). Further, a study was undertaken to assess the validity and internal reliability of the tool in the Malaysian context (Mohamad et al., 2014). This study found high reliability; and concluded that the tool can be utilized in the Malaysian context. Therefore, the 47 item HLS-EU questionnaire is used to determine the health literacy of participants under study. The 47 items are assessed using a 4-point self-report scale (very easy, easy, difficult, and very difficult).

**Body mass index (BMI):** BMI is calculated as the weight in kilograms divided by height in meters squared. A study identifies the difference between self-reported and actual measurements of adolescents (Ekström et al., 2015). When using self-reported technique 86.4% of adolescents are categorized into the correct BMI classification. The result of the study by Ekström et al. (2015) indicates that the underestimation of weight by 1.1kg and the overestimation of height by 0.5cm contribute to the underestimation of BMI by 0.5kg/m<sup>2</sup>. Ekström et al. (2015) suggest that collecting body height and weight using a web-based tool may be a valid, quick, and cost-effective alternative to measure BMI among adolescents. Hence, in this study, body weight and height is collected based on self-reports of study participants.

**Perceived susceptibility to a health problem** refers to an individual's subjective perception of the risks of acquiring that particular health problem. This can be measured

using a single item question: “*How concerned are you about getting any health problem?*” (National Cancer Institute, 2014).

**Perceived Severity of a health problem** refers to an individual’s subjective perception of the seriousness of a particular health problem. The United States Health Information National Trend Survey uses a single question to measure this variable, which is “*How getting a health problem would affect your life?*” (National Cancer Institute, 2014).

**Developing country:** is operationalized as a country whose gross national income per capita, is less than USD12, 736. Therefore, in this study, countries with low-income and middle-income economies, which are listed by The World Bank (2015a), are considered developing countries.



- j. Difficult to access doctors because they are too busy ☐ ☐ ☐ ☐ ☐
- k. Difficult to access doctors because they are ☐ ☐ ☐ ☐ ☐  
unaffordable for consultation ☐ ☐ ☐ ☐ ☐
- l. Difficult to access health magazines/brochures/  
newsletter ☐ ☐ ☐ ☐ ☐
- m. No time to attend Television/Radio health programs ☐ ☐ ☐ ☐ ☐
3. In the past 12 months, how often have you looked for health-related information from any sources? \_\_\_\_\_ times
4. In the past 1 month, how often have you looked for health-related information from any sources? \_\_\_\_\_ times
5. Have you **PURPOSELY OR INTENTIONALLY SOUGHT** health-related information in any of the following sources? Please select the most appropriate choice by ticking one of the boxes for each source. You are requested to give an answer to all items of the question.

- |                               | <b>Always</b>         | <b>Very often</b>     | <b>Sometimes</b>      | <b>Rarely</b>         | <b>Never</b>          |
|-------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 1. Television/Radio           | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 2. Internet                   | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 3. Family                     | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 4. Friends/co-workers         | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 5. Doctor/healthcare provider | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 6. Books                      | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 7. Brochures/Pamphlets        | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 8. Magazines                  | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 9. Newspapers                 | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
6. Have you **ACCIDENTALLY OR UNINTENTIONALLY** looked for health related information in any of the following sources although you were not seeking for these information? Please select the most appropriate choice by ticking one of the boxes for each source. You are requested to give an answer for all items of the question

- |                     | <b><u>Very often</u></b> | <b><u>Rather often</u></b> | <b><u>Rather seldom</u></b> | <b><u>Very seldom</u></b> | <b>Never</b>          |
|---------------------|--------------------------|----------------------------|-----------------------------|---------------------------|-----------------------|
| 1. Television/Radio | <input type="radio"/>    | <input type="radio"/>      | <input type="radio"/>       | <input type="radio"/>     | <input type="radio"/> |
| 2. Internet         | <input type="radio"/>    | <input type="radio"/>      | <input type="radio"/>       | <input type="radio"/>     | <input type="radio"/> |
| 3. Family           | <input type="radio"/>    | <input type="radio"/>      | <input type="radio"/>       | <input type="radio"/>     | <input type="radio"/> |

