

Mobile Self-Triage Applications: A Usability Perspective

Mir Kamyar Ziabari

Thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science in Systems Science



uOttawa

University of Ottawa

Ottawa, Ontario, Canada

August 2018

© Mir Kamyar Ziabari, Ottawa, Canada, 2018

Abstract

Each person faces a health problem differently. Many people tend to visit an emergency department immediately without any regards to the scale of acuity of their disorder. Some of these visits are not necessary and hence lead to overcrowding of emergency departments, increase patients' waiting time, and ultimately lead to overutilization of healthcare resources. At the other end of the spectrum, some people do not visit their physician when they should. One solution to this problem is to provide a reliable source of information that would allow people to assess their health condition (self-triage) in a way that helps reduce the number of inappropriate visits and promote warranted visits to emergency departments.

This research proposes a new and user-friendly self-triage mobile application called *Symptoms Pal*, which allows lay people to use their health data to assess the presence of potential health concerns. Using diagnostic algorithms developed by physicians on the basis of medical evidence, the application requires users to answer a short series of questions about their symptoms. Symptoms Pal then provides some basic information about the possible problem and aims to inform users about whether they should seek immediate medical attention such as visiting an emergency department, or see a family doctor, or do self-care. The usability of Symptoms Pal was evaluated by conducting a study involving 34 participants. Several strengths and weaknesses of the usability and perceived usefulness of the application were identified and led to a revised version and additional recommendations.

The thesis contributes 25 reusable requirements and validated user interface design artefacts for self-triage mobile applications.

Acknowledgment

First, I would like to express my deepest gratitude toward my co-supervisor, Professor Daniel Amyot, for his unconditional and untiring help and support. I would not have been able to finish this work without his patience, support, and guidance. The experience of working with Professor Amyot not only made me a better researcher, but his attitude and manners taught me how to try every day and keep up with good energy. I will always remain thankful to him for giving me this experience.

I also want to thank my other co-supervisor, Professor Wojtek Michalowski. He was supportive and his guidance was always helping me move forward, challenge myself, and also learn. Working with him was a fantastic experience over the past two years.

I would like to thank Dr. Nassoh Hafez and Dr. El Mostafa Bouattane, from Montfort Hospital, for giving me the opportunity to work on this project and for allowing me to work in a healthcare environment, which was my desire.

A special thanks to Prof. Natalie Baddour and the NSERC CREATE funded “Bio-medical Engineering Smartphone Training” program (CREATE-BEST) for funding this project, and to Dr. Karen Fung-Kee-Fung for providing me funding during my Master’s program.

I should also thank the family that I chose. To Aynur, Kiana, Koosha, Pirouz, Pourya, and Setareh, my loved ones who helped me a lot back home while I was moving out of my country to come to Canada. I will never forget your love and true friendship. I am happy that every time I feel overwhelmed I have you people to turn to, although each one of you is living in a different corner of this planet.

Many thanks to my neighbour and friend Sepideh for helping me a lot in the past two years. Sincere thanks to Malak Baslyman, for helping me many times in various situations and patiently guiding me towards making the right decisions.

I thank my brothers and role models, Amir Koushyar and Amir Kavyan, for helping me keep up, teaching me how to love, and being there for me during the past 26 years. Their love and support always encourage me to do my best and be a better person. I also

thank my father Farokh and my mother Sonia. My father was the one who taught me to always learn and my mother taught me to always love and always give back to the community. Today, wherever I am, is because of my family's unconditional love.

Whatever I do in this life is to make this planet a better place for the future generation.

Dedicated to all the troubled kids in the world.

I am always worried about you.

Table of Contents

Abstract	ii
Acknowledgment	iii
Table of Contents	vi
List of Figures	ix
List of Tables	x
List of Acronyms	xi
Chapter 1 Introduction	1
1.1 Motivation	1
1.2 Problem Context	2
1.3 Self-Triage Applications	3
1.4 Pros and Cons of Self-Triage Applications.....	5
1.5 Research Objectives and Thesis Goals	6
1.6 Methodology	6
1.7 Thesis Contributions	10
1.8 Thesis Outline	10
Chapter 2 Literature Review	12
2.1 Literature Review Methodology	12
2.2 Finding Keywords.....	12
2.3 Research on Symptom Checkers.....	13
2.4 Search Process.....	14
2.4.1 Diagnostic Accuracy.....	14
2.4.2 User Interface Design and Usability Evaluation.....	16
2.4.3 Critical Analysis	17
2.5 Summarization and Discussion.....	17
2.6 Chapter Summary.....	19
Chapter 3 Review of Commercial Applications	20
3.1 Collection of Applications	20
3.2 Description of the Applications	22
3.3 Comparison	28
3.4 Chapter Summary.....	32
Chapter 4 Application Design and Implementation	33
4.1 Goals and Requirements	33
4.2 Architecture.....	35
4.3 Data Layer	36
4.4 Logic Layer	38

4.4.1	Application Class.....	38
4.4.2	Data Structures	39
4.4.3	Activities.....	41
4.5	Presentation Layer.....	42
4.6	Interfaces of the Application.....	42
4.6.1	Language Page.....	43
4.6.2	Start Page and Symptom List	44
4.6.3	Question Page	46
4.6.4	Result Page	47
4.7	Chapter Summary.....	50
Chapter 5	Application Evaluation	51
5.1	Objective	51
5.2	Usability Study Design	51
5.3	Recruitment	53
5.4	Participation	54
5.5	Demographics	55
5.6	Predefined User Scenarios	55
5.7	Task Result Analysis.....	57
5.7.1	Task 1 – Symptom Selection	57
5.7.2	Task 2 – Answer Medical Questions	57
5.7.3	Task 3 – Receive Map Directions.....	58
5.8	Questionnaire	59
5.9	Usability Barriers	65
5.10	Findings from Participants	69
5.10.1	Perceived Usefulness.....	69
5.10.2	Additional Findings.....	70
5.11	Proposed Modifications	73
5.12	Chapter Summary.....	77
Chapter 6	Discussion.....	79
6.1	Technical Comparison	79
6.2	Limitations and Threats to Validity	80
6.2.1	Construct Validity.....	81
6.2.2	Internal Validity.....	81
6.2.3	External Validity.....	81
6.3	Chapter Summary.....	82
Chapter 7	Conclusion and Future Work	83
7.1	Conclusion	83
7.2	Future Work	86
References	87
Appendix A:	Email to Participants	92
Appendix B:	Consent Form	94
Appendix C:	Pre-Test Interview.....	97
Appendix D:	Study Protocol	98

Appendix E: Task List.....	100
Appendix F: Questionnaire.....	101
Appendix G: User Scenarios.....	106

List of Figures

Figure 1 - Examples a) of query-based UI b) selection-based UI (right)	4
Figure 2 - Design-science research guidelines [28].....	7
Figure 3 - Thesis methodology	9
Figure 4 - Methodology steps used while conducting the literature review	13
Figure 5 - Ada symptom checker.....	23
Figure 6 - Isabel symptom checker.....	25
Figure 7 - Mediktor symptom checker.....	26
Figure 8 - Your.MD symptom checker ChatBot	28
Figure 9 - Symptoms Pal architecture.....	36
Figure 10 - Overview of data included in decision trees (algorithms)	37
Figure 11 - Illustrative example – decision tree for fever.....	38
Figure 12 - Tree structure	40
Figure 13 - UML class diagram of the data structures.....	40
Figure 14 - Interfaces of Symptoms Pal	43
Figure 15 - a) Language page, b) Legal notice and disclaimer.....	44
Figure 16 - a) Woman pictogram, b) Man pictogram with head zone selected.....	45
Figure 17 - a) Zonal symptoms list for head & neck, b) Common symptom list ..	45
Figure 18 - a) Symptoms list b) Queried symptoms list	46
Figure 19 - a) Question page, b) Progress bar	47
Figure 20 - a) Emergency/Call911 b) Emergency room.....	48
Figure 21 - a) Visit a doctor, b) Self-Care	49
Figure 22 - a) Emergency rooms in Google Maps, b) Walk-in clinics in Google Maps.....	49
Figure 23 - Predefined user scenarios	56
Figure 24 - Questionnaire results for questions 1 to 4.....	59
Figure 25 - Questionnaire results for questions 6 to 10.....	61
Figure 26 - Questionnaire results for questions 11 and 13	62
Figure 27 - Questionnaire results for questions 12 and 14	62
Figure 28 - Questionnaire results for question 5.....	63
Figure 29 - Questionnaire results on start page UI (Appendix F, question 20).....	64
Figure 30 - Satisfaction scores	65
Figure 31 - a) Common symptoms interface, b) Modified start page.....	75
Figure 32 - a) Old question page, b) Modified question page	76
Figure 33 - a) Old result page, b) Modified result page.....	76
Figure 34 - a) Self-Care color coded in green, b) Emergency Room color coded in red.....	77

List of Tables

Table 1 - Summary of related works.....	18
Table 2 - Selected commercial mobile applications	21
Table 3 - Comparison of existing applications	31
Table 4 - Primary goals of the application.....	33
Table 5 - Requirements of the self-triage application.....	35
Table 6 - Application activities.....	41
Table 7 - Application design issues	66
Table 8 - Application clinical content issues	67
Table 9 - Users' preferred number of questions to ask.....	71
Table 10 - Additional requirements	73
Table 11 - Comparison with other applications.....	80

List of Acronyms

Acronym	Definition
APP	Application
CSV	Comma-Separated Value
DSRM	Design Science Research Methodology
EBSE	Evidence-based Software Engineering
ED	Emergency Department
ENT	Ear, nose, throat
GUI	Graphical User Interface
IDE	Integrated Development Environment
NHS	National Health Service
SP	Symptoms Pal
UI	User Interface
XML	Extensible Markup Language

Chapter 1 Introduction

One major feature of the healthcare system in Canada is the availability of health services to the population regardless of their income and employment status. However, despite this high level of availability, there are drawbacks in this system that affect the overall quality of the care being provided to people. One of the most important drawback is long waiting times resulting in part from inappropriate visits to Emergency Departments (ED) in hospitals.

This thesis introduces and studies the usability of a *self-triage mobile application*. This application, called *Symptoms Pal*, intends to inform any users of this application, whether they become patients or not, about the potential severity of their health problem and redirect them to the most suitable healthcare facility when needed. In this study, a self-triage application is a *symptom checker* whose main focus is to provide accurate triage advice. The purpose of such an application is to reduce the high number of inappropriate visits to ED, to expedite warranted visits, and ultimately to improve the experience of users during their healthcare journey. This chapter includes the motivation of the thesis, the problem context, the thesis goals, important background on self-triage applications, and a presentation of the selected research methodology. An outline of the rest of the thesis concludes this chapter.

1.1 Motivation

People visit an ED for various reasons and there are many cases where such visits are not justified. Van den Heede and Van de Voorde observed that “ED attendances for conditions that do not require urgent attention or specialized input would be called *inappropriate visits*” [68]. These inappropriate visits have negative influences on a healthcare service as well as on people who are visiting EDs. For example, inappropriate visits will increase waiting times for triage and examination, and they will in some cases severely affect people

who actually need to urgently consult with a medical expert. One of the possible explanations behind these inappropriate visits are visitor's lack of certainty due to the unavailability of reliable information. Hence, the need to provide reliable information makes a *self-triage application* an appropriate intervention method that could be helpful to people. A properly designed and usable *self-triage mobile application* can proactively address uncertainty to some extent, as it will help users to assess severity of a health condition before visiting EDs and will ultimately direct users to the appropriate care providers (e.g., family physicians or walk-in clinics) or even provide them with the information necessary for self-care. The introduction of a usable self-triage mobile application is the main motivation behind this research.

The lack of research on the topic of *symptom checkers* (i.e., systems that allow the evidence-based detection of a medical abnormality or disorder), usability of *symptom checkers*, and particularly usability of *self-triage mobile applications*, was another motivation for this study. There are hundreds of mobile (and web-based) applications on the market, and yet no study was found that systematically discussed the design, development, and usability evaluation of a *self-triage application*. In addition, there are few research papers discussing the different features and necessary requirements of these mobile applications. Last but not least, working in a real healthcare environment was another motivation for this thesis. This study is resulting from a collaborative project conducted by the University of Ottawa and Montfort Hospital in Ottawa, Canada.

1.2 Problem Context

One of the standing issues existing in EDs in Canada is overcrowding [3]. An *overcrowded ED* can be defined as a situation where there is more demand for healthcare services than the ED's capability to serve people with suitable care within a reasonable time. There are multiple impacts of overcrowding that can directly affect patients and healthcare providers, and Derlet [13] describes four important ones:

1. **Long waiting time:** One of the most common negative influences of an overcrowded ED are delays related to patient care, which have the potential to make a minor medical problem more serious.

2. **Extended pain and suffering:** One of the results of increasing waiting time due to overcrowding is that patients' pain and suffering may be extended because they are not being taken care of.
3. **Losing "golden hours":** Patients with severe conditions, such as strokes, may miss the "golden hour" when the treatment provided is most effective.
4. **Increased errors in treatment:** Increased waiting time leads to practitioners feeling rushed, and this time pressure can increase errors in judgment, diagnoses, and eventually treatment processes.

Inappropriate visits are one of the most important factors affecting overcrowding in EDs. Afilalo et al. reported, based on a study from 2004 [4], that 25% of visits to the ED in Canada are considered inappropriate. People without a regular physician or a regular source of care, or who are not referred to the ED by a physician, were shown to have more inappropriate ED use [12]. In addition to ED, these visits are decreasing the quality of primary care and increasing the pressure on physicians as well. Hammond [22] reported, in a study conducted in England in 2004, that about 50% of the visits to general practitioners' offices were unnecessary and that up to 70% were minor health problems. In such contexts, effective triage is necessary to deal with cases that cannot wait, assign people with non-urgent conditions to healthcare facilities other than EDs, and provide appropriate self-care suggestions for minor health problems that do not require a physician consult.

1.3 Self-Triage Applications

People are increasingly querying the Internet to study their health conditions. In 2013, Fox and Duggan reported that about 72% of Internet users experienced looking online for health information [19]. The use of search engines such as Google for evaluating either non-urgent symptoms (such as a cold) or urgent conditions (such as chest pain) is one of the most common self-diagnosis approaches that exploit the Internet. However, despite the wealth of information existing on the web, such searches can often produce disorienting and unverified information that can confuse users [50].

Another popular resource used to obtain health information are *symptom checker applications* [19]. Symptom checkers are systems that allow the evidence-based detection

of a medical abnormality or disorder. The first symptom checkers reported were websites and software distributed on CD-ROMs [7], but these applications were unable to gain much trust among their users. However, with an increasing pressure on primary care and the fast expansion of the Internet, health organizations as well as lay people started investing in better versions of these tools.

A symptom checker requires its users to input their symptoms and/or answer a series of questions about their physical signs. There are two common user interface (UI) patterns that applications employ in order to allow users to enter their symptoms: i) a *query-based* interface that provides users with the ability to write their symptoms in the application, and ii) a *selection-based* interface that offers a list of various symptoms and diseases known to the application and asks users to select their main observable symptom. Figure 1 shows examples of query-based and selection-based user interfaces.

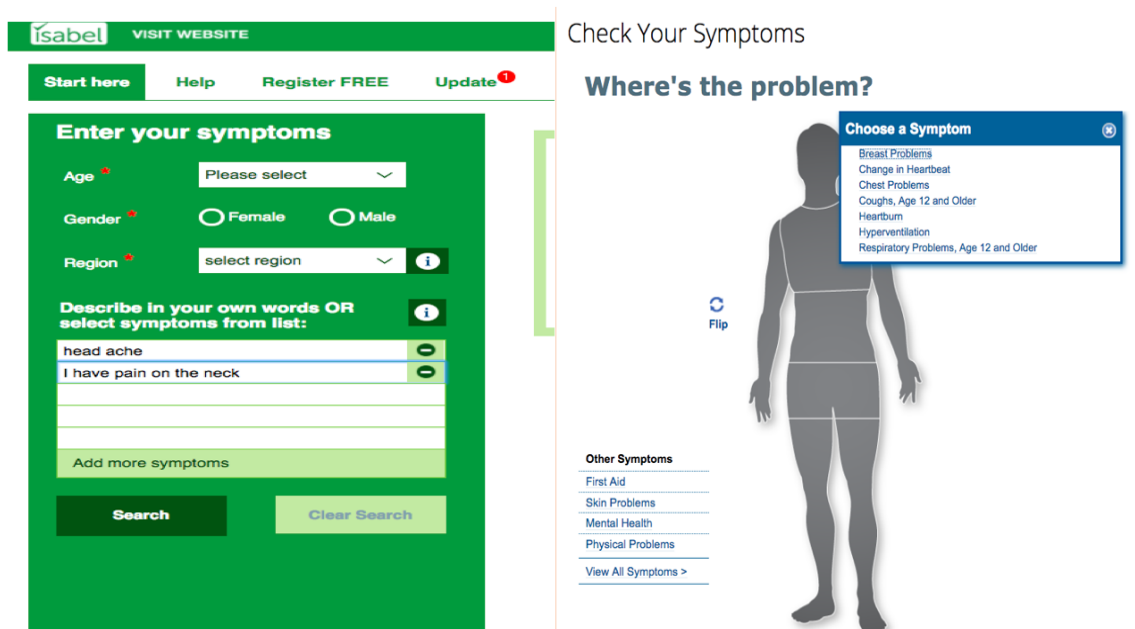


Figure 1 - Examples a) of query-based UI b) selection-based UI (right)

Following the evaluation of user inputs, a symptom checker provides its users with feedback such as self-diagnosis and/or self-triage suggestions. In this research, symptom checkers whose main focus is to provide self-triage advice are called *self-triage applications*. Self-diagnosis feedback involves potential health problems, often ranked by likelihood of occurrence. Self-triage feedback informs users about the severity level of their (likely)

problem and whether they should seek care or not, for example by recommending the most suitable healthcare facility (ED, walk-in clinic, pharmacy, etc.). Symptom checkers, including self-triage applications, can be accessed through websites or with mobile applications, which can be downloaded from stores such as Google Play¹ and Apple's App Store², and then used on smartphones and tablets.

Many such software applications are using various decision algorithms such as decision trees, machine learning, branching logic, and Bayesian inference to infer and store decision knowledge. Symptom checker applications can be categorized into two types:

- **Static** applications store health decision knowledge statically in their data layer. Upon collection of a symptom and user answers to questions regarding the symptom, a static symptom checker provides feedback to the user. Such systems do not have the ability to learn and usually do not allow their users to enter or select multiple symptoms simultaneously.
- **Dynamic** applications have the ability to expand their initial knowledge base and learn new medical conditions from different sources of data, for instance by using machine learning techniques. Dynamic symptom checkers often use natural language processing in order to find relationships between a user's text inputs and diseases or other health conditions.

1.4 Pros and Cons of Self-Triage Applications

A well-designed and precise symptom checker or self-triage application, which can correctly inform and guide its users, can offer several benefits by:

- Expediting warranted visits to ED [59].
- Decreasing the pressure on healthcare care providers[44].
- Reducing user uncertainty about complex medical conditions [20].
- Increasing patient knowledge [20].
- Providing access to health care in low resources settings [45].

¹ <https://play.google.com>

² <https://www.appstore.com/>

At the same time, these applications can cause important problems to both medical practitioners and users, including:

- Hindering patient-doctor relationships [41].
- Misleading users by not correctly diagnosing them, and by causing them potentially severe negative financial and health effects [54].
- Providing users with an incorrect care suggestion, causing them to visit a healthcare facility that is not entirely suitable and hence negatively impacting physicians, nurses, and patients and their family members on numerous aspects such as time, costs, stress, and workload [54].
- Raising concerns regarding the privacy of the personal data (such as names, emails, symptoms, etc.) that would be collected by some of the symptom checkers [41][44].

1.5 Research Objectives and Thesis Goals

The main research objective in this thesis research is to determine the usability of a *mobile self-triage application* that evaluates multiple symptoms and provides its users with the opportunity to self-triage their condition following a recommendation on the most suitable healthcare option and closest healthcare facility (if necessary). In addition, the secondary objective of this research is to improve the understanding of the features and functionalities that users would like to see in such an application, in order to improve overall user experience. In view of these research objectives, the main goals of this thesis are to:

- 1- Identify technical requirements necessary to develop a usable self-triage mobile application.
- 2- Design and implement a prototype of such an application.
- 3- Conduct a usability study of the application.
- 4- Evaluate the results and improve the requirements and prototype accordingly.

1.6 Methodology

Nayak observes that it is necessary to have a proper methodology in order to organize research effectively [46]. The *design-science* paradigm, in particular, has its underlying

foundations in the engineering and science of artifacts [28]. This paradigm provides a framework to make developments that characterize ideas, practices, technical capabilities, and products through which the analysis, design, implementation, management, and use of information systems can be effectively and efficiently done. One of the most common design-science research methodologies in this research context is Hevner’s Design Science Methodology in Information System (IS) [28], which inspired the research methodology used in this thesis. Hevner provides a set of guidelines that are presented in Figure 2.