- |                               |                       |                       |                       |                       |                       |
|-------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 4. Friends/co-workers         | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 5. Doctor/healthcare provider | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 6. Books                      | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 7. Brochures/Pamphlets        | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 8. Magazines                  | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 9. Newspapers                 | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

7. The most recent time you looked for information about health related, where did you go first? Please indicate in the following order:

- |                               | First                 | Second                | Third                 | Fourth                | Never used            |
|-------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 1. Television/Radio           | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 2. Internet                   | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 3. Family                     | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 4. Friends/co-workers         | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 5. Doctor/healthcare provider | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 6. Books                      | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 7. Brochures/Pamphlets        | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 8. Magazines                  | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |
| 9. Newspapers                 | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> | <input type="radio"/> |

8. For the above question number 7, please indicate the specific source

- If from television/radio, specify the program \_\_\_\_\_
- If from the Internet, specify the source \_\_\_\_\_
- If from family, specify the family member (mother, father, brother, sister, other relatives...) \_\_\_\_\_
- If from books, specify the title of the book \_\_\_\_\_
- If from brochures/pamphlets/Magazines/Newspaper, specify the title \_\_\_\_\_
- If from brochures/pamphlets/Magazines/Newspaper, specify the owner/Author/drug company/Ministry of Health \_\_\_\_\_

9. The most recent time you looked for information about health related, who was it for?    1. Myself ☐        2. Family ☐        3. Friend ☐

10. The most recent time you looked for information about health related, what was the **MAIN REASONS**? Please indicate in the following order of reasons:

	<b>1<sup>st</sup></b> <b><u>Reason</u></b>	<b>2<sup>nd</sup></b> <b><u>Reason</u></b>	<b>3<sup>rd</sup></b> <b><u>Reason</u></b>	<b>4<sup>th</sup></b> <b><u>Reason</u></b>	<b>Never</b>
1. Specific disease or health problem	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. Treatment or procedure	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. Medication	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. Diet	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5. Exercise	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6. Environmental hazard	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7. Parenting	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8. Other, Specify_____	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

11. About how much time (hours and/or minutes) during a typical week did you spend to look for the following information?

	<b><u>Hours spend</u></b>	<b><u>Minutes spend</u></b>
1. Specific disease or health problem	_____	_____
2. Treatment or procedure	_____	_____
3. Medication	_____	_____
4. Diet/Nutrition	_____	_____
5. Exercise	_____	_____
6. Environmental hazard	_____	_____
7. Parenting	_____	_____
8. Other, Specify_____	_____	_____

12. The most recent time you looked for information about health related, what was the **MAIN HEALTH TOPICS**? Please indicate in the following order:

	<b><u>1<sup>st</sup> topic</u></b>	<b><u>2<sup>nd</sup> topic</u></b>	<b><u>3<sup>rd</sup> topic</u></b>	<b><u>4<sup>th</sup> topic</u></b>	<b><u>Never</u></b>
1. Fitness/exercise	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. Cancer	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. Diabetes	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. Heart disease	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5. Hypertension	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6. Lung disease	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7. Medicines/medications	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8. Sexual health	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9. Skin health	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