Guideline	Description
Guideline 1: Design as an Artifact	Design-science research must produce a viable artifact in the form of a construct, a model, a method, or an instantiation.
Guideline 2: Problem Relevance	The objective of design-science research is to develop technology-based solutions to important and relevant business problems.
Guideline 3: Design Evaluation	The utility, quality, and efficacy of a design artifact must be rigorously demonstrated via well-executed evaluation methods.
Guideline 4: Research Contributions	Effective design-science research must provide clear and verifiable contributions in the areas of the design artifact, design foundations, and/or design methodologies.
Guideline 5: Research Rigor	Design-science research relies upon the application of rigorous methods in both the construction and evaluation of the design artifact.
Guideline 6: Design as a Search Process	The search for an effective artifact requires utilizing available means to reach desired ends while satisfying laws in the problem environment.
Guideline 7: Communication of Research	Design-science research must be presented effectively both to technology-oriented as well as management-oriented audiences.

Figure 2 - Design-science research guidelines [28].

There are two guidelines that are central to the work described in this thesis, Design as an Artifact (#1), and Design Evaluation (#3). An artifact should be purposefully created to address an important organizational problem and must effectively enable its implementation and application in the appropriate domain. The artifact here is a *mobile application* with the purpose of *enabling self-triage* to its users, in order to help them in their decision-making process regarding a health concern. The evaluation is yet another crucial component of the research process. The utility, quality, and efficacy of a design artifact should be

demonstrated via well-executed evaluation methods. This thesis uses two complementary evaluation methods:

- The first one is descriptive (informed argument [28]). While developing the application, Nielsen's usability heuristics were followed in order to design seamlessly [49].
- The other is experimental [28]. To evaluate the artifact's learnability, efficiency, and user interface attributes, a usability study was designed and performed with different users.

Based on the design science methodology, this thesis methodology, shown in Figure 3, is composed of several steps:

- **Problem identification and motivation:** In this step, as discussed in sections 1.1 and 1.2, the need for a new *self-triage mobile application* artifact is identified (*Design as an Artifact*).
- **Identifying existing systems:** Various features, benefits and problems of similar mobile/web applications are studied in a literature survey done according to Kitchenham's approach [35]. A few of the surveyed problems are directly affecting the current state of symptom checker and self-triage applications, and they are hence adding to the main problem (*Problem Relevance*).
- **System requirements:** The requirements are gathered from stakeholders, the literature, and existing symptom checker and self-triage applications. In this step, existing artifacts and peer-reviewed papers are searched because it is required to utilize available means to discover an effective solution to the problem (*Design as a Search Process*).
- **System development:** This step focuses on the development and design of the user interface of a mobile application (Symptoms Pal) that aims to reduce users' uncertainty when facing symptoms of health problems (*Research Contributions*). The clinical content (decision algorithms with triage advice and other medical suggestions) of this application was not developed in this stage but was developed and provided by physicians from Montfort Hospital.

- System evaluation:** A usability study was conducted in order to understand the perception of users regarding the efficiency, learnability, intuitiveness, and ease of use (usability) of the suggested prototype, leading to new requirements and improvements of the prototype. The accuracy of the clinical suggestions is not evaluated in this stage as it was out of the context of this research. A comparison between features of this application with features of other existing applications is also provided (*Design Evaluation*). The descriptive statistical analysis of the data collected from the usability study and the comparison with existing systems (*Research Rigor*) allow us to provide conclusions about the perceived usefulness of such artifact (*Research Contributions*).
- Communication:** This step involves sharing the research results by writing this thesis, presenting the work (e.g., through posters), and eventually producing a publication (*Communication of Research*).

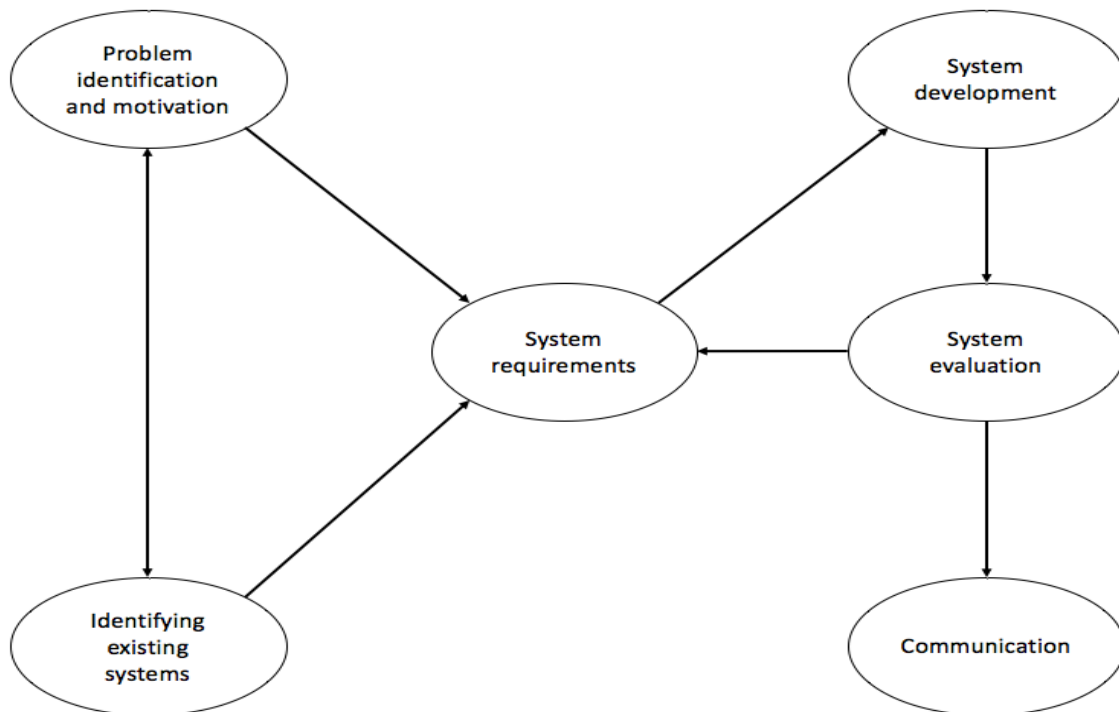


Figure 3 - Thesis methodology

1.7 Thesis Contributions

To satisfy the research objectives mentioned in section 1.5, a functional self-triage mobile application was designed and implemented, and a study involving 34 participants was conducted to assess the usability of this software application.

The main research contributions of this study include the (reusable) requirements, design, implementation, and usability assessment of a self-triage mobile application. The design was based on the requirements identified during several meetings with the stakeholders, based on an earlier prototype developed by Montfort Hospital, from the review of literature related to this topic, and from the review of related systems.

This work was showcased in two poster competitions, at the University of Ottawa and at the Knowledge Institute of Montfort Hospital, with positive informal feedback. In future work, and as a part of the methodology explained in the previous section, a publication will be produced to share the results of this study with other researchers.

1.8 Thesis Outline

The rest of this thesis is structured as follows:

- Chapter 2 presents a review of academic literature on symptom checkers, including self-triage applications.
- Chapter 3 provides a review of existing commercial symptom checkers existing on the market. Different features of these systems are reviewed and assessed based on measurable criteria.
- Chapter 4 first presents goals and requirements for the self-triage application (Symptoms Pal) contributed by this thesis. Then, it discusses the architecture, interfaces, and implementation of this application.
- Chapter 5 describes the usability study conducted with Symptoms Pal, from which new requirements and potential improvements are extracted. Some modifications are also made to the initial prototype application.
- Chapter 6 first provides a comparison between this application and closely-related systems and then discusses threats to the validity of this research.
- Chapter 7 concludes the thesis and presents important future work.

Appendices A to G describe the content of letters, protocols, and forms used for the usability study described in Chapter 5.

Chapter 2 Literature Review

This chapter presents a review of relevant academic papers. The purpose of this literature review is to present important background information and studies about symptom checkers and self-triage mobile applications. This review also aims to find gaps in existing research and enable the reuse of good requirements. A brief comparison of existing systems along four criteria is also provided.

2.1 Literature Review Methodology

As Kitchenham et al. have observed, empirical software researchers should consider evidence-based studies in their work and decision-making process [34]. This literature review's methodology is inspired from Kitchenham et al.'s framework: Evidence-based Software Engineering (EBSE). This review is not a fully systematic according to EBSE's definition, but it attempts to cover its main phases: Planning the Review, Conducting the Review, and Reporting the Review [35]. Figure 4 highlights the methodology steps followed here.

2.2 Finding Keywords

This review looks into *symptom checker mobile/web applications* that provide triage advice with or without the diagnosis. Generic keywords used for querying bibliographical databases include “symptom checker” and “self-triage”. Common technical and medical databases were queried in this review: Scopus (which covers IEEE Xplore, ACM Digital Library, Elsevier, Springer, and others), Web of Science, Medline (Ovid), PubMed and Google Scholar.

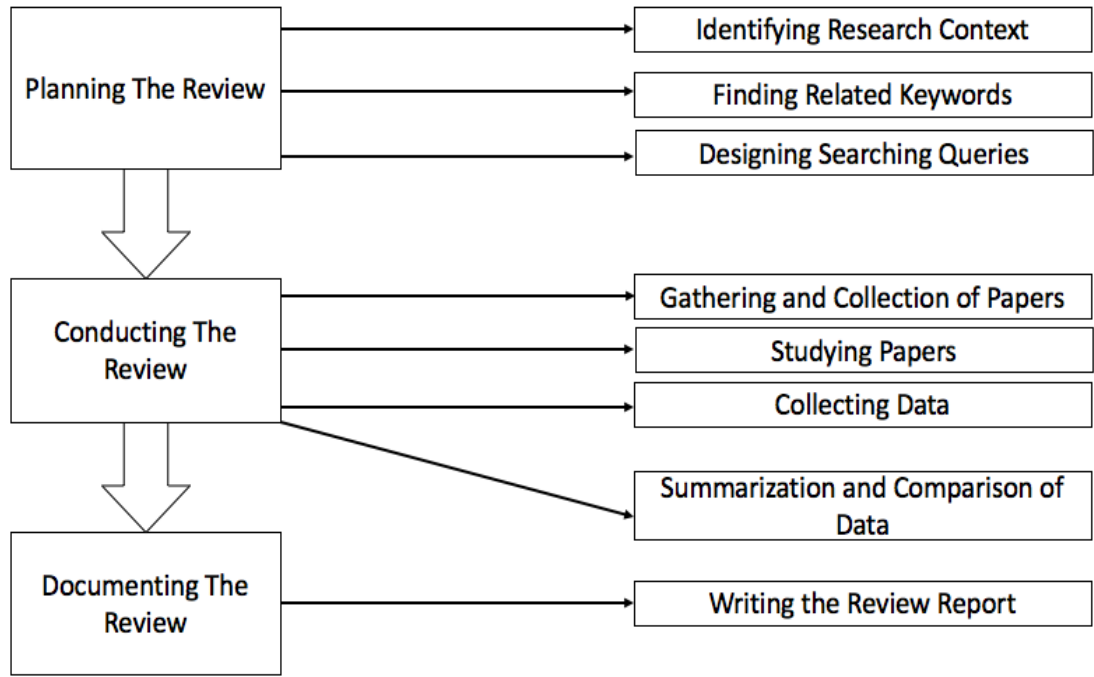


Figure 4 - Methodology steps used while conducting the literature review

2.3 Research on Symptom Checkers

The general query that was used to find research papers is:

`("symptom checker" OR ("self-triage" AND "app*"))`

All the selected databases were searched based on abstracts, titles, and keywords (except Google Scholar, where only the titles were searched as abstracts/keywords are not available to queries). No specific time limit was used. The query is general by design in order to find as many relevant papers as possible, at the risk of also finding many irrelevant ones.

As a result of the query, 71 unique papers were found. Many irrelevant papers were removed based on titles and abstracts through exclusion criteria: papers that discuss symptom checkers as part of a discussion on another research topic (e.g., mobile applications used in healthcare [8][60]), papers about applications that target only one specific type of symptom (e.g., chest pain [69] or knee pain [9]), papers that focus on a single type of users (e.g., older adults' experience on online diagnosis [40]), and papers that particularly discuss decision knowledge mechanisms of such applications [55][56]. One research paper [44] has the right scope but was excluded as it was written in Spanish.

2.4 Search Process

Out of the 71 papers found, only nine papers are related to symptom checkers/self-triage as their main focus and are included in our review. The categorization in this section is done based on their main discussion topic. These nine papers contribute to this study along three important axes:

- **Diagnostic accuracy** [18][21][53][58][59]: Papers that discuss diagnostic accuracy can help us understand the extent of reliability of these applications. This category also helps us understand what form of medical suggestion should be provided to users.
- **User interface design and usability evaluation** [1][38][43]: Papers that discuss usability and user interface design can help us understand the methods used to evaluate the usability of such applications. The information provided by these papers can also help us find out about features that are necessary and the ones that should be avoided.
- **Application barriers** [41]: Papers that discuss the barriers faced by applications can help us realize current shortcomings of such applications and how these applications can be improved.

2.4.1 Diagnostic Accuracy

In **2011**, Farmer et al. [18] reported the accuracy of the WebMD symptom checker [70] in diagnosing ENT (ear, nose, throat) symptoms. The data were collected from 61 patients that visited a specialist over a month. The patient information then was entered into the WebMD application and the diagnoses provided were compared with clinical diagnoses. The application provided a list of 1 to 20 possible diagnoses per patient. The symptom checker correctly diagnosed 70% of patients. However, the clinical diagnosis only matched the first diagnosis on the list for 16% of patients, which is deemed unsatisfactory.

In **2014**, a study similar was conducted by Hageman et al. [21]. This new study compared the first three diagnoses provided by WebMD for 86 patients, who had hand-related problems, with a hand surgeon's diagnoses. The results showed a somewhat poor performance of the application since only 33% of the diagnoses derived by WebMD

matched the final diagnosis of the hand surgeon. They mentioned that lack of users' health literacy was a challenge.

In **2015**, Semigran et al. [59] studied the accuracy of diagnosis and triage advice in 23 symptom checker mobile/web applications. To evaluate diagnosis and triage performance, 45 pre-defined patient scenarios (a.k.a. *clinical vignette*) were input to these applications. The correct diagnosis was listed by applications (with 95% confidence level), first in 34% of the cases, and within the first 20 diagnoses in 58% of the cases. Appropriate triage advice was given in 57% of the cases. The results imply that, in many cases, symptom checkers can give the user a sense of possible diagnoses but should not be entirely relied on, as these applications are often incorrect. It was also suggested that symptom checker applications might be valuable if the alternative is not seeking for advice using an Internet search engine.

Following their earlier paper, Semigran et al. [58] performed another study on this topic in **2016**. In their new study, they compared the diagnostic performance of physicians with symptom checker applications on the same clinical vignettes using the Human Dx online platform [67]. Human Dx is a web/mobile application on which physicians generate differential diagnoses for clinical vignettes. Physicians produced better diagnoses on these vignettes. They listed the correct first diagnosis in 72.1% of the cases, which is far better than application-generated diagnoses (33% correct first diagnosis). In the top 3 diagnoses listed, physicians provided 84.3% correct diagnoses that, again, outperforms symptom checkers.

In **2016**, Powley et al. [53] conducted a study with 34 patients (21 inflammatory arthritis patients, 4 inflammatory arthralgia patients, and 9 patients with other inflammatory problems), evaluating the accuracy of triage advice provided by the *NHS choices* self-triage application [47] and the diagnostic accuracy of the WebMD symptom checker. The patients were interviewed after receiving their diagnoses from physicians. The advice given by NHS choices to patients was inaccurate. Almost all the patients should have sought help from primary care but the application inappropriately advised only 56% of patients to seek primary care. The results of diagnoses provided by WebMD showed that, out of 21 patients with inflammatory arthritis, only four received a correct first diagnosis. It was concluded

that giving online suggestions for this particular condition is inappropriate since the recommendations are often inaccurate.

2.4.2 User Interface Design and Usability Evaluation

In **2016**, Abdullah et al. [1] studied proper navigation style (paging vs scrolling) in mobile symptom checkers. Three types of questions were designed and the experiment was conducted with 30 users. The results of this study showed that, on average, 63% of the participants preferred paging over scrolling, although the recorder task completion time indicated that participants spent more time on paging. In this study, only one UI feature of a self-triage mobile application was evaluated, which makes this study in a sense very limited, since other usability aspects of these applications were not discussed.

In **2017**, Marco-Ruiz et al. [43] designed and conducted a two-phase usability assessment method to find human-computer interaction barriers in symptom checkers. The Erdusyk symptom checker (available in Norwegian) was evaluated using this method. Various usability barriers were detected in this study that can be generalized to most symptom checkers: navigational problems could cause distractions to users while interacting with the application, long sentences and paragraphs in the application's questionnaire could be tedious, and the repetitive use of complex medical terms could slow down the interpretation of the feedback provided by the application. The authors detected many other usability barriers preventing users from using the application appropriately. They suggested that questions should be in form of short texts, and the number of medical terms should be minimized to improve the usability of such application.

In **2017**, Li et al. [38] compared the usability of three different online symptom checkers, as well as the plain Google search engine. The online symptom checkers were EverydayHealth.com, FamilyDoctor.org, and WebMD.com. Each symptom checker uses a unique user interface. Ten users participated in the study. Each participant first completed questionnaire and then each completed three symptom checking tasks. Following each task, participants were asked to answer a success task questionnaire, which asked the user to rate the difficulty or the easiness of the task if the task was completed, or a non-success questionnaire if the task was not completed. The results showed that 70% of the participants preferred the FamilyDoctor.org symptom checker. However, all participants still preferred

to search for self-diagnosis information using the Google search engine over the three symptom checkers evaluated in this study.

2.4.3 Critical Analysis

In **2014**, Lupton et al. [41] highlighted some issues that using symptom checkers applications could cause. The main concern reported by this research was hindering the doctor-patient relationship, in addition to concerns regarding the privacy of users' personal data. It was mentioned that the final diagnosis is always contingent even if it is done by a physician. However, the physician's authority tends to obscure the contingency of their decisions. The availability of symptom checkers would allow a patient to seek out a doctor not to offer a diagnosis, but rather to endorse a diagnosis that a patient brings to the consultation, which would negatively affect the doctors' authority. Lupton et al. also elaborated on how symptom checker developers failed to properly inform their users about how they are using their personal data such as email addresses, names, locations, and users' inputs on different conditions.

2.5 Summarization and Discussion

This section summarizes, in Table 1, the nine papers included in the literature review. The summary is presented along four criteria derived from the analysis of these papers (C1-C4), to illustrate their similarities and differences. The criteria are as follow:

- C1: If the paper is a review paper, then this criterion reports the number of applications studied.
- C2: Whether the paper reports diagnostic accuracy.
- C3: Whether the paper presents UI design and usability evaluation.
- C4: Whether the paper discusses barriers of symptom checkers/self-triage applications.

As shown in Table 1, five papers discuss diagnostic accuracy: two papers (P1, P2) evaluate the diagnostic accuracy of the WebMD symptom checker on a specific type of medical condition, P5 evaluates the diagnostic and triage advice accuracy of two systems on a spe-

cific type of medical condition, and P8 and P9 review the diagnosis accuracy of 23 applications and compare their results with physician-derived diagnoses. Three papers focus on usability evaluations or UI elements: P7 discusses the possible barriers that could result from the design of the different elements and features in UI as a secondary subject, P8 compares the UI attractiveness of three applications, and P9 presents possible barriers that users of symptom checkers and physicians are concerned with by evaluating 35 applications.

Table 1 - Summary of related works

Article Code	C1	C2	C3	C4	Reference
P1		Yes			[18]
P2		Yes			[21]
P3	23	Yes			[59]
P4	23	Yes			[58]
P5	2	Yes			[53]
P6			Yes		[1]
P7			Yes	Yes	[43]
P8	3		Yes		[38]
P9	35			Yes	[41]

Reviewing these papers helped us realize different strengths and weaknesses of symptom checker mobile/web applications. Most of the papers discussing diagnostic accuracy concluded that these tools are not providing reliable medical advice that could help a patient with a decision regarding their condition. Also, it was concluded that the provision of self-diagnosis might affect the doctor-patient relationship since a patient could rely on the information that is available in such applications instead of a professional advice.

Additionally, it was mentioned that as a substitute for online diagnosis using plain search engines, symptom checker applications represent a better alternative. However, one study mentioned that when users are unable to relate to the user interface of such applications, they tend to use search engines to seek medical information. This makes it crucial

for developers to create UI that improves the attractiveness of such applications. As suggested by various studies, different usability guidelines should be taken into account. For instance, using paging instead of scrolling and avoiding too much text in the questionnaire part of the application. In addition, minimizing the use of medical terms and navigational problems could improve the user experience.