10. Depression, anxiety, stress, or

other mental health issues      ☐      ☐      ☐      ☐      ☐

11. Diet or nutrition      ☐      ☐      ☐      ☐      ☐

12. Smoking, alcohol and drug      ☐      ☐      ☐      ☐      ☐

13. Environmental Hazard      ☐      ☐      ☐      ☐      ☐

14. Dental Health      ☐      ☐      ☐      ☐      ☐

15. Others, specify \_\_\_\_\_      ☐      ☐      ☐      ☐      ☐

13. In general, how much would you trust health information from each of the following? Use the following scale:

**1. Not at all      2. A little      3. Some      4. A lot**

1. A doctor      \_\_\_\_\_

2. Family or friends      \_\_\_\_\_

3. Newspapers or magazines      \_\_\_\_\_

4. Radio      \_\_\_\_\_

5. Internet      \_\_\_\_\_

6. Television      \_\_\_\_\_

7. Government health agencies      \_\_\_\_\_

8. Charitable organizations      \_\_\_\_\_

9. Religious organizations and leaders      \_\_\_\_\_

14. What type of challenges do you face while seeking health information? Use the following scale:

**1. Always      2. Very often      3. Sometimes      4. Rarely      5. Never**

1. Doubt about the existence of relevant information      \_\_\_\_\_

2. Doubt about the quality of health information      \_\_\_\_\_

3. Health information is hard to understand      \_\_\_\_\_

4. Takes too much time and effort to get relevant information      \_\_\_\_\_

5. Unsatisfactory past search results      \_\_\_\_\_

6. Uncertainty about where to look for information      \_\_\_\_\_

7. Unavailability of health information system      \_\_\_\_\_

8. Lack of knowledge and skill on the use of information resources      \_\_\_\_\_

9. Online resources are too expensive      \_\_\_\_\_



10. Lack of access to medical library, journal, and books \_\_\_\_\_
11. Difficult to access doctors because they are too busy \_\_\_\_\_
12. Difficult to access doctors because they are unaffordable for consultation \_\_\_\_\_
13. Difficult to access health magazines/brochures/newsletter \_\_\_\_\_
14. No time to attend Television/Radio health programs \_\_\_\_\_

### Part III: Online Health Information Seeking

1. How long do you use the Internet per day? \_\_\_\_\_ hours
2. Have you ever use the Internet to look for health information?
1. Yes → go to Part III, question number 4
2. No → answer question number 3 and go to Part IV
3. If the answer for question number 2 is **NO**, what would be the reason? Choose from the list below and →go to Part IV

	<u>Always</u>	<u>Very often</u>	<u>Sometimes</u>	<u>Rarely</u>	<u>Never</u>
a. Poor Internet connection	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b. Doubt about the existence of relevant information	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c. Doubt about the quality of online health information	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d. Health information is hard to understand	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
e. Takes too much time and effort to get relevant Information	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
f. Uncertainty about where to look for information	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
g. Unavailability of health information system	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
h. Lack of knowledge and skill on the use of the Internet for health information	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
i. Online resources are too expensive	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
j. Lack of access to online medical library, journal, and books	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

4. In the past 12 months, how often have you looked for health related information from the Internet? \_\_\_\_\_ times
5. In the past 1 month, how often have you looked for health related information from the Internet? \_\_\_\_\_ times
6. How many years of experience do you have on online health information searching? \_\_\_\_\_ years

7. The most recent time you looked for online health information, where did go first?

Please indicate in the following order:

	<u>First</u>	<u>Second</u>	<u>Third</u>	<u>Fourth</u>
a. Google	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b. Yahoo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c. Web MD	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d. Medline Plus	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
e. Myhealth (Malaysian Health Portal)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Or Ethiopian Ministry of Health website				
f. Wikipedia	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
g. Others, specify _____	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

8. What were the difficulties when using the Internet to seek health information?

a. Time-consuming to wade through volumes of data to find required information

1. Yes      2. Maybe      3. No

b. Not able to obtain information immediately      1. Yes      2. Maybe      3. No

c. Too much/excessive irrelevant information      1. Yes      2. Maybe      3. No

d. Lack of information/inability to find what you are looking for online

1. Yes      2. Maybe      3. No

e. Low credibility of the information sources      1. Yes      2. Maybe      3. No

f. Difficulty to understand the information      1. Yes      2. Maybe      3. No

g. Difficult to judge the quality (reliability and accuracy) of the information

1. Yes      2. Maybe      3. No

h. Felt frustrated during searching for online health information

1. Yes      2. Maybe      3. No

i. Others, specify \_\_\_\_\_

9. In the last 12 months, have you used the Internet for the following reasons?

a. Used a social networking site such as Facebook, Twitter, YouTube, LinkedIn, Google Plus+, and Instagram for health information      1. Yes      2. No

b. Shared health information on social networking sites such as Facebook, Twitter, YouTube, LinkedIn, Google Plus+, and Instagram      1. Yes      2. No

c. Wrote in an online diary or blog about health information      1. Yes      2. No