In terms of data collection, the security and privacy of user data is a pressing concern for this type of applications. Developers of symptom checker applications that are collecting user information are often not properly informing their users about the storage and use of this information. Yet, the collection of personal data could help related health organizations to better understand and evaluate users' problems. For instance, if an application is specifically designed for a hospital and its patients, that hospital might benefit from their patients' inputs and provide them with more accurate care suggestions compared to other organizations that do not have access to such information. However, organizations and developers should properly inform their users about possible usage of data (as required by law) and address security concerns from the start.

2.6 Chapter Summary

In this chapter, nine peer-reviewed papers with different subjects related to symptom checkers and self-triage application were reviewed. These papers are briefly summarized and discussed. The provision of diagnoses by applications is interpreted as somewhat problematic. As suggested by reviewed papers, it is important to provide well-designed applications to attract users to work with them and to prevent users from facing usability barriers.

The next chapter presents the review of 17 *commercial* symptom checker/self-triage applications.

Chapter 3 Review of Commercial Applications

Following our review of literature on symptom checkers and self-triage applications, we found it necessary to also review existing commercial applications. Systems from mobile application markets are included in this review based on a set of inclusion criteria. Seventeen selected applications are described and summarized. A comparison is done based on a set of measurable criteria determined after several meetings with stakeholders from Montfort Hospital and after the literature review.

3.1 Collection of Applications

Two important mobile application markets (Google Play and Apple's App Store) and Google's search engine were queried using "symptom checker" and "self-triage application" as keywords. Several applications were found from the first 100 search results of Google's search engine. In addition, 64 applications were suggested by Google Play for Android devices, and 49 applications were suggested by App Store for Apple devices. Some of the applications were found through a snowballing approach exploiting references from other research or from inside research papers. The inclusion criteria that were used are that applications:

1. Provide triage advice to users or be developed for self-triage.
2. Focus on multiple problems/disorders/signs and are not discussing single types of problems like heart diseases or muscular problems [69].
3. Discuss symptoms of adult humans, not children or animals.
4. Work on mobile devices.
5. Are accessible for free.
6. Are available in English.
7. Are not asking for registration codes issued by a third party.
8. Are accessible from different geographical areas. Applications that are focusing on one particular area are not included in this study.

9. Are not providing advice based on the logical/medical content of some other application. For instance, patient.info [51] and netDoctor.com [30] are using the same medical content and logic as the Isabel symptom checker [29].

All the applications were checked individually, and 16 applications were initially selected for this study. However, while this research was being conducted, six applications went out of the market [66]. Four of these six applications were excluded from our study, but iTriage [31] and Easy symptoms [16] were still included since their iOS versions are still functioning. The assessments provided by iTriage and Easy symptoms are however not considered valid, as these applications are not supported anymore. Finally, five more applications were added to this research while this thesis was being written. Systems included in this study are either symptom checkers that provide triage advice or self-triage applications. Table 2 lists the 17 applications included in this analysis:

Table 2 - Selected commercial mobile applications

Application Code	Applications Name	Search Engine	References	Platforms
S1	ADA	Google Play & App Store	[2]	iOS and Android
S2	Buoy	Google	[11]	Web based
S3	Doctor Diagnose	Google Play	[14]	Android
S4	Drugs.com	Paper Reference	[15]	Web based, iOS and Android
S5	FamilyDoctor.org	Paper Reference	[17]	Web based
S6	Easy Symptoms	App Store	[16]	iOS
S7	Health direct	Google	[23]	Web based
S8	HealthNAV	Google Play	[27]	Web based, iOS and Android
S9	Healthwise	Paper Reference	[24]	Web based
S10	iTriage	Google Play & App Store	[31]	iOS
S11	Isabel	Google	[29]	Web based, iOS and Android

Application Code	Applications Name	Search Engine	References	Platforms
S12	Mediktor	Google Scholar	[44]	Web based, iOS and Android
S13	NHS Direct	Google	[48]	Web based
S14	Symcat	Paper Reference	[63]	Web based
S15	Symptify	Google Play & App Store	[64]	Web based, iOS and Android
S16	Symptomate	Google Play & App Store	[65]	Web based, iOS and Android
S17	Your.MD	Google Play & App Store	[71]	Web based, iOS and Android

3.2 Description of the Applications

The main features of the selected applications are described below. Four of these applications will be illustrated with screenshots (three of the most popular ones, and one that includes a chatbot). Note that with the exception of Mediktor, no evidence was found in the peer-reviewed literature or web sites about usability and functionality evaluations of these applications.

- **Ada** [2] is a symptom checker mobile application accessible on iOS and Android platforms (Figure 5). This project is funded by the European Union and was developed by a group of physicians and computer scientists. It is not providing healthcare facility suggestions. The application requires its users to create an account by submitting an email address, or via existing Facebook or Google accounts. The application provides possible diagnoses and their likelihood based on a user's inputs about his/her health condition. Another notable functionality of the application is that it allows users to save their medications and re-visit the assessments and recommendations regarding their old symptoms. Additionally, the physician can use this tool as clinical decision support system to get a list of possible existing diagnoses.

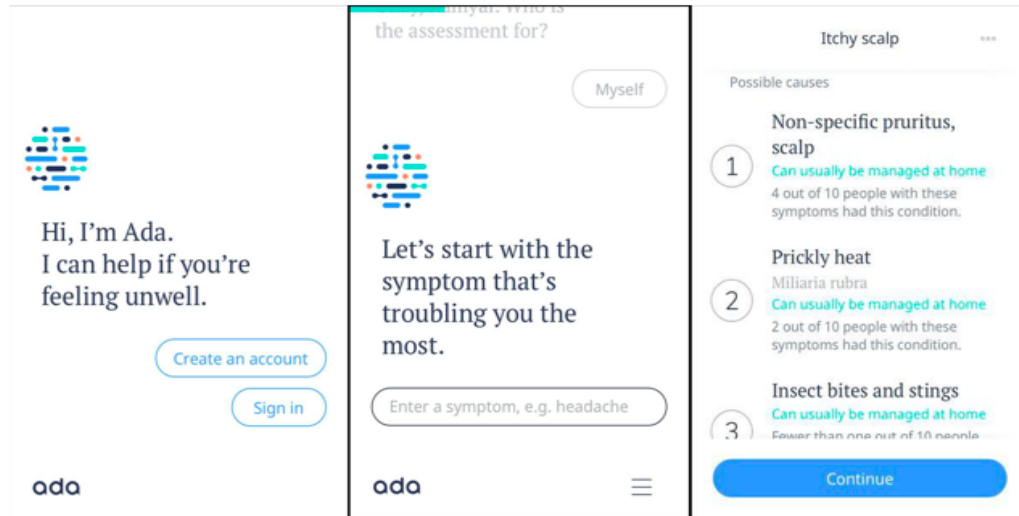


Figure 5 - Ada symptom checker

- **buoy** [11] is a symptom checker web application developed by Buoy Health, Inc. The application is accessible on multiple platforms via Internet browsers. This application is not customized to any location and is available in English only. Users can benefit from this symptom checker without disclosing any personal data. However, by creating an account, users become able to re-visit the symptoms that they checked before. The application provides possible diagnoses and their likelihood.
- **Doctor Diagnose** [14] is a symptom checker mobile application designed for Android platforms. The source of medical information used was not specified. The application is not customized to any location and is not supporting multiple languages. It is not required to provide personal data to use the application. The application provides diagnoses to users. In addition to symptom checker functionalities, the application is providing access to a medical database that has information about the relationships between medical conditions, symptoms, and medications.
- **Drugs.com** [15] is a website that includes a symptom checker application, besides many other features such as an introduction to different drugs, alerts from the US Federal Drug Administration, and side effects of drugs. This symptom checker can be accessed via Internet browsers and used on different screen sizes. The medical content of this application is based on Harvard Health Publications. The developers have made their content compliant with HONcode (a standard for trustworthy health information) [6]. This application is not customized to any location and is

- not supporting multiple languages. The application does not ask for personal data, but it provides users with a possible diagnosis for their medical condition.
- **familydoctor.org** [17] is a web-based symptom checker accessible from various platforms via Internet browsers. The application is supported by American Academy of Family Physicians [5]. This application is not customized to any location, but it supports Spanish and English. The application does not ask for personal data, but it provides users with a possible diagnosis for their medical condition.
 - **Easy Symptoms** [16] is a symptom checker mobile application designed for iOS platforms. The source of medical information was not provided by the application. The application is not customized to any location and is not supporting multiple languages. It also, does not ask for personal data. The application provides users with possible diagnosis for their health condition.
 - **Health Direct** [23] is a multi-platform symptom checker application that has Android and iOS versions, in addition to a web-based version. The mobile applications of this symptom checker can be accessed only by Australian residents. However, its website is accessible globally. This application is a government-funded project and its content is compliant with HONcode standards. This application also can help a user, who is residing in Australia, to find a suitable healthcare facility upon request (customized to a particular location). This application is only available in English. The application does not ask for personal data, but it provides users with a possible diagnosis for their medical condition.
 - **HealthNAV** [27] is a multi-platform self-triage application that has Android and iOS clients, in addition to a web-based version. The application is supported by BayCare Health System, Inc. It is customized to a particular location (Florida) and is only available in English. Users are not required to provide any personal data, but they can create an account to share their health information with the organization.
 - **Healthwise** [24] is a web-based self-triage application accessible via internet browsers on various platforms. The application is developed by Healthwise Inc., a not-for-profit health company. The application is not customized to any location and is not supporting multiple languages. It does not collect personal information

and does not provide users with diagnoses. The logic and medical content of this application are being used by many other healthcare organizations [25][26]. This application provides a user with a treatment advice if the problem possibly can be resolved without medical assistance.

- **Isabel** [29] is one of the most well-known symptom checker applications and is available on multiple platforms. This application is supported by Isabel Healthcare, Inc. In this application, users are able to enter their symptom(s) and, without asking any question, the application lists possible diagnoses (Figure 6). This application links its users to other informative medical websites upon request. The application is not customized to any location and is not supporting multiple languages.

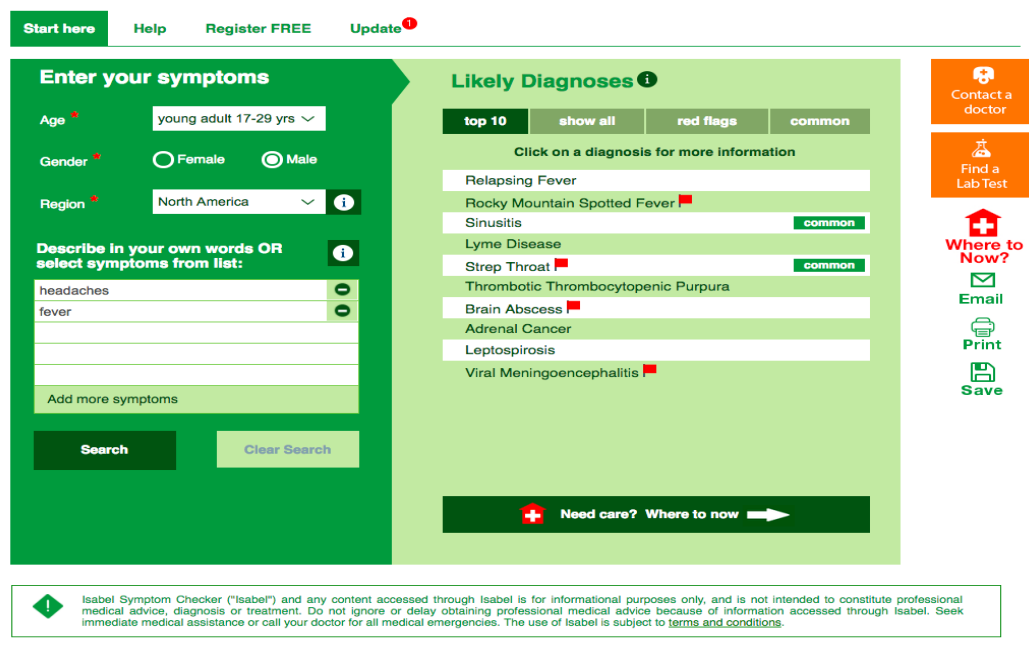


Figure 6 - Isabel symptom checker

- **iTriage** [31] was another well-known symptom checker but it is no longer available for download on App Store or Google Play. The application was downloaded more than 15 million times on mobile devices since its deployment. The application provides information about doctors and facilities that users can visit regarding their condition (customized to locations). It also provides a tracking feature that allows users to reassess their old symptoms when they share their personal information through an account. Additionally, users can find information about different medications and procedures.

- **Mediktor** [44] is one of the most recent products on the market. It is a multi-platform symptom checker that has Android and iOS clients, as well as a web-based version. This application is developed by The San Carlos Sanitary Research Institute and the University of Barcelona. It is available in four languages: English, Portuguese, Italian and Spanish. The developers of this application claim that diagnoses suggested by this application coincide with physician diagnoses in 91.3% of the cases. Mediktor (Figure 7) also makes contacting specialists possible by providing their availability, contact information, and price for providing suggestions via chat conversations or telephone calls.

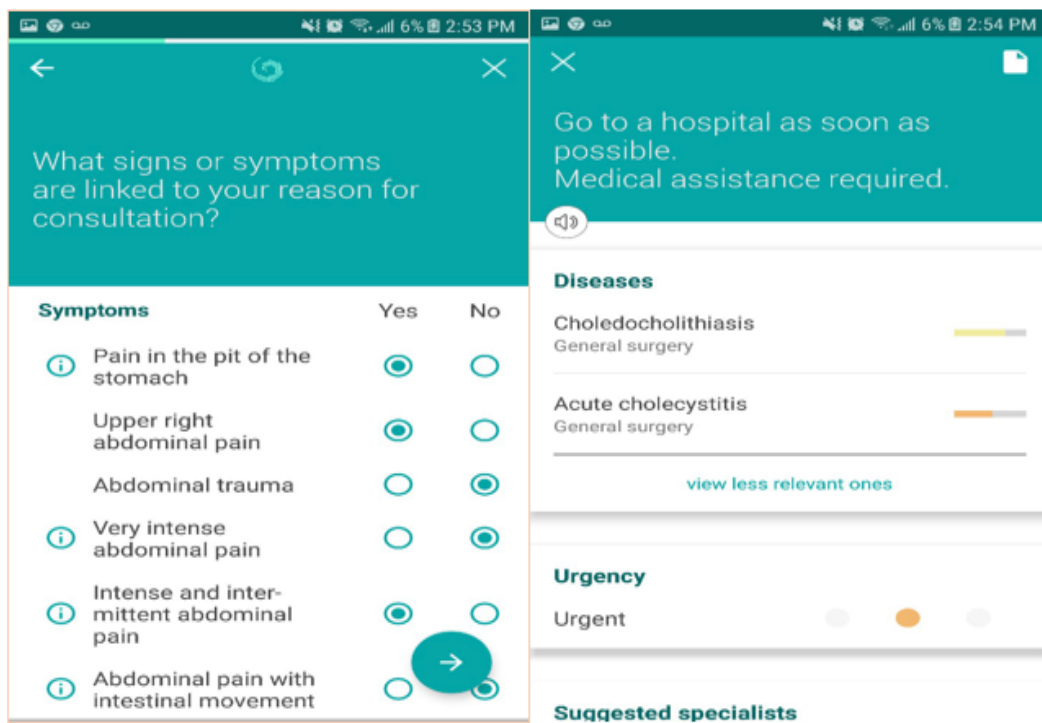


Figure 7 - Mediktor symptom checker

- **SymCat** (Symptom-based, Computer Assisted Triage) [63] is a web-based symptom checker accessible from various platforms via Internet browsers. The medical information and clinical content behind this application were designed by physicians involved in a research group called AHEAD. This application provides healthcare facility suggestions with the help of Google Maps, which makes it customized to user locations. This application is not supporting multiple languages. Users are not required to provide any personal data. However, they can create an

account to save their health information and re-visit symptoms previously checked. SymCat provides diagnoses to users as well as their likelihood.

- **Symptify** [64] is a multi-platform symptom checker that has Android and iOS clients, as well as a web-based version. It is developed by a group of emergency physicians and software engineers. It provides healthcare facility suggestions to US residents and it supports multiple languages (Spanish and English). It is not required to register to this application, but many additional features of this application get activated upon registration. The application provides possible diagnoses and their likelihood.
- **Symptomate** [65] is a multi-platform symptom checker that has Android and iOS clients, as well as a web-based version. The medical content of this application and algorithms are designed and provided by a software company called Infermedica [72]. The application is not customized to any location. It supports multiple languages (English, Spanish, French, German, Russian, and Polish). It is not required for users to register to the application. The application provides possible diagnoses and their likelihood.
- **Your.MD** [71] is a multi-platform symptom checker that provides self-diagnosis and self-triage advice (Figure 8). In addition to their symptom checker features, this application allows its users to ask questions regarding different health concerns in the form of a chat conversation. A unique feature of this application is that it provides integrated chatbots with different instant messaging technologies, such as WhatsApp, Telegram, and Facebook Messenger. This feature allows users to simply use the system to assess their symptoms and discover health information.

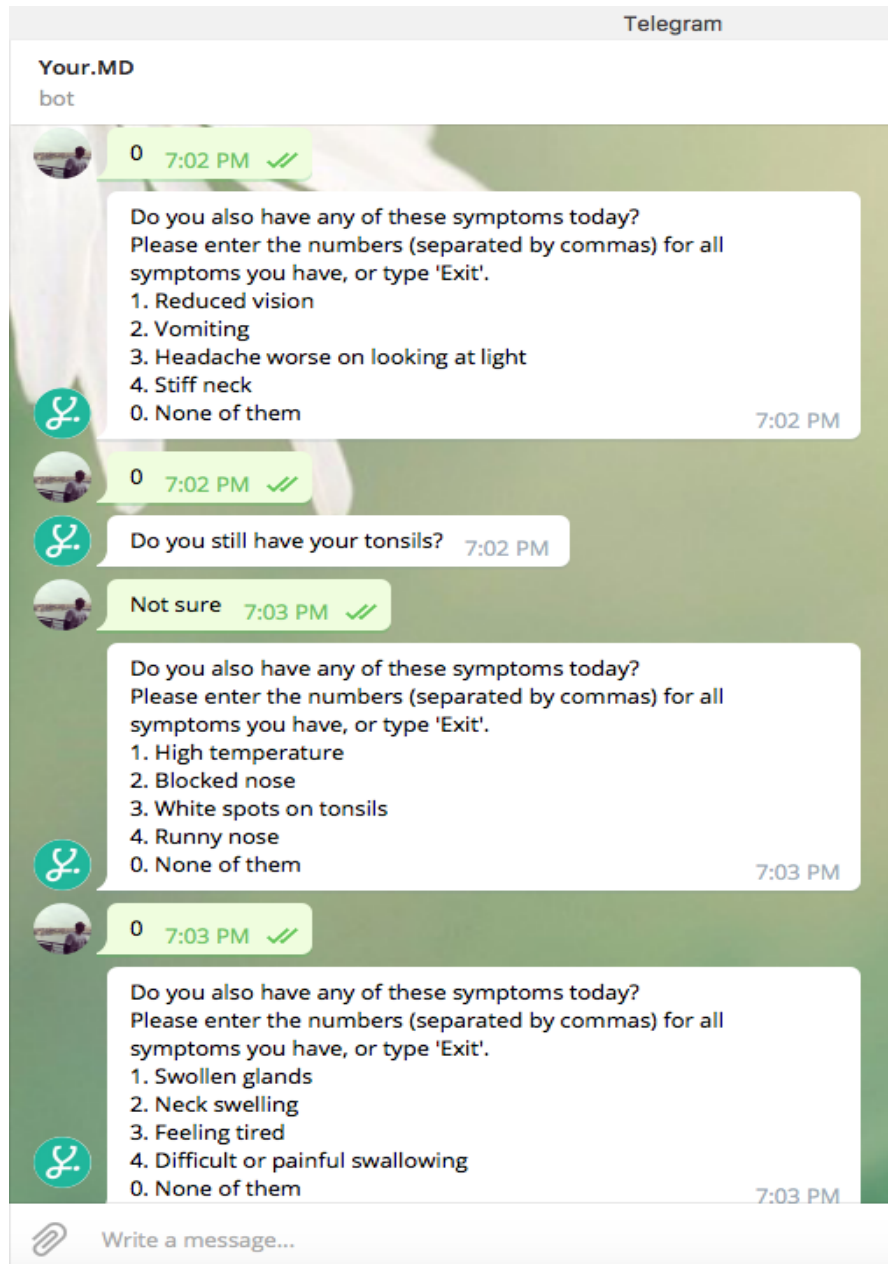


Figure 8 - Your.MD symptom checker ChatBot

3.3 Comparison

The categorization used here is based on ten measurements criteria (see below). These criteria were selected after various meeting with stakeholders from Montfort Hospital and after reviewing the literature [41]. The results are summarized in Table 3.