- d. Participated in an online forum or support group for people with similar health issue    1. Yes    2. No
- e. Watched a health-related video on YouTube    1. Yes    2. No
10. Which social media have you used **MOST** to receive or share health information?
1. Never used    2. Facebook    3. Twitter    4. YouTube
5. LinkedIn    6. Google Plus+    7. Instagram    8. Other

## Part IV: Behavioural Characteristics

1. In a typical week, how many days do you do any physical activity or exercise of at least moderate intensity, such as brisk walking, bicycling at a regular pace, and swimming at a regular pace? \_\_\_\_\_ days per week
2. On the days that you do any physical activity or exercise of at least moderate intensity, how long do you typically do these activities? \_\_\_\_\_ hours \_\_\_\_\_ minutes
3. Have you ever smoke cigarettes?
  1. Never Smoked
  2. Ex-smoker
  3. Occasional Smoker
  4. Daily Smoker
4. Have you ever smoke electronic cigarettes?
  1. Never Smoked
  2. Ex-smoker
  3. Occasional Smoker
  4. Daily Smoker
5. Have you ever smoke shisha?
  1. Never Smoked
  2. Ex-smoker
  3. Occasional Smoker
  4. Daily Smoker
6. Have you ever drunk alcohol?
  1. Never drunk
  2. Ex-drinker
  3. Occasional drinker (less than once per month)
  4. Monthly drinker (1 – 3 times per month)
  5. Weekly drinker (at least once per week)

## Part V: Psychological characteristics

1. How concerned are you about getting any health problem?
1. Very Concerned                      2. Concerned                      3. Not at all concerned
2. How getting a health problem would affect your life?
1. Make me very sick                      2. Sick                      3. Not a big deal

3. Overall, how confident are you that you could get advice or information about health if you needed it?

1. Completely confident      2. Very confident      3. Somewhat confident  
4. A little confident      5. Not confident at all

## Part VI: Health-related information

1. In general, would you say your health is:

1. Excellent      2. Very good      4. Good      3. Fair      5. Poor      6. I don't know

2. Do you live with any of the following health problems or conditions?

	<u>Yes, currently</u>	<u>Yes, in the past 12 months</u>	<u>Not at all</u>
1. Diabetes	1	2	3
2. High blood pressure	1	2	3
3. Asthma and lung conditions	1	2	3
4. Heart disease	1	2	3
5. Cancer	1	2	3
6. Other health problems, specify	1	2	3

3. Do your parents live with any of the following health problems or conditions?

	<u>Yes, currently</u>	<u>Yes, in the past 12 months</u>	<u>Not at all</u>
1. Diabetes	1	2	3
2. High blood pressure	1	2	3
3. Asthma and lung conditions	1	2	3
4. Heart disease	1	2	3
5. Cancer	1	2	3
6. Other health problems, specify	1	2	3

4. Health literacy – on a scale from very easy to very difficult, please fill the following

Item No	On a scale from very easy to very difficult, how easy would you say it is to:	Very difficult	Difficult	Easy	Very easy	Don't know
1.	find information about symptoms of illnesses that concern you?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.	find information on treatments of illnesses that concern you?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3.	find out what to do in case of a medical emergency?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4.	find out where to get professional help when you are ill?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5.	understand what your doctor says to you?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6.	understand the leaflets that come with your medicine?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7.	understand what to do in a medical emergency?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8.	understand your doctor's or pharmacist's instruction on how to take prescribed medicine?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9.	judge how information from your doctor applies to you?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10.	judge the advantages and disadvantages of different treatment options?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

11.	judge when you may need to get a second opinion from another doctor?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
12.	judge if the information about illness in the media is reliable?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
13.	use information the doctor gives you to make decisions about your illness?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14.	follow the instructions on medication?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
15.	call an ambulance in an emergency?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
16.	follow instructions from your doctor or pharmacist?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
17.	find information about how to manage unhealthy behaviour such as smoking, low physical activity and drinking too much?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
18.	find information on how to manage mental health problems like stress or depression?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
19.	find information about vaccinations and health screenings that you should have?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
20.	find information on how to prevent or manage conditions like being	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	overweight, high blood pressure or high cholesterol?					
21.	understand health warnings about behaviour such as smoking, low physical activity and drinking too much?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
22.	understand why you need vaccinations?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
23.	understand why you need health screenings?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
24.	judge how reliable health warnings are, such as smoking, low physical activity and drinking too much?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
25.	judge when you need to go to a doctor for a check-up?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
26.	judge which vaccinations you may need?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
27.	judge which health screenings you should have?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
28.	judge if the information on health risks in the media is reliable?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
29.	decide if you should have a flu vaccination?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
30.	decide how you can protect yourself from illness based on advice from family and friends?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

31.	decide how you can protect yourself from illness based on information in the media?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
32.	find information on healthy activities such as exercise, healthy food and nutrition?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
33.	find out about activities that are good for your mental well-being?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
34.	find information on how your neighbourhood could be more health friendly?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
35.	find out about political changes that may affect health?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
36.	find out about efforts to promote your health at work?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
37.	understand advice on health from family members or friends?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
38.	understand information on food packaging?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
39.	understand information in the media on how to get healthier?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
40.	understand information on how to keep your mind healthy?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
41.	judge where your life affects your health and well-being?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



42.	judge how your housing conditions help you to stay healthy?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
43.	judge which everyday behaviour is related to your health?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
44.	make decisions to improve your health?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
45.	join a sports club or exercise class if you want to?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
46.	influence your living conditions that affect your health and wellbeing?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
47.	take part in activities that improve health and well-being in your community?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## Appendix C: Simulated Situation

### Task I

**Simulated situation:** Imagine that your housemate has recently been diagnosed with Dengue fever. You want to help your housemate by finding as much information about the disease as possible on any website you want.

**Indicative Request:** Find information on how your housemate possibly experienced the disease. Indicate the signs and symptoms, and treatment of the disease.

### Task II

**Simulated situation:** In your formal education, you have learnt that smoking is one of the risk factor for cancer, diabetes, heart disease, stroke, and chronic respiratory diseases which account for more than 80% of premature deaths in developing countries and 73% of total deaths in Malaysia. Smoking is a risk factor responsible for about 10,000 Malaysians deaths each year. Assume you have a friend who smokes 14 cigarettes per day and you are concerned about your friend. You would like to provide some information for your friend.

**Indicative Request:** Find information about treatments and guidelines to help your friend to quit smoking.

### Task III

**Simulated situation:** Assume you have read the following information in a magazine: *“cardiovascular diseases (heart and blood vessel diseases) are the number one cause of death globally. Obesity is one of the major risk factors for diseases. In 2014, globally, 600 million people of age 18 years and above were obese”*. Assume you have a best friend whom you think is having an obesity problem. You are concerned about your friend’s body weight and want to provide information on how to have a healthy body weight.

**Indicative Request:** Find information about the classification of bodyweight, and the measures or guidelines to help your friend to have a healthy body weight.

#### **Task IV**

**Simulated situation:** Imagine last night you listened to a radio program about physical inactivity. You have heard that physical inactivity is a key risk factor for cancer, diabetes, heart disease, stroke, and chronic respiratory diseases. The radio program also pointed out that more than 80% of the world's adolescent population is insufficiently physically active. In comparison, you also heard that sufficient physical activity gives benefits of reduced stress, improved sleep, and lead to a better quality of life. You have noted the benefit of physical activity at the same time you are concerned that you are physically inactive. You want to find out proper guidelines or recommendations on physical activity.