- 1) Offline Accessibility: if the application can generate results without a connection to the Internet, then the value is set to Yes.

- 2) Multi-Platform Access: if the application can be accessed through various mobile platforms such as Android or iOS, then the value is set to Yes.
- 3) Medical Information Supervision: if the medical contents used in the application are derived from academic literature or if a medical team (from a health organization or a hospital) validated the content, then the value is set to Yes. If the source of medical contents is unknown, then the value is set to N/A.
- 4) Customized to Location: if the application provides the location of nearby and accessible healthcare facilities, then the value is set to Yes.
- 5) Multi-Language Support: if the application is supporting multiple languages, then the value is set to Yes.
- 6) Validation: if the application's usability or/and its diagnostic accuracy were validated with results reported in peer-reviewed literature, then the main reference is provided.
- 7) Data Collection: if the application requires users to provide personal data, then the value is set to Yes. Else, if it is optional to create account but some of the features are provided only after creating an account, then the value is set to +/-, otherwise the value is set to No.
- 8) Availability of Self Diagnoses: if the application is providing diagnoses to users, then the value is set to Yes.
- 9) Type: whether the application is static or dynamic (see section 1.3) is specified.
- 10) Number of symptoms that are being assessed by the application. For some applications, this number is unavailable and so the value is set to N/A.

The colors in Table 3 are interpreted as follows: green indicates satisfactory, yellow indicates partially satisfactory, and red indicates unsatisfactory. The availability of self-diagnosis and the collection of data by an application cannot be assessed as being satisfactory or unsatisfactory (so they are not color-coded). Both these features can negatively affect users, yet if implemented correctly, users and healthcare services can benefit from them.

Applications S3, S6 and S10 provide offline accessibility to their users. Developers should consider that in moments when users do not have any access to the Internet, what might be a case in some urgent situations. Only three applications (S3, S6 and S10) are not

being supported by multiple platforms. iTriage (S10) used to allow access to multiple platforms, covering both offline accessibility and multiple platform access, however, right now just the iOS version of this application is functional. No information was found to verify the validity of the medical content of two applications (S3 and S6). The medical content of iTriage (S10) is not considered verified since the application is no longer being supported by its developers. In terms of being customized to a location, five applications are able to provide healthcare facility suggestions to users (S7, S8, S10, S11, S14). Four applications are providing multilingual support: familydoctor.org (S5) and Symptify (S15) support Spanish and English whereas Mediktor (S12) and Symptomate (S16) support more than two languages. Only one application supports the French language (S16).

Among all the applications included in this review, only one provided peer-reviewed validation (S12), which is surprising given the sensitiveness of the topic. For some of these applications, online reviews could have been retrieved (e.g., from an application store, from Twitter, etc.). Evaluating these reviews could be helpful in order to determine new requirements for such applications [42]. However, these reviews would not satisfy the need for evaluating the usability of an application.

One application required registration in order to provide the assessment to its users (S1). Ten other applications provide this feature, but users are still able to benefit from the applications' medical suggestions without sharing any personal data. Three applications are pure self-triage applications and do not provide any kind of diagnosis to users (S8, S9, S13). The drawbacks of self-diagnosing using such applications were already discussed in the previous chapter. However, receiving a correct diagnosis could have a huge impact on the user experience and also improve the quality of health services provided to users. Nearly 50% of the selected applications are dynamic, which means these applications can learn from users' inputs and their data entries. Static applications are simply focused on a set of predefined symptoms and are unable to expand their decision-knowledge resources.

In a nutshell, most applications are attempting to provide verified medical diagnoses and triage advice to a vast group of users, by supporting multi-platform access. However, applications are often not providing healthcare facility suggestions and are not accessible offline. In addition, the usability and functionality of most of the applications are either not evaluated or their evaluation is not reported in peer-reviewed publications.

Table 3 - Comparison of existing applications

Apps	Offline Accessibility	Multi- Platform Access	Medical Information Supervision	Customized to Location	Multi-Language Support	Usability and Functionality Validation	Data Collection	Availability of Self-Diagnoses	Type	Number of Symptoms
S1	No	Yes	Yes	No	No	No evidence	Yes	Yes	Dynamic	236
S2	No	Yes	Yes	No	No	No evidence	+/-	Yes	Dynamic	1000+
S3	Yes	No	N/A	No	No	No evidence	No	Yes	Static	57
S4	No	Yes	Yes	No	No	No evidence	No	Yes	Dynamic	140
S5	No	Yes	Yes	No	Yes	No evidence	No	Yes	Static	48
S6	Yes	No	N/A	No	No	No evidence	No	Yes	Static	46
S7	No	Yes	Yes	Yes	No	No evidence	No	Yes	Static	223
S8	No	Yes	Yes	Yes	No	No evidence	+/-	No	Static	150
S9	No	Yes	Yes	No	No	No evidence	No	No	Static	150
S10	Yes	No	N/A	Yes	No	No evidence	+/-	Yes	Static	N/A
S11	No	Yes	Yes	Yes	No	No evidence	+/-	Yes	Dynamic	N/A
S12	No	Yes	Yes	No	Yes	[44]	+/-	Yes	Dynamic	N/A
S13	No	Yes	Yes	No	No	No evidence	+/-	No	Static	28
S14	No	Yes	Yes	Yes	No	No evidence	+/-	Yes	Dynamic	474
S15	No	Yes	Yes	Yes	Yes	No evidence	+/-	Yes	Static	220
S16	No	Yes	Yes	No	Yes	No evidence	+/-	Yes	Dynamic	N/A
S17	No	Yes	Yes	No	No	No evidence	+/-	Yes	Dynamic	N/A

3.4 Chapter Summary

As a complement to the review of scientific literature from the previous chapter, this chapter evaluated 17 commercial applications available on the market. Applications were briefly discussed and compared against 10 criteria. Applications are mainly providing (English-speaking) users with triage and diagnosis advice. Only a few applications are supporting offline accessibility and/or are providing advice customized to locations. The scientific evaluation (usability and functionality) of most of these applications is so far questionable, at least in terms of peer-reviewed publications.

The next chapter presents the goals, architecture, design, and implementation of our proposed self-triage application.

Chapter 4 Application Design and Implementation

This chapter discusses the goals, requirements, architecture, and implementation of the Symptoms Pal mobile application. Montfort Hospital researchers have created an initial prototype of Symptoms Pal for the Android and iOS platforms. However, since they wanted to do further research on the application (which led to this thesis) before releasing it to the market, they decided to focus on one platform and hence stopped temporarily the development of the iOS version. The Android application source code (without any goal, requirement, architecture, or design information) was provided to us for further development and major improvements, and for the usability study. Different aspects of the improved application and its users interface are elaborated here.

4.1 Goals and Requirements

People's inability to access reliable information when facing a symptom might cause them anxiety and drive them to visit an ED, or might make them neglect their situation and avoid needed care. In both contexts, it is important to provide users with a reliable source of information that will reduce uncertainty about how to react to symptoms. The primary goals of the application, determined in collaboration with researchers from Montfort Hospital, are shown in Table 4. These goals go beyond the goals of the thesis itself.

Table 4 - Primary goals of the application

Goal ID	Goal Description
G1	Improve the overall user experience.
G2	Reduce the number of inappropriate visits to emergency department.
G3	Expedite warranted visits to emergency departments.

It is concluded in section 2.5 that the diagnosis and triage advice provided by other applications are not completely providing reliable answers to its users. Additionally, diagnosing by use of the search engines may not direct users to the most appropriate healthcare facility [40]. Hence, G1 is to improve the user experience in comparison with other applications

that are providing triage advice and also comparing to user's experience while searching for their chief health complaints in search engines such as Google and Yahoo!

Additionally, as discussed in Chapter 1, inappropriate visits are one of the factors that influence overcrowding in an ED. Hence, this new mobile application aims to reduce the number of inappropriate visits to healthcare services and to satisfy G2. G3 is about expediting warranted visits to emergency departments. For instance, there are cases that users have doubts about existence of a health problem, however, since they are not certain they tend to ignore their problem. Provision of an application that can reduce uncertainty by providing reliable information can eventually expedite warranted visits to EDs (i.e., minimize the time wasted at home when someone is uncertain about the seriousness of his/her medical condition) and also satisfy G3.

In this research, no baseline is determined to evaluate if these goals are entirely satisfied. As was described earlier in section 1.5, the goal of this thesis is to evaluate the usability and user interface of such an application.

To design and develop this application with regards to these goals, 13 requirements were identified from several meetings with stakeholders from Montfort Hospital and from related work in peer-reviewed publications (Chapter 2) and in existing symptom checker applications (Chapter 3). Table 5 shows these self-triage application requirements and their corresponding goals. Requirements highlighted in green are the ones that are measured through this research and usability study. The requirements in red are necessary to develop a self-triage application, but it was not possible to evaluate them through this research due to their clinical nature. Requirements in yellow are features that are partially measured within this study from a usability perspective; the correctness of the medical suggestions provided cannot be evaluated in this study.

Table 5 - Requirements of the self-triage application

Id.	Requirements	Goals
R1	Support using existing application information without forcing users to sign-up or provide any data.	G1
R2	Provide offline accessibility.	G1
R3	Provide reliable information about different medical conditions.	G1,G2, G3
R4	Provide user interfaces enabling users to select easily their symptoms and navigate through the application.	G1
R5	Recommend the best triage option (e.g., go to ED, visit a family doctor, go to a walk-in clinic, or go to a pharmacy).	G2,G3
R6	Provide the right level of acuity of a user's problem. (e.g., emergency, non-emergency, self-care).	G2,G3
R7	Suggest self-treatment options only for self-care cases.	G1,G2
R8	Provide additional links to reliable online informative healthcare resources to help the user with the current problem.	G1
R9	Embed the source of information within the application.	G1
R10	Direct the user to the best healthcare setting based on the level of acuity and suggested care option.	G2,G3
R11	Support a multi-lingual interface.	G1
R12	Support multi-platform access (minimally for smartphones and tablets).	G1
R13	Provide additional medical information to educate users.	G1

4.2 Architecture

A 3-tier architecture, shown in Figure 9, was developed to integrate the requirements and guide the implementation. This architecture shows the overall structure of the components, relationships between these components, and technologies used in this application. Each of these three layers will be covered in the next sections. In order to complete the development of this application, the Android Studio v2.3.3 integrated development environment

(IDE)³, based on the Java programming language, was used for programming and debugging the application.

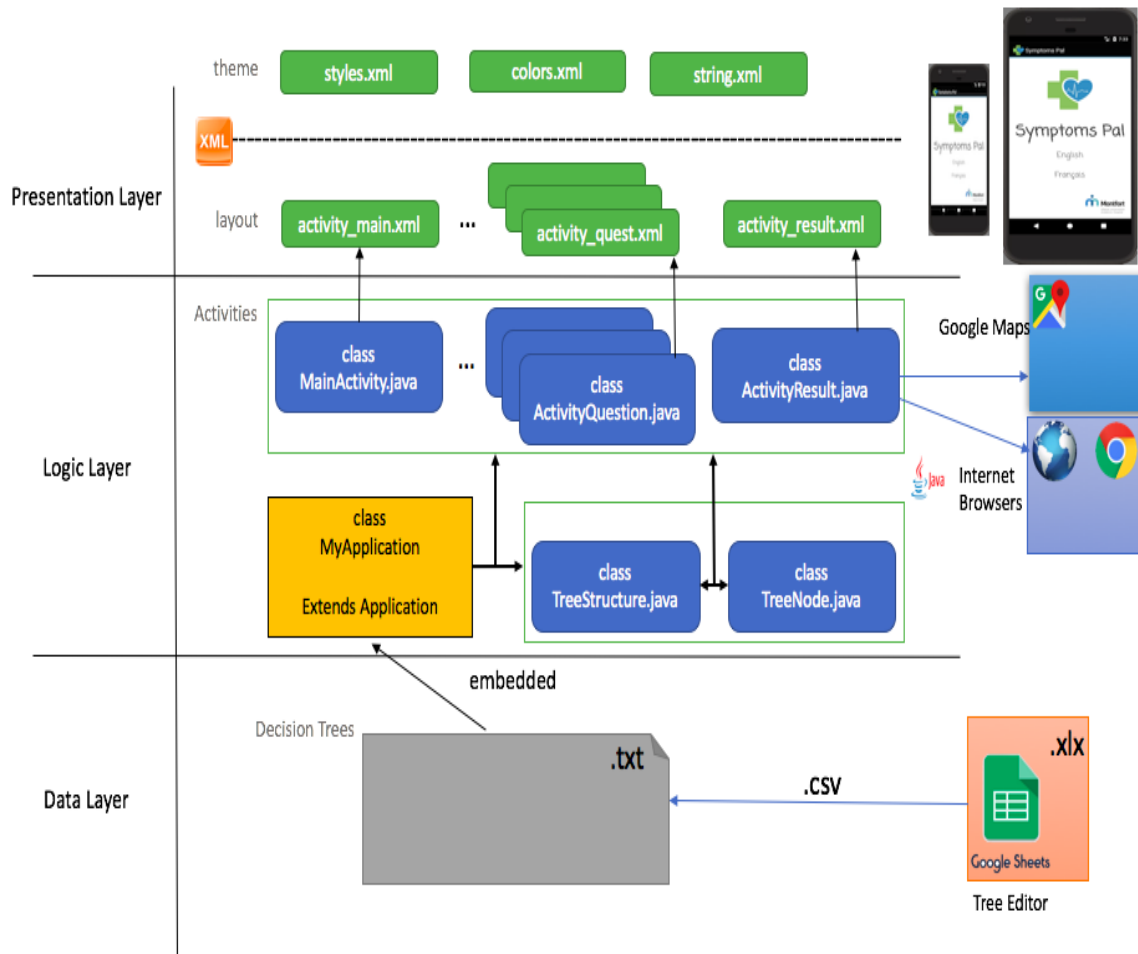


Figure 9 - Symptoms Pal architecture

4.3 Data Layer

The data layer stores the clinical content as symptom-oriented decision algorithms that will be used by the application. A decision algorithm specifies a tree of questions for a given symptom, potential answers to choose from, and ultimately the corresponding triage advice with additional information to be provided to users.

The original application initially received encoded these algorithms as decision trees in plain text. We have extended the initial formal to include additional information

³ <https://developer.android.com/studio/>

regarding human body regions and (more consistent) advice categories. We also decided to store these trees in Google Sheets in order to have one source of truth and enable multiple collaborators to edit trees online (in an Excel-like environment), support comments, keep track of changes, and enable better consistency across algorithms. These trees are converted to a comma-separated value (CSV) format and stored as a text file in our data layer. Figure 10 shows the pattern that is used to input information in each decision tree file. Each line of this file stores one of the following types of information:

- **Category Lines** are represented by C and include symptoms and signs (e.g., headache, fever, chest pain, etc.) and their related human body regions (e.g., head, abdomen, chest, etc.).
- **Question Lines** are represented by Q and include questions about a symptom and the list of possible answers to each question.
- **Result Lines** are represented by R and include triage advice, the description of the expected problem and reasons supporting the advice, further medical suggestions, and links to additional information.

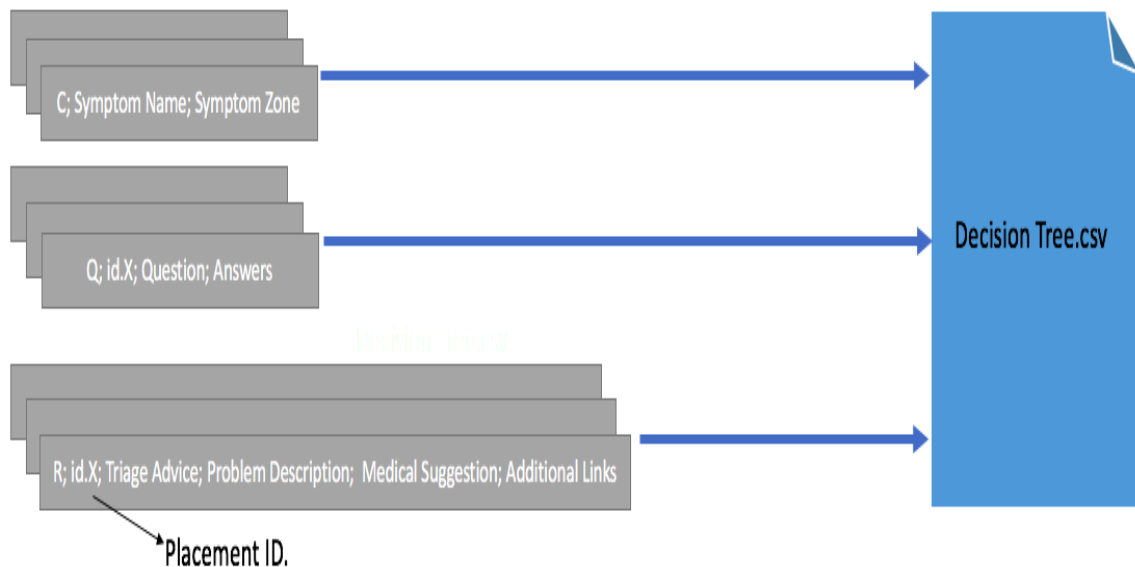


Figure 10 - Overview of data included in decision trees (algorithms)

Result lines and question lines both include a placement identifier (e.g., 8.1.1.0). This identifier is used to uniquely label each category and to determine the placement of questions and results in the tree (the longer the identifier, the deeper it is in the tree). Questions and

answers of each category are listed following the category. Figure 11 shows the decision tree for the fever symptom (as seen in Google Sheet) as an illustrative example.

C	Fever		
Q	8.	Is fever accompanied by any of the following:\n- Stiff neck.\n- Seizure.\n- Convulsion.\n	Yes
R	8.0	Emergency Room	▼ Your signs a
Q	8.1	Does any of the following apply?\n- Fever is accompanied by pain and warmth in a certain body areas that are swollen.\n- Fi	Yes
R	8.1.0	Visit a Doctor	▼ You may ha
Q	8.1.1	Do you have any of the following?\n- Urinary symptoms.\n- Ear ache.\n- Soreness in other area of the body.\n- Breathing di	Yes
R	8.1.1.0	Visit a Doctor	▼ Fever is a si
Q	8.1.1.1	Is fever accompanied by vomiting or diarrhea?\n	Yes
R	8.1.1.1.0	Visit a Doctor	▼ Seek medic
Q	8.1.1.1.1	Is the patient younger than 2 months?\n	Yes
R	8.1.1.1.1.0	Emergency Room	▼ Over the cou
Q	8.1.1.1.1.1	Is temperature over 40°C (104°F) or patient is confused or hearing noises?	Yes
R	8.1.1.1.1.1.0	Emergency Room	▼ The temper
Q	8.1.1.1.1.1.1	Has the fever lasted more than 2 days?	Yes
R	8.1.1.1.1.1.1.0	Visit a Doctor	▼ For fever las

Figure 11 - Illustrative example – decision tree for fever

4.4 Logic Layer

The logic layer contains various Java classes that determine the functionality of components and that are essential in the creation of blocks in any Android application. This layer handles the information exchanges between data and presentation layers. The central components of the logic layer here are:

- The application Java class (MyApplication in Figure 9) reads the data from the decision algorithm file and passes it to tree structure classes.
- Tree structure classes store the algorithm information and allow activity classes to use them properly.
- Activity classes define the logic of each user interface of the application.

4.4.1 Application Class

When an Android application starts, the system creates a thread of execution as Main thread or UI thread. This thread is in charge of dispatching events to the proper UI widgets. The *Application class* in Android works as the base class in the application and embraces all other components such as activities. The subclass of this class (here MyApplication.java) executes before any other class when the process of the application is created. Here, since

the algorithms/trees should first be streamed into the application, it was necessary to make a subclass of the Application class and read the text file as an input and in the background, so the UI thread would not be busy executing more important tasks, such as handling user interactions and launching activities. This also improves the efficiency of the application.

A Java AsyncTask class was used in MyApplication.java to allow using an asynchronous task in a background thread and pass the result to the UI thread. Here, MyApplication.java streams the tree file, buffers it, and parses each line of the decision algorithm to determine whether it is a category line, a question line, or a result line. Then, it stores the data that each line of the tree contains. The information regarding the question lines and result lines is then passed to our tree data structure for further use.

4.4.2 Data Structures

A customized tree data structure was designed in order to map the content of the decision tree into the application source code. Each node of this tree is a list and has a placement identifier (id). There are two types of nodes in our trees: *question nodes*, which store parent node placement ids and the content of the question line from the CSV file, and *result nodes* that include parent placement ids and the result information. The first question in each category is assigned with a placement identifier and becomes the root of the category tree. Answers would either add a question node as a child to its parent or a leaf corresponding to a result node. The placement id of each child node is determined by the parent node and answer that leads to that node. This pattern continues recursively until all the children nodes are tree leaves. Figure 12 illustrates the tree structure just explained.

Three Java classes were created in order to implement this tree structure:

- **TreeStructureCustom.java**, which creates the tree structure for each category.
- **TreeNodeCustom.java**, which creates question and result nodes of the tree and stores the related information into those nodes.
- **ResultInfo.java**, which includes variables that are used to store results.

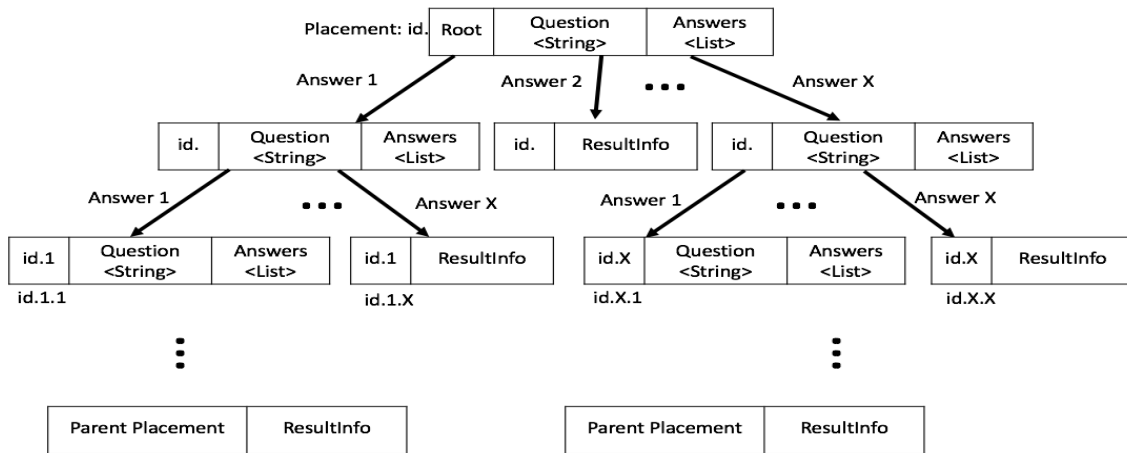


Figure 12 - Tree structure

Figure 13 shows the data structures and their relationships as a UML class diagram.

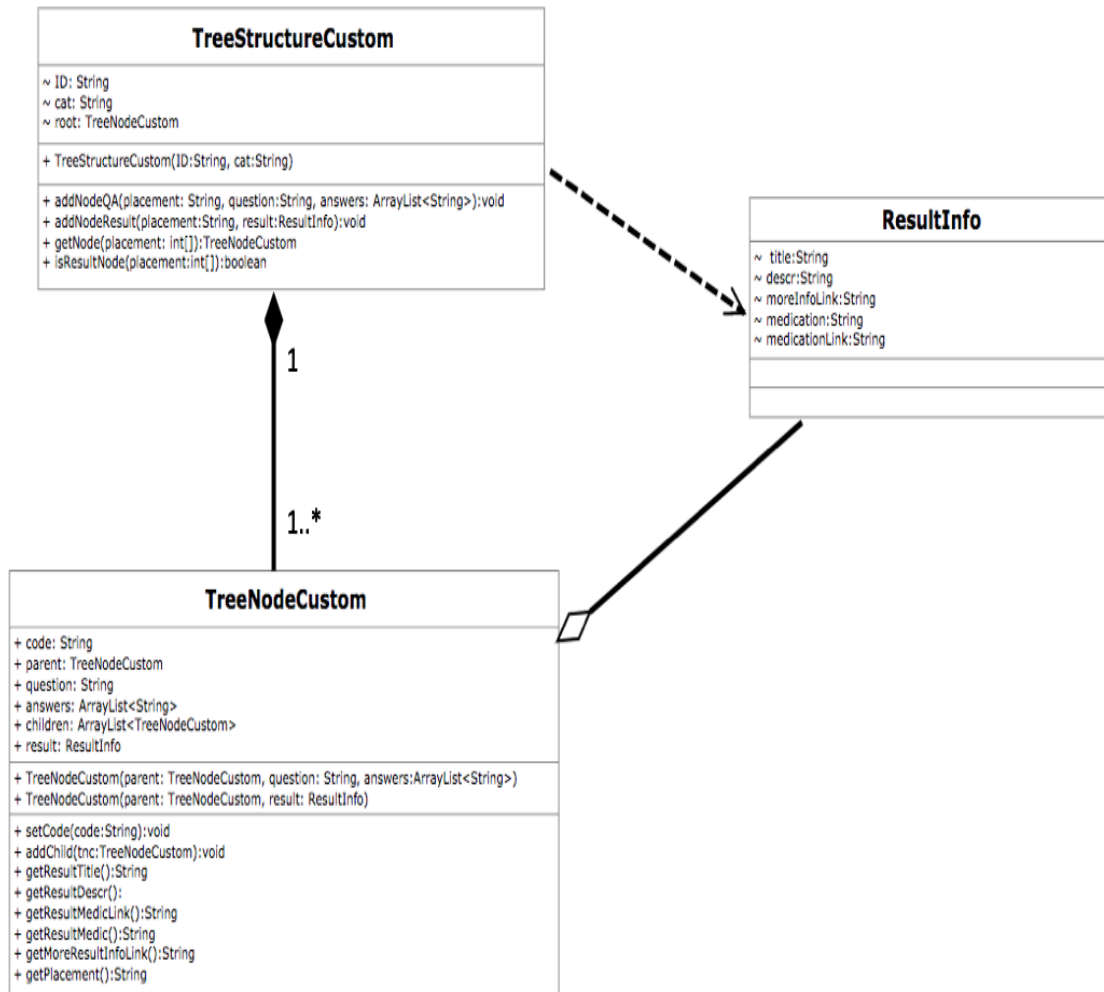


Figure 13 - UML class diagram of the data structures

4.4.3 Activities

Activity classes are among the most important components in the development of an Android application. An *activity* is a single screen of an application, similar a browser window with a UI [10]. It serves as the entry point for a user's interactions with the application and manages visual elements that present data such as an image or text. An activity also allows users to interact with the application by defining the functionality of widgets. The Symptoms Pal application has multiple activities that are independent of each other but work together to create a cohesive UI. The transition between different activities is instantiated by *intents*, which are asynchronous message objects that transfer information between activities and applications. The activities listed in Table 6 are needed to perform the following necessary interactions in this application:

- Keeping track of what the user is currently asking in the presentation layer (what is on the screen) to make sure that the application keeps providing that feature, which is hosted by the activity.
- Having data structures, the Application class and the presentation layer exchange information.
- Redirecting the user to other third-party applications (e.g., Google Maps) upon request.

Table 6 - Application activities

Activity Class	Description
ActivityLanguage.java	Allows the selection of a language (e.g., English or French) in the application. It provides the intent to ActivityBody.java.
ActivityBody.java	Stores a clickable human body pictogram. It retrieves symptoms zones from the Application class and matches each zone with a human body region. Upon selection of the affected region, it provides the intent to MainActivity.java.
MainActivity.java	Retrieves the message from ActivityBody.java and fetches the list of symptom names related to the selected body region. Upon selection of a symptom, it sends the symptom and its id to ActivityQuestion.java.

Activity Class	Description
ActivityQuestion.java	Receives the symptom name and its assigned id and instantiates the TreeStructureCustom.java class to retrieve the decision tree of that category. Keeps track of answers to either retrieve another question or send an intent to ActivityResult.java.
ActivityResult.java	Retrieves the result information and passes it to the presentation layer. Provides intents to Google Maps and Browsers upon selection.

4.5 Presentation Layer

The topmost tier of our application (Figure 9) is the presentation layer, which consists mainly of the graphical user interface (GUI). The GUI here is composed of two parts: layouts and themes. A *layout* makes the structure for the UI. Almost all the activities are incorporated with a layout (unless defined otherwise). *Themes* are common styles that can be applied to UI elements, widgets, and layouts of the application. Themes help improve the consistency of the application's appearance. Android provides a straightforward XML (Extensible Markup Language) vocabulary that corresponds to UI elements and widgets and was used to construct layouts and themes. Different layouts were provided for different screen sizes, hence accommodating different Android devices such as phones and tablets.

4.6 Interfaces of the Application

In this section, the Symptoms Pal application's interfaces and their features are introduced. Users start the application with the language page, where they can select their desired language. Following the selection of a language, users select their preferred type of symptom list (graphical or list) and then select their main symptom. They are then given a list of questions about their symptom and the triage advice will be provided based on their answers. Figure 14 shows the sequencing of the application's interfaces.

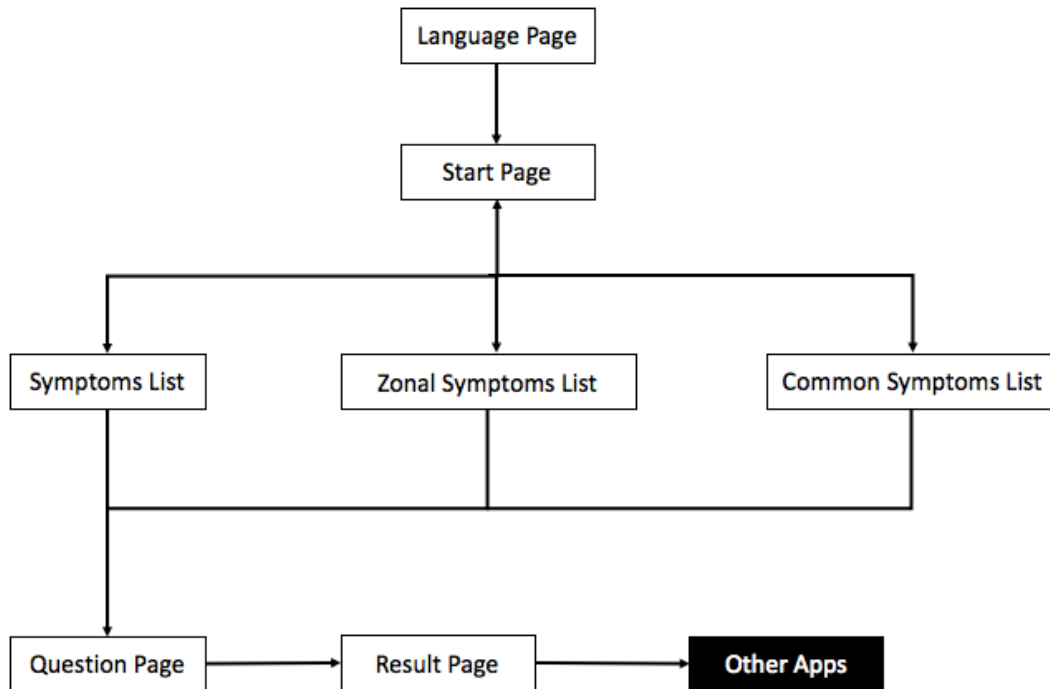


Figure 14 - Interfaces of Symptom Pal

4.6.1 Language Page

Users are able to select their desired application language, either English or French, as shown in Figure 15a. Upon selection of the language, the application’s interface is updated with that content. Users must read the legal disclaimer and terms & conditions of the application, then accept them in order to proceed (Figure 15b).

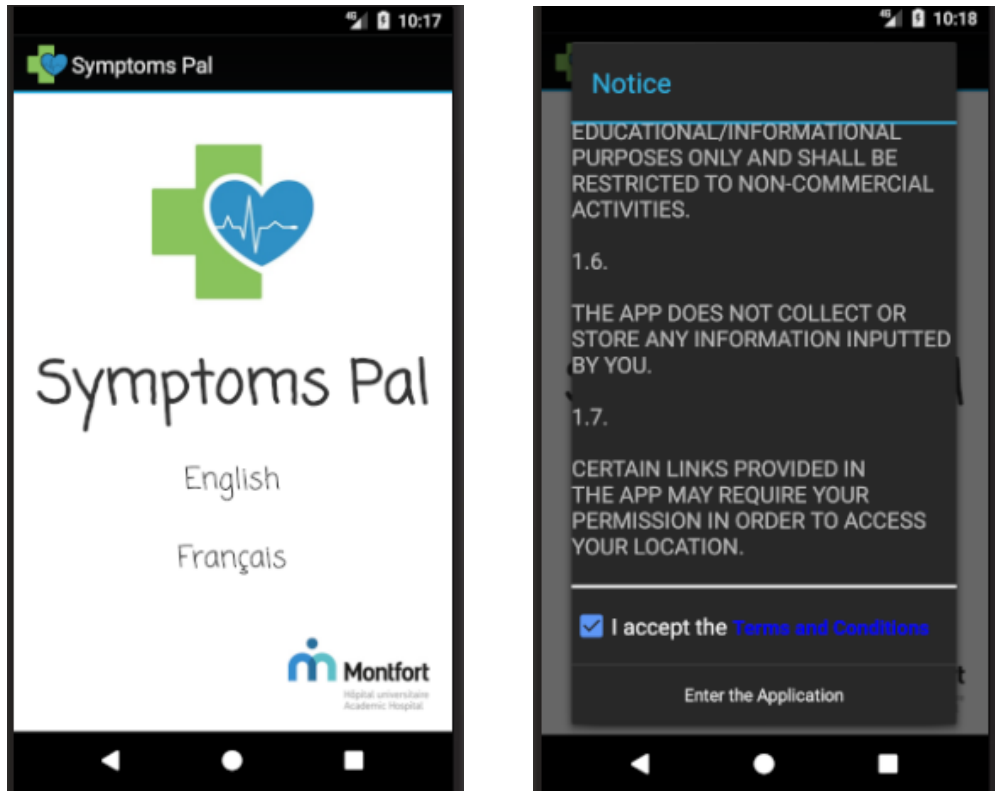


Figure 15 - a) Language page, b) Legal notice and disclaimer

4.6.2 Start Page and Symptom List

The start page presents to the users a human body pictogram. They can select their gender (Figure 16a) and click on the body part that is best associated with their main experienced symptom (Figure 16b and Figure 17a). The regions of the back of the body can be accessed by *flipping* the pictogram. A list of symptoms that are common can be accessed by clicking on the “Common Symptoms” button (Figure 17b). Also, by clicking on the “Symptoms List” button, users can access all of the symptoms covered in Symptoms Pal and query the list to find their desired symptom (Figure 18). Upon selecting their main symptom, users can proceed to the question page to answer some related questions.

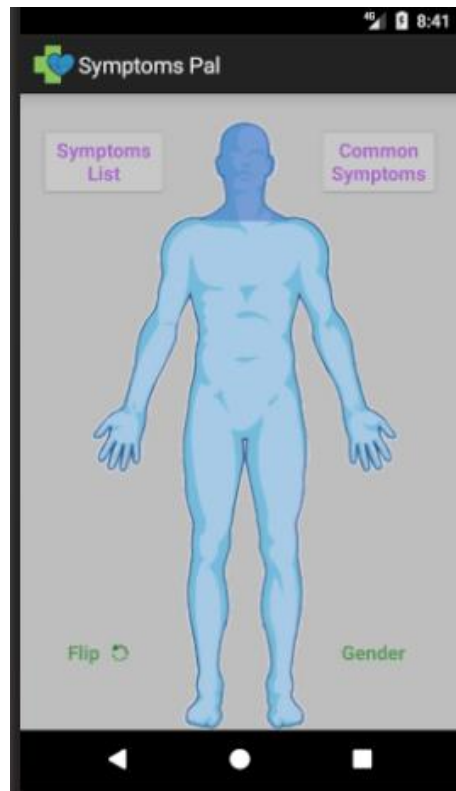
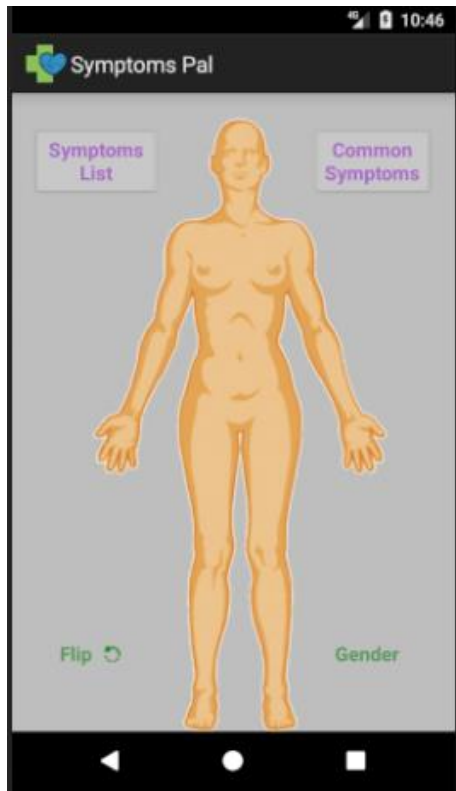


Figure 16 - a) Woman pictogram, b) Man pictogram with head zone selected

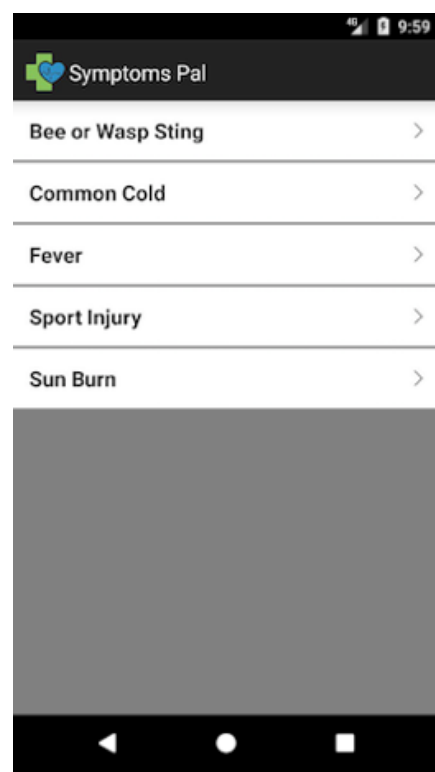
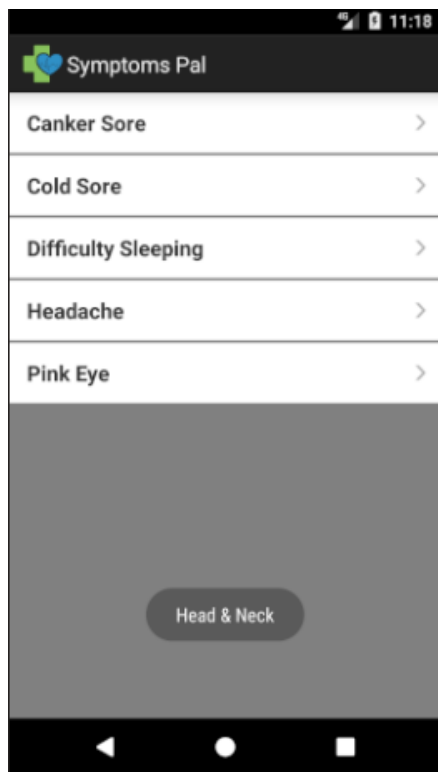


Figure 17 - a) Zonal symptoms list for head & neck, b) Common symptom list

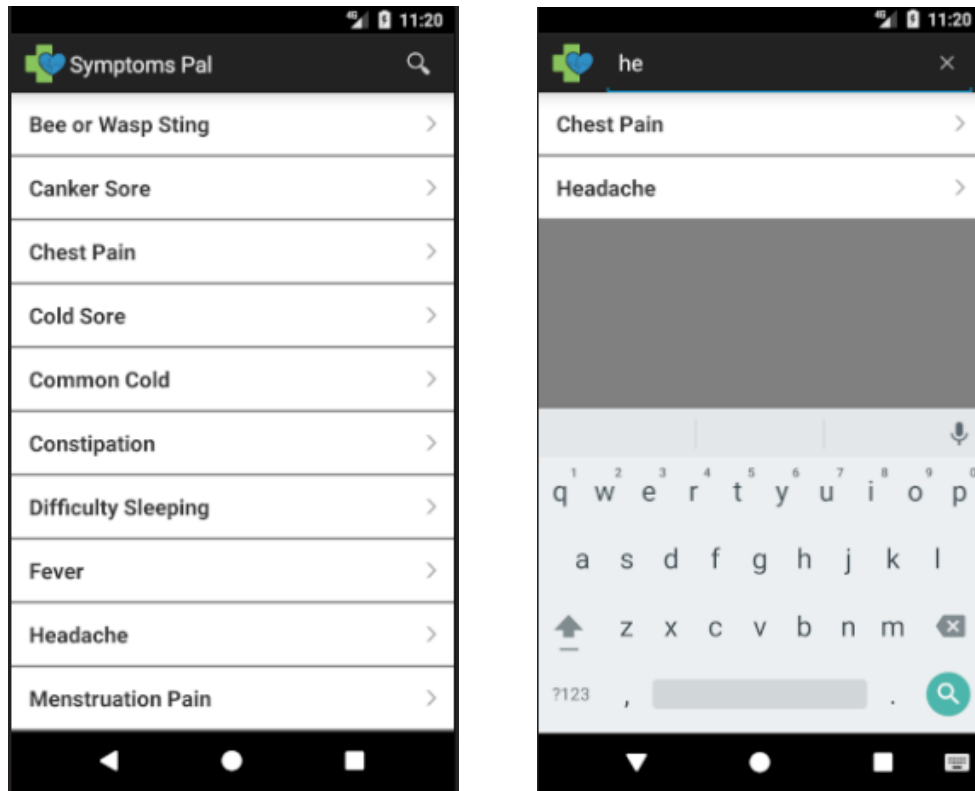


Figure 18 - a) Symptoms list b) Queried symptoms list

4.6.3 Question Page

Upon selection of a symptom, users are asked to answer questions regarding their physical signs and, in some cases, additional symptoms so the Symptoms Pal application could provide the appropriate triage advice. Throughout the process, a progress bar on top of the page tracks user answers and shows them when they will finish the questionnaire (Figure 19).