**Indicative Request:** Find information about guidelines or recommendations to help you to be sufficiently physically active.

## Appendix D: Observational Checklist

1. File name of the screen recorded \_\_\_\_\_
2. Task performed      1. Task 1 (Dengue)   Task 2 (Smoking)   3. Task 3 (Obesity)  
   4. Task 4 (Physical activity)      5. Task 5 (Personal Task)
3. Query issued to search engine \_\_\_\_\_
4. Type of query      1. Question-based   2. Keyword-based   3. Run together words based
5. How query formulated?
  1. Query is not from simulated situation
  2. Whole query is from Simulated situation
  3. Keyword from simulated situation
6. Query length \_\_\_\_\_
7. Keywords \_\_\_\_\_
8. Health terms \_\_\_\_\_
9. Stop words \_\_\_\_\_
10. Operators \_\_\_\_\_
11. Spelling error 1. Yes      2. No      3. Query suggestion used
12. Is dropped down query suggestion clicked
  1. Query not suggested      2. Clicked      3. Not clicked
13. Query reformulation that changes the form of query
  1. Reform to question   2. Change term sequence   3. Spelling correction
14. Query reformulation pattern
  1. Specification   2. Generalization   3. Parallel movement   4. Switch to new concepts
15. Result clicked?      1. Yes      2. No
16. What search engine is used? \_\_\_\_\_
17. What websites are viewed? \_\_\_\_\_
18. Result clicked page \_\_\_\_\_

## Appendix E: Post-experiment Survey

### Task 1: Dengue Fever

1. Have you ever search information about task 1 (about Dengue Fever) before?  
1. Yes    2. Somewhat    3. No
2. How easy was task 1 (about Dengue Fever) to perform?  
1. Very easy    2. Easy    3. Neutral    4. Difficult    5. Very difficult
3. How satisfied are you with regard of getting the search results of task 1 (about Dengue Fever)  
1. Very disappointed    2. Unsatisfied  
3. Neutral    4. Satisfied    5. Very satisfied
4. Please explain why you are satisfied OR NOT satisfied with search results of task 1 (about Dengue Fever) \_\_\_\_\_
5. How successful are you in the search session of task 1 (about Dengue Fever)  
1. Very successful    2. Successful    3. Neutral  
4. Unsuccessful    5. Very disappointed
6. Please explain why your search session is successful OR NOT successful  
\_\_\_\_\_

### Task 2: Smoking

7. Have you ever search information about task 2 (about smoking) before?    1. Yes  
2. Somewhat    3. No
8. How easy was task 2 (about smoking) to perform?  
1. Very easy    2. Easy    3. Neutral    4. Difficult    5. Very difficult
9. How satisfied are you with regard of getting the search results of task 2 (about smoking)  
1. Very disappointed    2. Unsatisfied  
3. Neutral    4. Satisfied    5. Very satisfied
10. Please explain why you are satisfied OR NOT satisfied with search results of task 2 (about smoking) \_\_\_\_\_
11. How successful are you in the search session of task 2 (about smoking)  
1. Very successful    2. Successful    3. Neutral  
4. Unsuccessful    5. Very disappointed
12. Please explain why your search session is successful OR NOT successful  
\_\_\_\_\_

### **Task 3: Obesity**

13. Have you ever search information about task 3 (about obesity) before? 1. Yes 2. Somewhat 3. No

14. How easy was task 3 (about obesity) to perform?

1. Very easy 2. Easy 3. Neutral 4. Difficult 5. Very difficult

15. How satisfied are you with regard of getting the search results of task 3 (about obesity) 1. Very disappointed 2. Unsatisfied