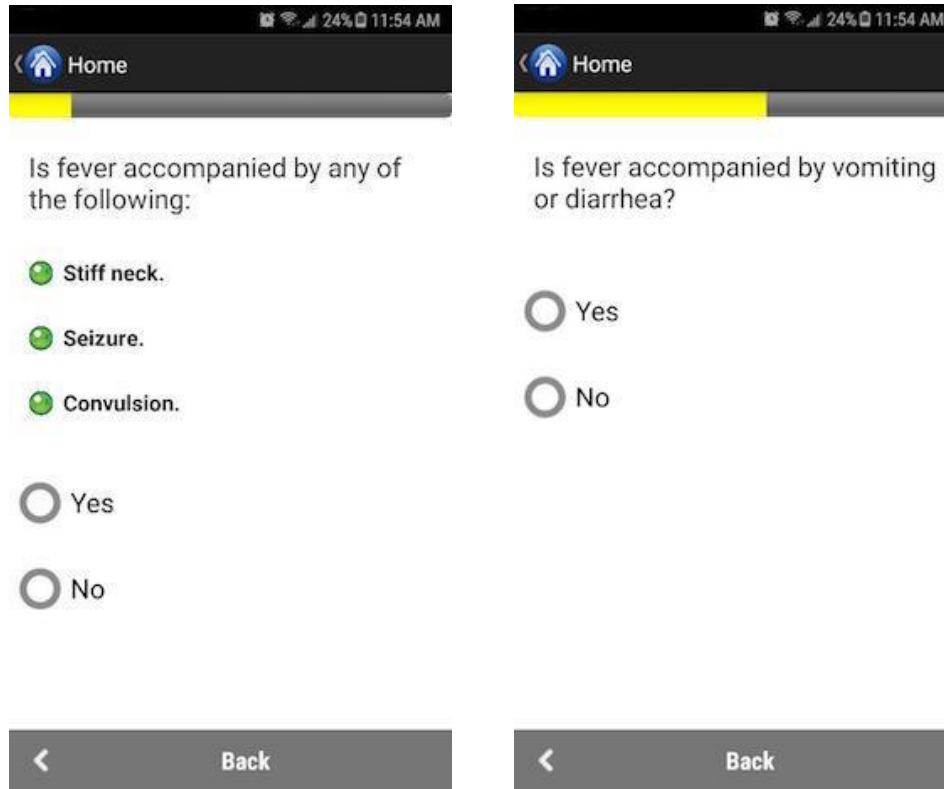


Figure 19 - a) Question page, b) Progress bar

4.6.4 Result Page

There are four different types of triage advice that Symptoms Pal provides:

- **Emergency/Call 911** is for urgent and very serious problems that require immediate attention. A button is provided and, by pressing it, users can call 911 to ask for help (Figure 20a).
- **Emergency Room** is again for urgent problems that necessitate an ED visit. A “Take me there” button redirects a user to Google Maps in order to suggest the nearest EDs (Figure 20b and Figure 22a).
- **Visit a Doctor** is for less urgent problems that still require a physician visit. In these cases, a user can visit their family physicians or visit a walk-in clinic for diagnosis and treatment. In case they want to visit a walk-in clinic, the nearest ones can be found by clicking on “Take me there” button (Figure 21a and Figure 22b).
- **Self-Care**, is for non-urgent problems that can be managed by users and with over-the-counter medications (Figure 21b).

Following the triage advice, Symptoms Pal briefly explains the reason behind its suggestion. In some cases, the application provides additional care suggestions that can be helpful for users prior their visit to a healthcare facility or in cases when self-care is sufficient. Additional information is also provided by linking to reliable medical websites.

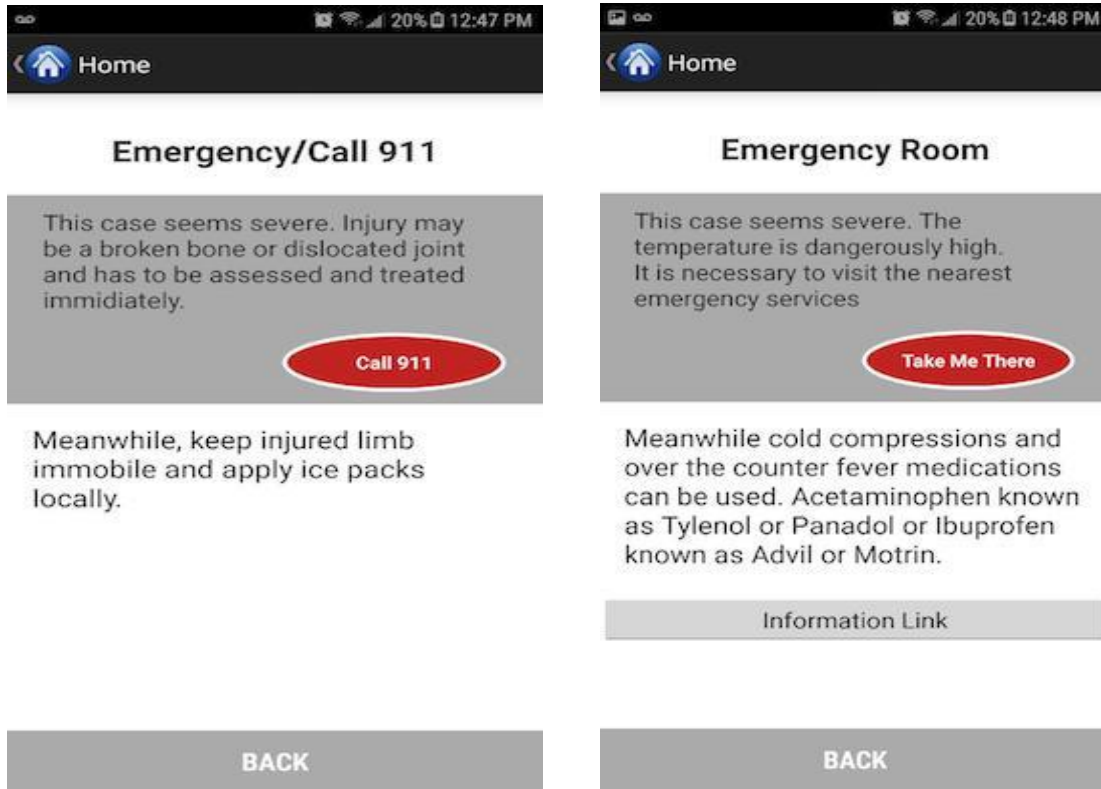


Figure 20 - a) Emergency/Call911 b) Emergency room

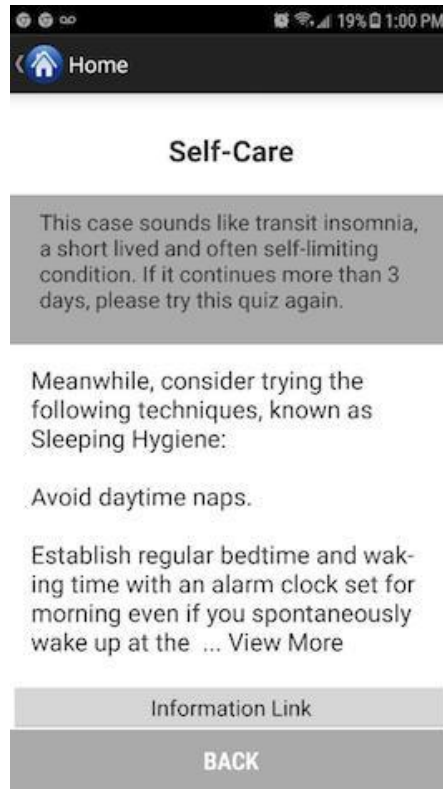
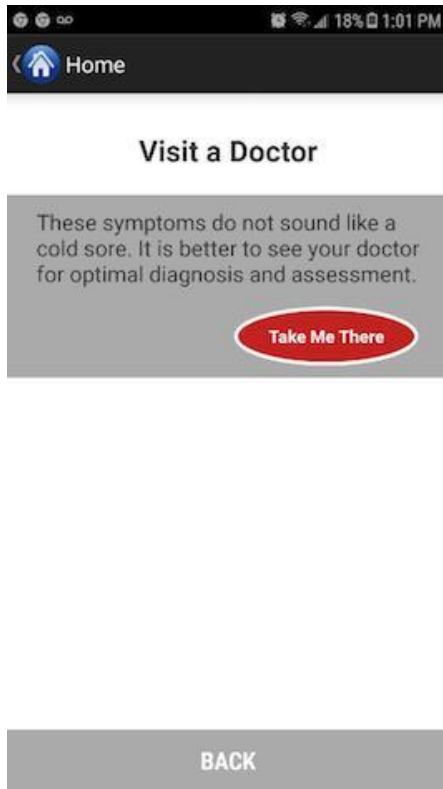


Figure 21 - a) Visit a doctor, b) Self-Care

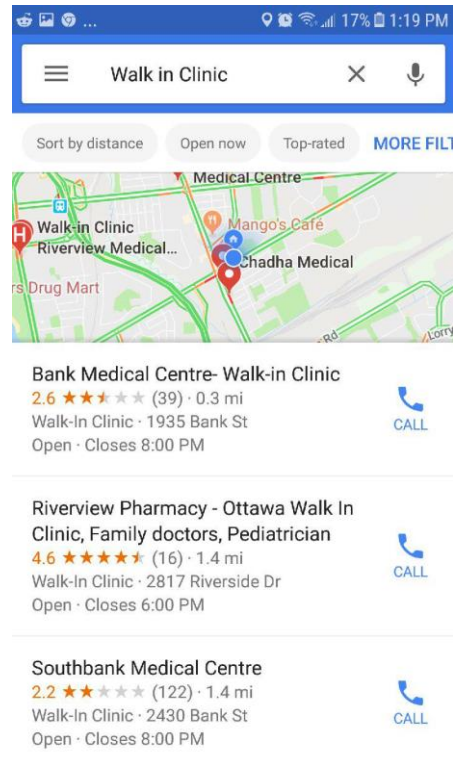
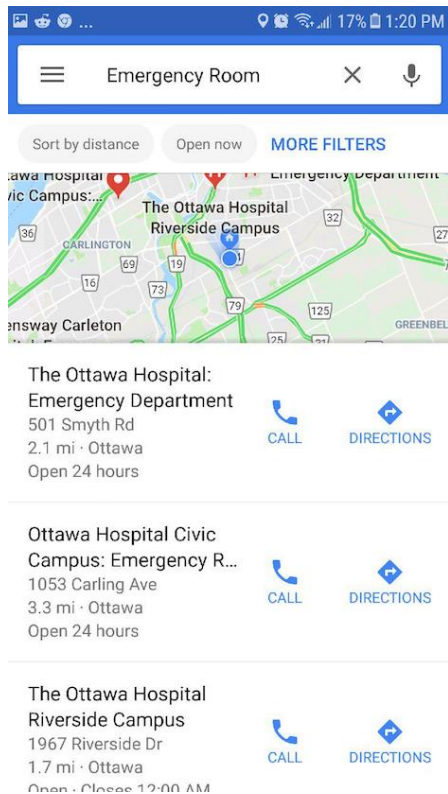


Figure 22 - a) Emergency rooms in Google Maps, b) Walk-in clinics in Google Maps

4.7 Chapter Summary

In this chapter, the Symptoms Pal's goals and requirements were defined in order to design an appropriate solution. The architecture of the application and its main components (data layer, logic layer, and presentation layer) were also presented. The implementation of the application's logic and user interfaces was done using Android Studio v2.3.3 for Android smartphones and mobile tablets.

The next chapter will present the formal usability evaluation that was done and report its results.

Chapter 5 Application Evaluation

As the evaluation stage of the Design Science Research Methodology (DSRM) process, a think-aloud usability study with participants that were not patients was conducted. During the study, user comments and the barriers that they faced were noted. The experiment's objective, methods, recruitment process, findings about users, and questionnaire results are discussed in separate sections. The evaluation described in this chapter and the related documents included in Appendices A to G were reviewed and approved by the University of Ottawa's Research Ethics Board, certificate H10-17-07.

5.1 Objective

The specific *objective* of this evaluation is to validate and refine the UI of the Symptoms Pal application before further development. The validity of clinical contents (the decision algorithms) and diagnostic accuracy are out of scope of this research and therefore were not evaluated. We seek to answer the following *research questions* within the scope of the usability study:

- 1- How effective is the self-triage mobile application in helping users select appropriate symptoms and make a decision?
- 2- What usability aspects of the self-triage mobile application require improvement?
- 3- What additional functional features would users like to see in a self-triage mobile application?

5.2 Usability Study Design

This usability study focuses on users' feedback to improve their experience and the UI design of Symptoms Pal. Two common methods in usability testing and user-interface analysis are used in this study:

1. **Think-aloud protocol** is a usability testing method to gather data, which requires the participant to speak out loud as they are performing a set of tasks. This allows observers to understand the participant's cognitive process [32].
2. **A/B testing**, or split testing, is an experimental method for UI design and user-experience analysis. Two versions of a same applications (A and B) are implemented and used during the evaluation. A/B testing here aims to identify changes to the UI that augment the attentiveness of users [61].

To gather feedback, standard procedures for UI evaluation were followed [33][62]. Specifically, after a brief introduction to the purpose of the study, each participant was asked to fill in a pre-test questionnaire (Appendix C). In this questionnaire, participants were asked to provide information about their age group, their familiarity with symptom checker and self-triage applications, and their past experiences with visiting ED, if any.

Following the pre-test interview, participants were provided with three different scenarios (Appendix G). The scenarios were designed to simulate a real user/patient experience to some extent for the participants. Two different versions of Symptoms Pal were provided to participants to perform A/B testing, one using a graphical human body interface (Figure 16) and another one using a list interface (Figure 18). Each participant should have read the first scenario assigned by the researcher and then complete a set of tasks while following the think-aloud protocol on either version 1 (the graphical interface) or version 2 (the list interface). After completing the tasks using one version, participants had to read the second scenario and redo all the tasks with the other version. The third scenario was read to participants by the researcher in order to simulate a situation where the participant himself/herself does not have any access to the application and someone else is checking the symptoms for him/her. In order to complete this scenario, participants were able to undertake tasks with their desired version. The researcher took hand-written notes about issues of the tool based on the participants' experiences and their thinking process.

Participants were asked to fill in a post-test anonymous questionnaire. In this questionnaire, they were asked to evaluate the usability of Symptoms Pal, the presentation of application's UI, and the perceived usefulness of such application. Answers regarding the usability of the tool and the presentation of user interfaces were recorded using a five-point

Likert scale, while questions about perceived usefulness were recorded through open-ended questions.

Observational notes about participants' performance, barriers that were experienced while performing tasks, and verbal comments were noted and coded, in order to identify emerging themes. A descriptive pattern was then developed to identify how many participants experienced which issues. This gave the thesis author an indication of recurring issues that needed to be resolved.

Answers to the questionnaire recorded on a Likert scale were aggregated (e.g., 9 out of 34 participants "strongly agreed" that the application is learnable). These results can provide good insight about the quality of each UI and the overall usability of this application. Answers to open-ended questions were analyzed in the same manner as the handwritten observational notes.

Note that, as mentioned in the previous section, the accuracy of the clinical content (decision algorithms, suggestions, and triage advice) of this application is not evaluated in this study. To conduct such study, it would be necessary to evaluate the application in appropriate settings (e.g., an ED) with people who are diagnosed by physicians as a comparison point (at least for goal G2). Additional input from people suffering from real symptoms at home would be needed for goal G3. Performing such study would require a higher budget and additional approvals from the research ethics boards of the University of Ottawa and of Montfort Hospital. However, the results of the usability study could still affect the wording and presentation patterns of the clinical content and ultimately the user experience while working with the application. In that context, the Montfort Hospital researchers decided to perform a *clinically-oriented functionality study* (on the accuracy of decision algorithms and triage advice, with ED patients) after finalizing the usability study (this thesis) and applying necessary changes to the wording and presentation patterns.

5.3 Recruitment

Participants are students from the School of Electrical Engineering and Computer Science and the Telfer School of Management at the University of Ottawa. This group was chosen because the goal of this study is to evaluate the usability of this application and not the accuracy of the medical suggestions. Hence, feedback from people who would likely use

such application in the future appeared to be sufficient. However, this study design still leaves some research questions unanswered, such as “what features would users expect in a user interface when they are in acute pain or suffering from recurring pain?”

A convenience sampling method was used to recruit participants [57]. In this method, the researcher/thesis author attended (after granting the permission of the class instructor) different undergraduate and graduate classes in the School of Electrical Engineering and Computer Science as well as in the Telfer School of Management. He then attempted to recruit students with a five minutes presentation describing the purpose of the study. Upon receiving of students’ consent, their emails were collected for future contact. Consenting students were sent an email (Appendix A) with further information and officially inviting them to participate. Students who answered the email were asked to choose a time slot at their convenience to participate to the experiment. It should be noted that the faculty members were not involved in the recruitment of the students, and for preventing biases such as perceived power relationships, the researcher did not make presentations in the classes that were taught by his thesis co-supervisors.

5.4 Participation

The assessment was conducted on an individual basis. Each participant was first presented with a consent letter (Appendix B). The participant was given a chance to ask questions and get clarifications before signing the letter. If the participant declined to sign the consent letter, he/she was excluded from the study.

Included participants were given an introduction to the mobile application and a pre-test interview (Appendix C). A list of tasks (Appendix E) and predefined user scenarios (Appendix G) were then provided to each participant. Since the goal of the assessment was to gather feedback rather than to evaluate participants’ abilities in using the application, each participant was given help to accomplish these tasks when needed.

Upon performing tasks, the participant was provided with a questionnaire intended to answer the research questions (Appendix F). Questions 1 to 4 have been asked in a similar study that aimed to evaluate the usability and learnability of a software tool [37][39]. Questions 5 to 15 were developed from the requirements and features used to develop the application. Questions 16-19 aimed to assess whether participants perceived the proposed

application as useful in real life. The last question was designed to understand users' preference for the UI version (Appendix D). These questions vary from recommendations targeting clinical information systems [36] as the focus here is solely about usability and not clinical content.

The consent letters that identify the names and numbers of the participants are kept locked, in secure location with controlled access. Questionnaires are stored separately, also in a secure location, to preserve anonymity. We aimed to recruit 35-40 participants for the assessment and were able to recruit 34 participants.

5.5 Demographics

Participants came from different age groups: 7 were between ages 17 to 25, 23 between ages 26 to 35, 4 between ages 36 to 45, and 1 was over 46. This variety of ages allowed us to sample different points of views about this application. From 34 students participated in this study, only 2 previously experienced working with a symptom checker application and 13 visited an ED in Canada.

5.6 Predefined User Scenarios

Three user scenarios were designed to provide realistic contexts for the participants in this study. These scenarios were designed by one of the co-supervisors (Prof. Wojtek Michalowski) instead of the Montfort Hospital physicians, in order to avoid possible biased contributions of these scenarios. Each scenario describes the symptoms and physical signs of a known medical condition. The expected outcome of each scenario includes the triage advice with the accompanying partial medical explanation. Participants were asked to accomplish a set of tasks based on these scenarios.

Task 1 asks participants to find symptoms (as described in a scenario) in Symptoms Pal. Task 2 asks participants to answer questions about health conditions as described in the scenario. Feedback then was provided with regards to users' answers. The feedback includes a triage advice, a partial medical explanation, an additional care suggestion, informative links, and the healthcare facility suggestion if relevant. In Task 3, participants simply had to find the healthcare facility suggestion.

Each participant was asked to read a scenario and complete all three tasks with the UI version that was assigned to them. As they completed all the tasks, they were asked to redo these tasks but for second scenario and using the other UI version. The third scenario was read by the researcher in order to simulate the situation where users are using a triage mobile application for a friend or a family member who is in distress. In this scenario, participants were allowed to select their preferred UI version. The difference between the two versions is in their start page. Version 1 (the graphical interface) provides users with a human body pictogram to filter the list of symptoms; however, the symptoms list itself can be accessed by clicking on a button on this interface. Version 2 (the list interface) provides a start page with a scrollable list that contains all the symptoms included in this application. In this interface, users can query the application to find their symptom.

During the entire session, participants were asked to provide comments about the usability of Symptoms Pal, barriers and issues they were facing, features of the application, and the design of the UI. They were also asked to provide feedback about features that could be included in the application, the number of questions they prefer to be asked in such an application, and how they prefer to receive feedback from the application. Figure 23 illustrates the mix of user scenarios and versions used in the usability study.

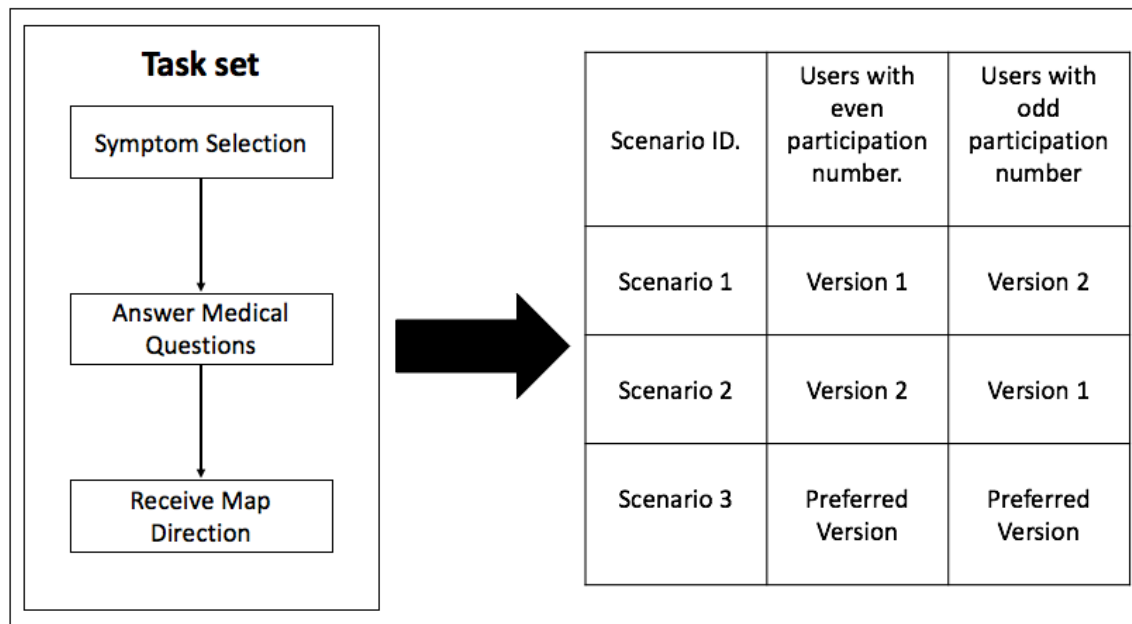


Figure 23 - Predefined user scenarios

5.7 Task Result Analysis

5.7.1 Task 1 – Symptom Selection

Each participant had to complete this task three times with two UI versions. All of the participants were able to find the required symptom using the two UI versions of the application for the scenario 1 and scenario 3. However, for the second scenario, participants had problems when they were attempting to select the right symptom. Participants should have selected *fever* as their symptom to receive the expected outcome for the second scenario. Overall, out of 34 participants, only 17 were able to select fever. Among these, 4 selected fever from version 1 (the graphical interface), and 13 were able to select fever from version 2 (the list interface). This shows that finding the “fever” symptom using the graphical interface is harder than finding it using the list. The reasons could be:

1. The scenario mentioned *headache* as another symptom for the user, so many participants thought that selecting the head for headache would lead to the same result as selecting fever.
2. When first facing the graphical interface, participants did not know what to do since the functionality of the “Common Symptoms” button was not clear and they also were not aware that the areas in the human body pictogram were selectable by clicking on them. As they realized that the pictogram is clickable, their main action was to click on the body parts, which made them neglect the “Common Symptoms” button that actually includes a list of common symptoms such as fever.

Based on the results of this task and on users’ comments, it was realized that the functionality of the human body pictogram with a “Common Symptoms” button was not clear to participants. It is essential to provide better instructions on this page, so users would understand these available features.

5.7.2 Task 2 – Answer Medical Questions

In this task, participants had to answer a set of questions with regards to the medical condition described in each scenario. The results of this task show that much has to be changed in terms of provision and wording of the questions. In the first scenario, 29 participants could not reach the expected outcome. In the second scenario, from the 17 participants who

selected fever, 15 answered correctly and 2 answered incorrectly. For the third scenario, only 14 out of 34 participants arrived at the expected outcome. Although it is difficult to exactly point out one single reason for these errors, the possible explanation for the poor performance of participants while performing task 2 based on scenario 1 and scenario 3 is that participants tend to focus on the physical signs and symptoms in a question and not so much on the actual question that is being asked.

To be more specific, when performing task 2 according to scenario 1, a multiple-item question asks participants to answer “Yes” if they experience a physical sign that is present among a set of options. The first option provided by the question is “vomiting that lasts more than 3 days”. Following their answers to questions, participants were asked the reasons for their answers. Many of the participants replied Yes as they read *vomiting* as one of the physical symptoms mentioned in the scenario. However, the scenario clearly stated that vomiting happened just for a day. Also, when answering the questions based on scenario 3, all 20 participants said that they arrived at a wrong outcome as they saw *burning in front of the chest* in the question and remembered it being mentioned in the scenario itself. However, the question asked by the application was “*Do you experience pain, ache, discomfort, burning, or tightness across the front of your chest after exertion*”. All 20 users ignored the “after exertion” part and answered the question incorrectly.

Based on the results obtained and on the performance of participants on this task, it was realized that the algorithm questions themselves should be better formulated in order to minimize cognitive load. It is better to avoid long sentences in the questions since users tend to read less and focus on their signs rather than on what is being asked in the question. It is also, necessary to break down complex or aggregate questions. For instance, instead of asking one multiple-item question, it is better from a cognitive load perspective to ask each option as a separate question.

5.7.3 Task 3 – Receive Map Directions

In this task, the goal was to determine whether participants could find the healthcare facility suggestion provided by the application. They had to press the “Take me there” button. All the participants were able to understand the functionality of this button and use this feature properly.

5.8 Questionnaire

At the end of the experiment, the participants were asked to complete a questionnaire intended to help us understand how they felt about the self-triage application (Appendix F). The first part of the questionnaire was about the overall impression of the “use and learn” experience, and 4 related questions were included. The results, presented in Figure 24, indicate that participants agreed that the proposed application is easy to learn and use, and that it is easy to understand the output presented (intuitiveness). Most of the participants considered that the solution could help them perform tasks efficiently.

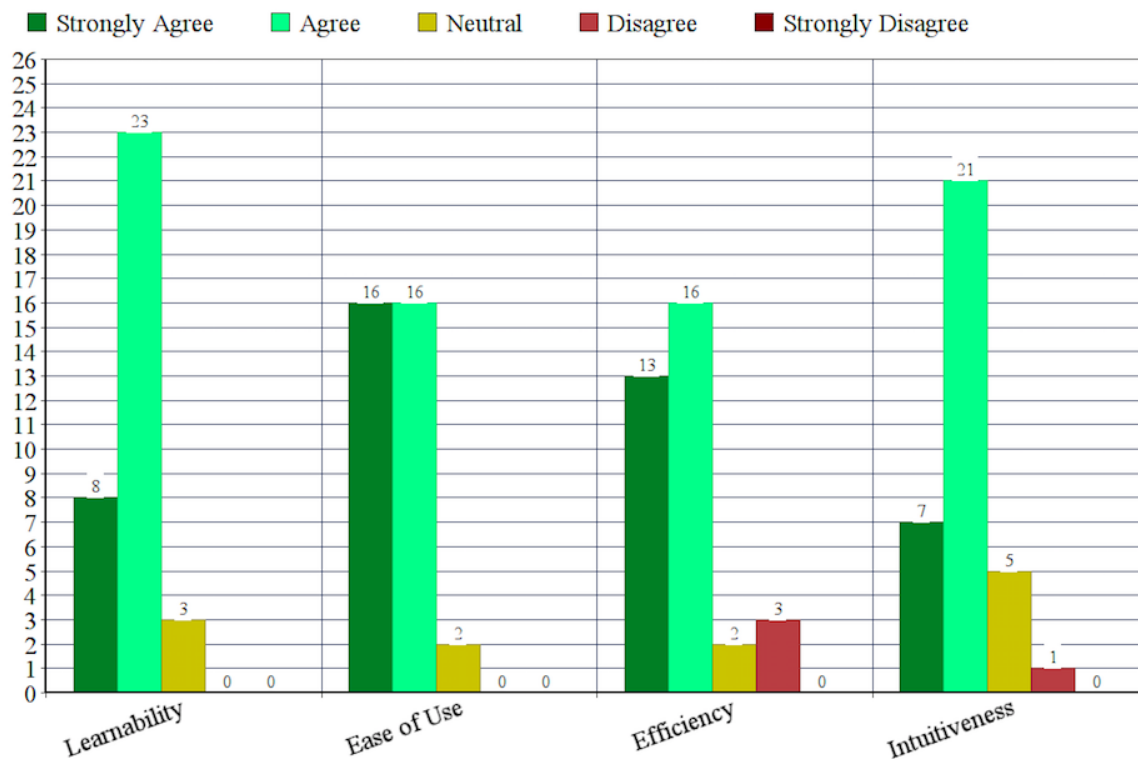


Figure 24 - Questionnaire results for questions 1 to 4

The second part of the questionnaire (Appendix F, questions 5 to 14) is about the UI presentation. The UI of this application should enable users to easily locate symptoms. Four questions were designed to assess symptoms location in the application. Two questions aimed to understand whether users are able to locate their symptom (as described in the scenarios) using the graphical and list interfaces. The other two questions aimed to find out whether users are able to locate an appropriate symptom while checking symptoms for someone

else. The participants had to answer one questions with regards to the version they picked to perform tasks for scenario 3. A question was aimed to find out whether users understood the selected symptom names.

Out of all the participants, 64% agreed or strongly agreed that it is easy to locate symptoms in the list interface. In addition, 73% of the participants agreed or strongly agreed that it was easy to locate symptoms in the graphical interface. From the 23 participants who selected the graphical interface to perform the third scenario, 17 agreed or strongly agreed that it was easy for them to locate symptoms for someone else from the graphical interface, whereas from the 11 users who did the third scenario with the list interface, 9 were able to locate symptoms easily. Out of the 34 users, 26 agreed or strongly agreed that symptom names are easy to understand.

The results show that it is rather easy to locate symptoms using the graphical interface; however, there are some difficulties whose resolution could improve the user experience. The results also tell us that, despite the fact that 9 participants (out of 11) who did the third scenario with the list interface agreed that it is easy to locate symptoms from the list, the 36% disagreement and neutral responses on the easiness of symptom location in the list interface show that this feature is not perfect. The results also show that symptom names are generally easy to understand, although some users explicitly mentioned that they had difficulty understanding the meaning of some of the symptoms. Figure 25 shows the results of the questionnaire about locating symptoms in the UI.

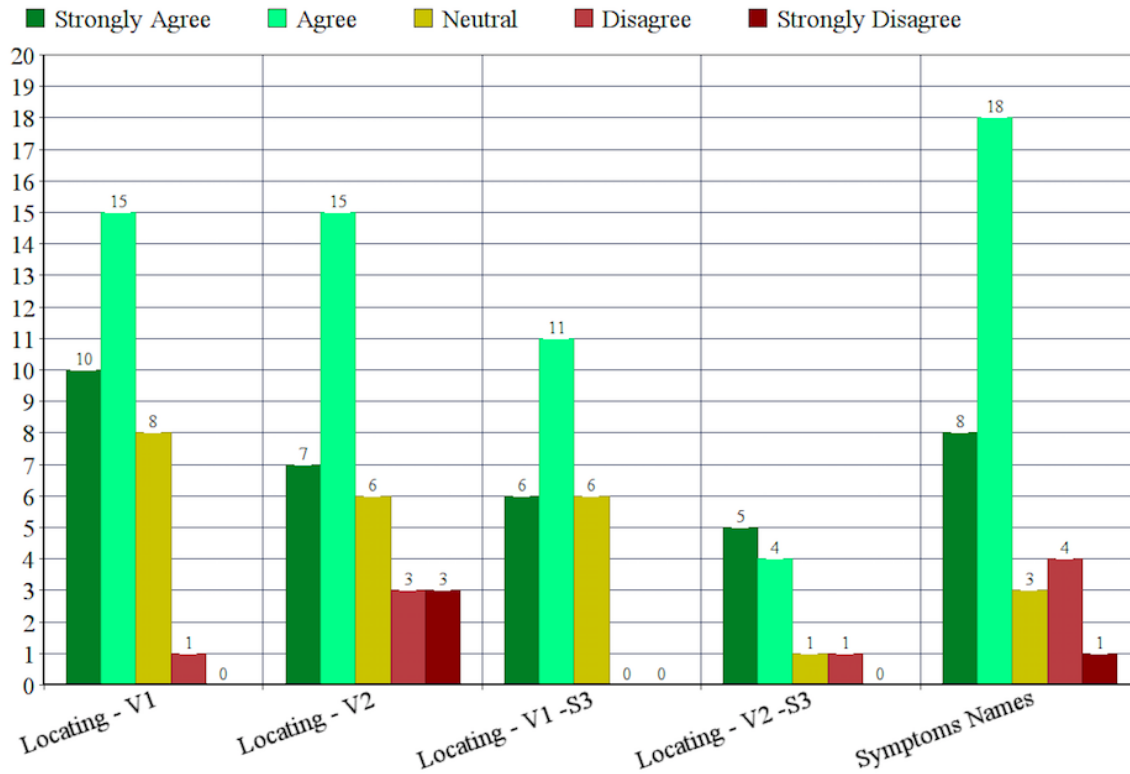


Figure 25 - Questionnaire results for questions 6 to 10

Participants' feedback on the presentation of questions and results (Appendix F, questions 11 and 13) are other aspects of the UI that were evaluated. Figure 26 shows that the presentation of questions needs revisions since 39% of the participants did not agree that the current presentation is appealing. The participants found the presentation of results reasonable with a 73% agreement. However, some useful improvement suggestions were provided.

Two other features of the UI were evaluated in this questionnaire: the progress bar on the question page and the healthcare facility suggestion (Appendix F, questions 12 and 14). Among all the participants, 50% agreed on the helpfulness of the progress bar (see Figure 27). The results indicate that this feature is either not useful or needs improvement. The healthcare facility suggestion by the application was very much appreciated since 32 out of 34 users agreed or strongly agreed that this feature provided useful information.

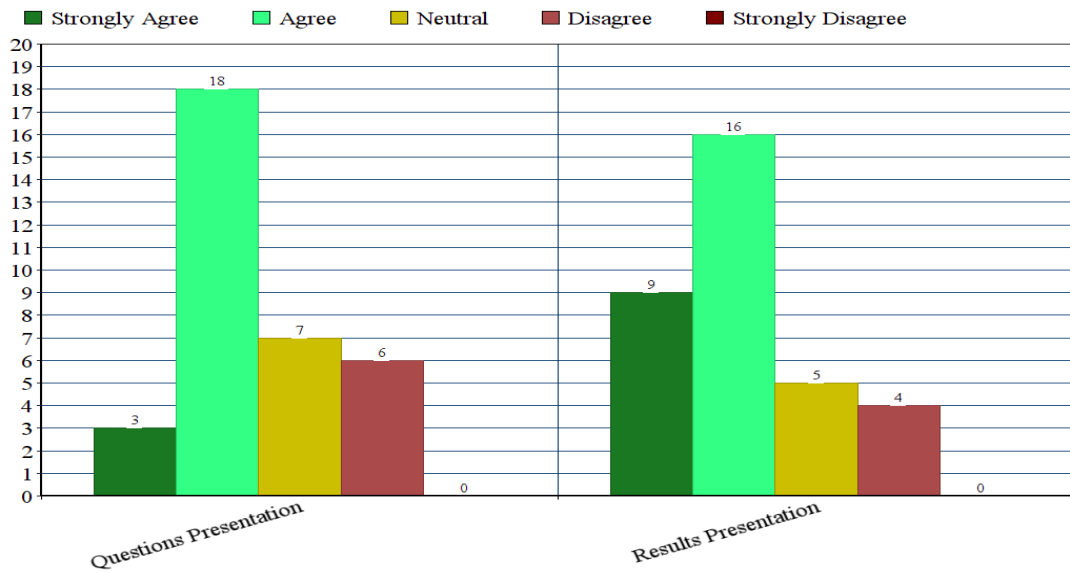


Figure 26 - Questionnaire results for questions 11 and 13

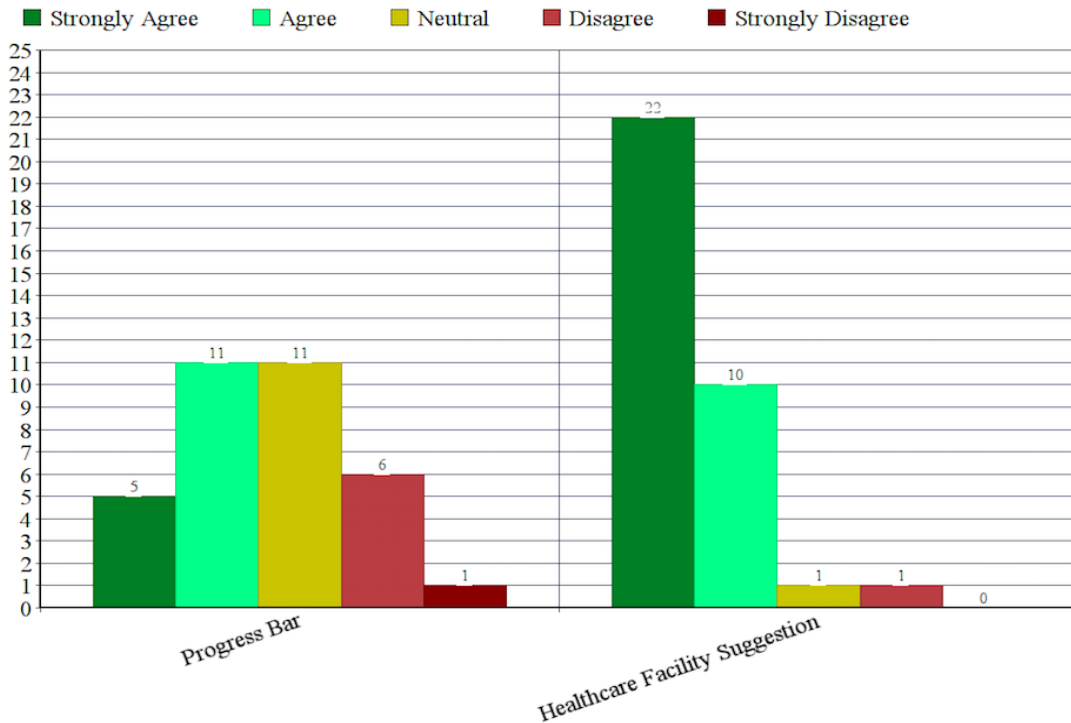


Figure 27 - Questionnaire results for questions 12 and 14

Cognitive load refers to the mental effort required to understand a context. Minimizing cognitive load helps preventing users from becoming overwhelmed while using the application. The cognitive load of an application should be minimized to improve its usability.

The last question about the UI presentation (question 5) aimed to find out whether participants perceived the cognitive load of the application as being low. The definition of cognitive load was provided to the participants and, based on that definition and the participants' experience with the application, participants were asked to assess their cognitive load. Out of 34 participants, 21 (61%) agreed or strongly agreed that the cognitive load of the application is low. This number indicates that there is room for improvement in terms of minimizing cognitive load. Figure 28 shows the results of the questionnaire for cognitive load.