3. Neutral 4. Satisfied 5. Very satisfied

16. Please explain why you are satisfied OR NOT satisfied with search results of task 3 (about obesity) \_\_\_\_\_

17. How successful are you in the search session of task 3 (about obesity)

1. Very successful 2. Successful 3. Neutral

4. Unsuccessful 5. Very disappointed

18. Please explain why your search session is successful OR NOT successful

---

### **Task 4: Physical Activity**

19. Have you ever search information about task 4 (about physical activity) before? 1. Yes 2. Somewhat 3. No

20. How easy was task 4 (about physical activity) to perform?

1. Very easy 2. Easy 3. Neutral 4. Difficult 5. Very difficult

21. How satisfied are you with regard of getting the search results of task 4 (about physical activity) 1. Very disappointed 2. Unsatisfied

3. Neutral 4. Satisfied 5. Very satisfied

22. Please explain why you are satisfied OR NOT satisfied with search results of task 4 (about physical activity) \_\_\_\_\_

23. How successful are you in the search session of task 4 (about physical activity)

1. Very successful 2. Successful 3. Neutral

4. Unsuccessful 5. Very disappointed

24. Please explain why your search session is successful OR NOT successful

## Personal Task

25. What is your personal task about? Can you provide a brief description about what you want search for? \_\_\_\_\_

26. How easy was your personal task to perform?

1. Very easy      2. Easy      3. Neutral      4. Difficult      5. Very difficult

27. How satisfied are you with regard of getting the search results of your personal task?

1. Very disappointed      2. Unsatisfied  
3. Neutral      4. Satisfied      5. Very satisfied

28. Please explain why you are satisfied OR NOT satisfied with search results of your personal task \_\_\_\_\_

29. How successful are you in the search session of your personal task

1. Very successful      2. Successful      3. Neutral  
4. Unsuccessful      5. Very disappointed

30. Please explain why your search session is successful OR NOT successful \_\_\_\_\_

## Overall search experience and challenges

31. How satisfied are you on the overall search experience?

1. Very disappointed      2. Unsatisfied      3. Neutral      4. Satisfied      5. Very satisfied

32. What are the challenges you faced when performing the tasks?

	Strongly <u>agree</u>	Agree	Somewhat <u>agree</u>	Disagree	Strongly <u>disagree</u>
a. Time consuming to search volumes of data to find required information	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b. Not able to obtain information immediately	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c. Too much/excessive irrelevant information	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d. Lack of information/inability to find what you are looking for online	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
e. Low credibility of the information sources	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
f. Difficulty to understand the information	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
g. Difficulty to formulate search tasks to query	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
h. Difficult to judge the quality (reliability and accuracy) of the information	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
i. Felt frustrated during searching for online health information	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

33. What other challenges did you face while performing the tasks? Please specify \_\_\_\_\_

## **Appendix F: Explanatory Statement**

### **Research Title: Health Information Seeking and Searching Behaviours of University Students: A Case from Two Developing Countries**

**Mr. Tesfahun Melese Yilma**  
**(Chief Investigator & PhD Student)**  
School of Information Technology  
Phone: +60176107554  
email: tesfahun.melese@monash.edu

**Dr Anushia Inthiran**  
**(Co-Investigator)**  
School of Information Technology  
Phone: +60355146090  
email: anushia.inthiran@monash.edu

**Prof. Daniel Reidpath**  
**(Co-Investigator)**  
School of Medicine and Health  
Sciences Phone: +603 5514 4962  
email: daniel.reidpath@monash.edu

**Dr Sylvester Olubolu Orimaye**  
**(Co-Investigator)**  
School of Information Technology  
Phone: + 60355145813  
email: sylvester.orimaye@monash.edu

You are invited to take part in this study which is conducted at Monash University Malaysia for a PhD. Please read this Explanatory Statement in full before deciding whether or not to participate in this research. If you would like further information regarding any aspect of this research, you are encouraged to contact the researchers via the phone numbers or email addresses listed above.

#### **What does the research involve?**

The aim of this study is to investigate health information seeking and searching behaviours of university students from developing countries and identify its associated factors. Understanding how university students acquire health information may eventually have significant implications for the design of effective public health communications that are geared towards the university population and young adults. Furthermore, university students' health information searching behaviours will be explored so as to better understand and suggest design strategies to improve health information retrieval systems.



This study involves a self-administered survey, information search session, and post-search survey. The total time taken for the entire process will not exceed the duration of 45 minutes. Also, there is no inconvenience or harm related to you throughout your participation in this research.

### **Consenting to participate in the project and withdrawing from the research**

You are chosen for this research because you have indicated that you are willing to take part in this study during a personal communication made either at a classroom, cafeteria, lobby, or students' recreational areas. Your participation is voluntary and you are under no obligation to consent to participate. However, if you consent to participate, you may only withdraw prior to the start of the data collection process.

### **Confidentiality**

Your name will not be associated with any information that you provide. It is not possible for you to be identified in any subsequent publications or presentations.

### **Storage of data**

The survey data will be collected using Google Forms. Whereas, observation data will be kept in hard drive. Storage of the data collected will be adhered to the Monash University's regulations and is kept on the University's premises in a locked cupboard/filing cabinet for a total duration of 5 years.

### **Use of data for other purposes**

Reports of the study may be submitted for publications, individual participants will not be identifiable in such reports. No information will be released to any other parties.

### **Results**

If you would like to be informed of the aggregate research finding, please contact Tesfahun Melese Yilma at [tesfahun.melese@monash.edu](mailto:tesfahun.melese@monash.edu).

If you would like to contact the researchers about any aspect of this study, please contact the investigator:

Mr Tesfahun Melese Yilma

Tel: +60176107554

Email: [tesfahun.melese@monash.edu](mailto:tesfahun.melese@monash.edu)

Project Number: CF15/4245 – 2015001820

### **Complaints**

Should you have any concerns or complaints about the conduct of the project, you are welcome to contact the Executive Officer, Monash University Human Research Ethics (MUHREC):

Executive Officer

Monash University Human Research Ethics Committee (MUHREC)

Campus Research Management

Monash University Malaysia

Jalan Lagoon Selatan

47500 Bandar Sunway

Selangor Darul Ehsan

Malaysia

Tel: +6(03) 5514 6000 Email: [researchoffice@monash.edu](mailto:researchoffice@monash.edu) Fax: +6(03) 5514 6001

Thank you

Tesfahun Melese Yilma



## Appendix G: Consent Form

**Research Title: Health Information Seeking and Searching Behaviours of University Students: A Case from Two Developing Countries**

**Chief Investigator: Mr Tesfahun Melese Yilma**

I have been asked to take part in the Monash University research project specified above. I have read and understood the Explanatory Statement and I hereby consent to participate in this project.

<b>I consent to the following:</b>	<b>Yes</b>	<b>No</b>
I agree to participate in the survey of health information seeking behaviour	<input type="checkbox"/>	<input type="checkbox"/>
I agree for activities to be logged and observed during health information searching	<input type="checkbox"/>	<input type="checkbox"/>
I agree to participate in the post-search survey after searching for health information	<input type="checkbox"/>	<input type="checkbox"/>

**Participant Name**

**Participant Signature**

**Date**

## **Appendix H: Instruction Letter**

Dear participants,

Thank you for your willingness to participate in this study which aims to explore health information seeking and searching behaviours of university students from developing countries.

This study has three parts: Online survey, experiment, and post-experiment survey. In the first part, you are required to fill up the online survey. Then you will be provided with four simulated situations which explain conditions that invite you to search for information on any website you want. This experiment also requires you to search for information on your own personal task which needs to be defined based on your health information need. You may stop searching after you get the required information or unable to locate the information. Upon completion of the experiment, you are required to fill up the post-experiment survey. Please note that do not come back to the experiment (search session) after you start filling the post-experiment survey for a specific simulated task.

Thank you.

With regards,

Tesfahun Melese Yilma