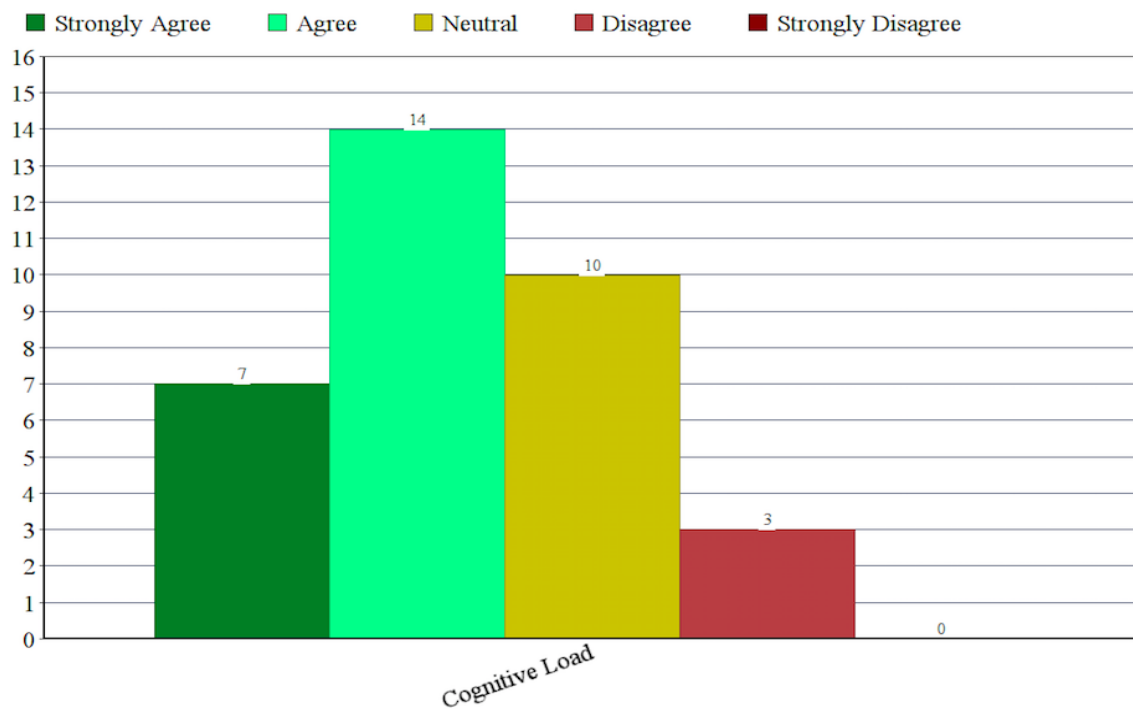


Figure 28 - Questionnaire results for question 5

Participants also had to select their preferred interface for the start page of the application. It was possible to infer the favorite interface based on the participants' selection (23 out of 34 participants ended up selecting the graphical interface, see Figure 25). However, to make sure that the inference is correct, an explicit question was asked about the preferred interface. Figure 29 shows participants' responses of participants to the question. The results show that the graphical interface was favored by users.

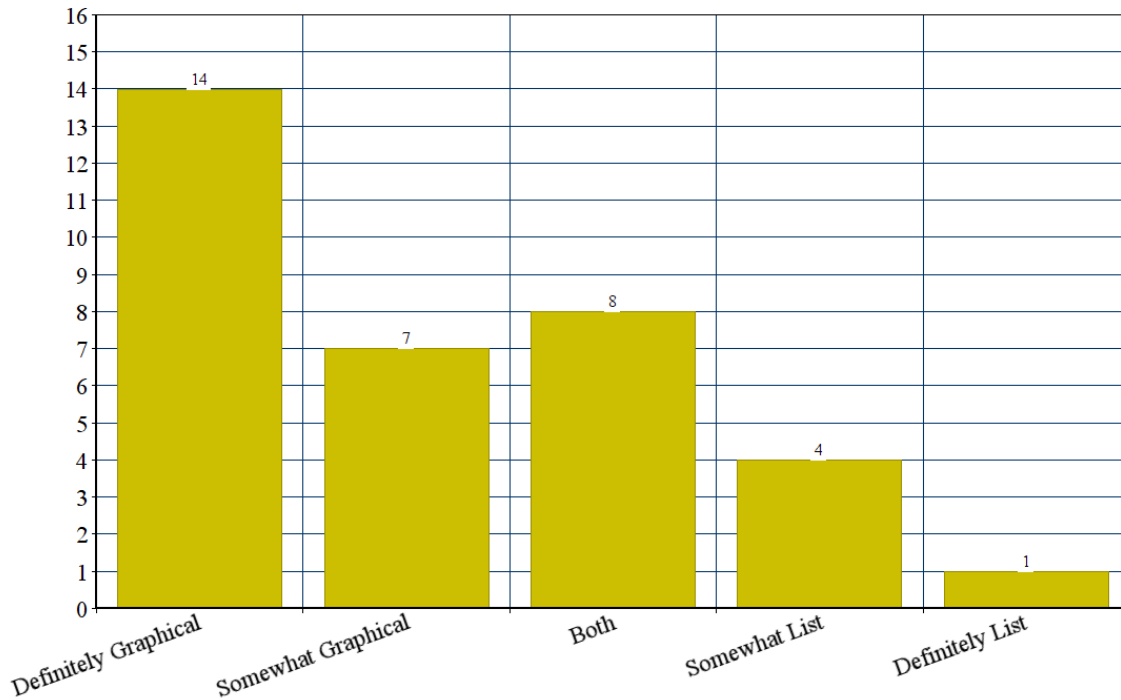


Figure 29 - Questionnaire results on start page UI (Appendix F, question 20)

To analyze the aggregated user satisfaction of the proposed solution, a weighted score scale was developed to calculate the satisfaction score for questions 1 to 14 in the questionnaire. The denominator of the formula comes from the total number of participants, with a maximum of 4 points each. This score is applied to the 12 questions of the questionnaire that all the users answered to:

Satisfaction Score

$$= \frac{4 \times \text{Strongly Agree} + 3 \times \text{Agree} + 2 \times \text{Neutral} + 1 \times \text{Disagree}}{4 \times \text{Total number of participants}} \times 100$$

The radar chart in Figure 30 shows the aggregated satisfaction scores for the 12 evaluated aspects of the Symptoms Pal application. Results indicate that healthcare facility suggestion and “Ease of use” score more than 85%. For 8 questions, the satisfaction scores fall into the 70%-80% satisfaction range. The satisfaction scores for minimized cognitive load, progress bar, questions presentation, and symptoms location from the list are however inferior to 70%.

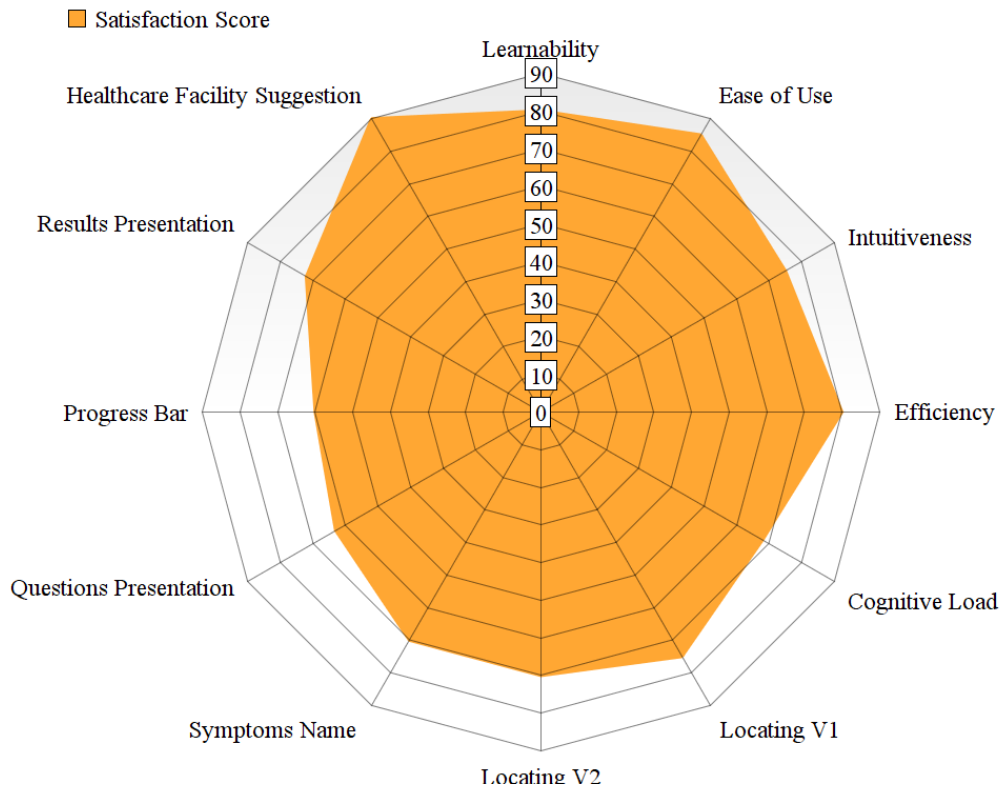


Figure 30 - Satisfaction scores

The users' feedback on different UI features of the application show that the application is learnable, easy to use, efficient and intuitive. However, some features of the application, such as the presentation of questions and the level of cognitive load, should be re-evaluated and if necessary modified. Observational findings from users' performance gave us insight on what should be done in order to improve the users' perception of the features that were less liked. Also, these findings will let us improve the overall quality of the approved features of this application as well. In the next section, barriers that were affecting the users' interactions with different features of this application are further discussed.

5.9 Usability Barriers

Various reasons affected the participants' perception about the usability of Symptoms Pal and its UI. These reasons were interpreted by reviewing participant comments and by studying their answers to open-ended questions (questions 15 and 18 in Appendix F). Issues

identified by the participants were categorized into two groups: design issues and content issues.

Design issues can be defined as deficits in the design of UI elements that cause difficulties in navigating properly through the application. Table 7 summarizes the UI design issues of the Symptoms Pal application.

Table 7 - Application design issues

Issue Code	Description
D1	In the graphical interface (version 1), what users should do next is not clear. They do not know the functionalities associated with the “common symptoms” button and the human body pictogram.
D2	In the question page, with multiple-item questions, there was a confusion regarding the interpretation of some of the GUI elements (a green bullet in particular).
D3	In the question page and the result page, the scrollbar was not visible. When a question has many parts, the answers will be available just by scrolling down the page. When users do not see the question and scrollbar, they assume that a part itself is clickable. That again slows down their navigation in the application.
D4	In the question page, not all the users were able to notice the progress bar on top of the page. Moreover, its shape, color and position do not reflect its functionality to users.
D5	In the result page, the grey and white background did not please the users’ eyes. They prefer to see more contrast and attractive colors.
D6	In the result page, the “Take me there” button with an oval shape and red background immediately grabs the users’ attention, and make them click on that button. This distracts users from reading the explanation provided.
D7	In the result page, when a triage suggestion is to visit the ED by calling 911, the oval shaped button that allows users to call 911 seemed intimidating to them.
D8	In the result page, users prefer to see colors different from the background of the triage advice suggested on the top of the page. Many users suggested that the colors can indicate the level of urgency of their problems even better than the triage advice itself.
D9	In the result page, when providing additional information, the font of the text is larger than the font of the medical explanation in the grey box. This inconsistency in the font size can lead users to think that the content of the additional information is more important than the medical explanation.

Issue Code	Description
D10	In the result page, when providing additional information, it would be better to have a small title, so users would realize how this additional information can help them.
D11	In the result page, sometimes the additional information is long. A “view more” link is provided so they would be able to expand the text. This “view more” link has the same color as the text. This makes users not notice the view and so they ignore it.

The second group of issues is related to the medical content and design of the decision trees / algorithms. These issues are listed in Table 8

Table 8 - Application clinical content issues

Issue Code	Description
C1	Some symptoms’ names are too medical and difficult for users to understand. They would prefer having lay and common names.
C2	In questions, too much text is usually provided. That was increasing the level of cognitive burden to understand the question. Hence, by having long text, users might miss some important information in the question and make an error while selecting an answer.
C3	In questions, sometime there are too many parts/items. Such composite questions, again, increase users’ mental effort to understand the questions and might lead them to make an error while selecting their answer.
C4	In questions, users tend to focus on keywords and so they have a tendency to ignore logical “AND” and “OR” conjunctions. It is better to avoid such conjunctions.
C5	In questions, immediately providing advice after receiving answers to a small number of questions (1, 2, or 3) is perceived as not useful since users might answer wrong to a question. Users also might think that not enough information was not collected from them, make users question the reliability of the content/algorithms.
C6	In questions, when required to consider options in an aggregation, users tend to select an option rather than answering “Yes” and “No” to the question. To be more specific, if question asks whether users have “any of the following conditions” and then lists a number of conditions, users tend to select the condition itself rather than check all conditions and answer Yes or No to the question.

Issue Code	Description
C7	In questions, users sometimes do not know the answer to a question. They would prefer having an “I do not know” or “I am not sure” option for such situations. This, again, is related to clinical content since by adding this option to answers, the structure of current decision trees would need to change.
C8	In questions and results, sometimes many medical terms were used and users did not understand these terms.
C9	In questions and results, if having complex medical terms is necessary, users do not want to read the explanation of such terms on the same page. Instead, they would prefer having a help button or a hyperlink underlying the medical terms.
C10	In results, while providing advice, labeling a case as severe or serious will scare users about their situation.

These barriers partially explain why users did not respond well to some features of the UI. The progress bar, the presentation of questions, and the level of cognitive load received the lowest satisfaction scores (60%, 63%, and 68% respectively). D4 and C5 could explain why the progress bar received the lowest score. Three design issues related to the question page (D2, D3, D4) and eight content issues related to question content (C2-C9) could explain why users were not satisfied with the presentation of the questions. D1, D2 and D3 indicate that users were confused while working with different interfaces of the application. C2 and C3 point out issues that increase the level of cognitive load to absorb application content. Together, these barriers can show, to some extent, why some users determined the level of cognitive load as high, however, it is not possible to confirm the reasons.

Other aspects of the application received reasonable satisfaction scores. However, symptom names (70%) and feedback presentation (72%) could still be improved. Some suggestions from the participants about design issues were implemented in a revised version of the Symptoms Pal application reported in section 5.11.

The next section discusses other findings from participants about perceived usefulness of this application and their preference on patterns of presentations of the questions and the results.

5.10 Findings from Participants

Two open-ended questions in the questionnaire (Appendix F, questions 16 and 17) were designed to help better understand whether users perceive this application as useful or not. Comments were also collected from participants informal verbal questions asked during the evaluation session.

5.10.1 Perceived Usefulness

Questions 16 and 17 of the questionnaire aimed to understand the perceived usefulness of this application. Question 16 asked users about “*in what types of health concern decisions (if any) do they think that a self-triage mobile application could help?*”. Eight participants did not respond to the question and Three participants answers were not related to the question. Participants answers are categorized as follow:

- P1: Non-urgent or less urgent health conditions (headache, cold, etc.). (12 participants)
- P2: Urgent health conditions (stroke, heart attack, etc.). (6 participants)
- P3: Any unknown medical situation. (5 participants)

Among 12 participants who stated P1, two participants added that if they had a common problem such as a headache, and they had done nothing for a couple of days about that headache they would use the application. Three of the users categorized as P5 added additional explanations. One participant added that he will use the application instead of search engines. Another one stated that he will use the application for unknown situations (whether urgent or not), that do not seem to need an expert knowledge for diagnosis. Lastly, a user mentioned that he also looks for the suggestions about whether he should buy over the counter the medicine to take care of his problem.

Question 17 asked participants whether they would use the application in its current form when having some health concern(s) and if Yes, based on what feature(s) they want to use the application. Four users did not answer the question. Nine answered No to this question. Twenty-one participants replied Yes and 19 of them identified favourite features. Participants’ answers are categorized as follow:

- H1: The graphical interface. (5 participants)
- H2: Triage advice provided by the application. (8 participants)
- H3: Reliable information. (3 participants)
- H4: The list interface. (2 participants)
- H5: Healthcare facility suggestion. (1 participant)

Two users who are categorized as H2 mentioned multiple features as their favourites. A user stated that in addition to the triage advice, the healthcare facility suggestion is another intriguing feature of this application. Another participant mentioned that besides, triage advice and the healthcare facility suggestion he/she likes to use the application because it is a reliable source of information.

Based on the collected answers, it could be concluded that the tool is perceived as useful for most of the users. Users tend to use this application for different urgency levels, and because of different features that the application provides. However, some users mentioned that they would not use this application in its current state. About 50% of users who answered to the question 12, replied that in case they want to use this or any similar application, they prefer to check the symptoms that are less likely to be urgent. 6 mentioned that in case they use the application, they would like to make sure that their concern about the urgency of a situation is valid. Also, some users mentioned that such tools are helpful for any kind of medical situation.

5.10.2 Additional Findings

In each session, before asking participants to fill out the questionnaire, users were verbally asked: “How do you want questions to be presented and how many questions do you want to see in such applications?” Most of the participants replied that they would prefer to see a short text for a question that does not have multiple items/parts/options. In such case, participants would prefer to select an item rather than reading all items (aggregated by a AND or OR conjunction) and give answers somewhere else. Some participants mentioned that they would prefer to type down their answer to the question. They also mentioned that they do not want to see medical terms in the questions.

In terms of the number of questions, 32 participants mentioned that they would not trust an application that provides triage advice and feedback with a few questions (1, 2, or 3), and they would prefer to answer more questions. Different opinions exist about the preferred number of questions that should be asked by the application. Two participants mentioned that they are not sure about the answer to this question since they thought that if more questions were asked the results should be more accurate, however, they would prefer not having to answer too many questions while checking symptoms. Also, one participant would prefer to measure the effort by the time that it takes for an application to provide a useful advice. That participant mentioned that if it takes more than 10 to 15 minutes for the application to provide a suggestion, it is no longer helpful to the users.

A total of 27 participants gave a range of values for the ideal number of questions. The number of participants and the ranges suggested by them are categorized in Table 9. A few participants mentioned that they are not comfortable with a few questions (up to 3), however, they did not state how many questions they would like to see. These participants are categorized as ≥ 4 . One participant clearly stated that he/she wants 10 questions at least, and the more questions the better (≥ 10).

It should be mentioned that, for categorization purposes, some of the ranges were merged together. For instance, four participants chose a [2-5] range, two participants chose [3-5], and one mentioned [1-5]. Since the purpose is to find a ballpark figure for the number of questions, such answers were categorized as [2-5]. Also, one participant stated that he/she would prefer “on average 6 or 7 questions”, which is also categorized as [5-10].

Table 9 - Users’ preferred number of questions to ask

Range	1-4	2-5	4-6	5-10	≥ 4	≥ 10
Number of participants	1	7	2	7	9	1

Participants also were asked, to provide their suggestions about how they would like to see the organization of a presentation of feedback. Out of all the participants, 30 responded to this question, 21 of whom stated that they want something close to the pattern that is implemented in the application. The pattern in the application includes triage advice, a partial but not detailed explanation regarding the triage advice, some additional care suggestions

about the case, and a healthcare facility suggestion. Also, additional informative links for some cases are provided, depending on the case and the triage advice.

The participants preferred to see more information about how to approach the problem in case of self-care, such as suggestions for treatment, but less information for urgent cases. Many participants wanted to have an option to read more from valid sources for *all* the cases. There were participants who suggested that upon receiving the suggestion, it would be good to have a link telephone health line information to confirm the recommendation of the application. Also, some participants suggested that the application should always provide partial explanation and additional care suggestions within the current pattern.

Five participants stated that they would prefer to see many diagnostic results based on their likelihood of occurrence and a short explanation about each option with one triage advice. In that group, some participants mentioned that, along these explanations, additional information should be available, whether via a link to valid medical sources or within the application, but it should be provided upon request. Two participants stated that they want long explanations about the conditions with treatment suggestion and medication, whereas two participants just stated that the triage advice and a healthcare facility suggestion would be enough.

As participants' responses show, it is hard to propose a definitive pattern that could satisfy all needs, but it could be concluded with high confidence that users do not want to invest too much time working with the application. They tend to suggest answering between 4 and 10 questions and want to be provided with feedback that gives them insight into their problem with triage advice and healthcare facility suggestion. They do not want to get information about the treatment and medication advice as they prefer to receive this information from medical professionals. It should be mentioned that most of the participants did not have any experience with such application. Hence, participant feedback could be different if they had previously used other symptom checkers or self-triage applications with different presentation patterns different than the one used in Symptoms Pal.

5.11 Proposed Modifications

As a result of the assessment, 12 new requirements were identified. These requirements are corresponding to the design and clinical content issues discussed in Table 7 and Table 8. All of these new requirements satisfy goal G1 mentioned in Table 4, which is to improve the overall patient/user experience, reduce uncertainty, and provide reliable information. Eleven design issues were noticed from the users' perspective. Also, some users presented design suggestions that could have improved the quality of this application. These suggestions were taken into accounts in addition to issues that were discussed earlier in Table 7. Modifications were applied to resolve many of the mentioned issues. clinical content issues still exist in the application since only physicians are allowed to modify and edit the algorithms and their questions. Table 10 shows the new requirements and the corresponding issues.

Table 10 - Additional requirements

Id.	Requirement	Is- sue	Satis- fied
R14	Necessary instructions for navigation through the application shall be included.	D1	Yes
		D2	Yes
		D3	Yes
		D4	No
		D10	Yes
		D11	Yes
R15	The shape of the UI elements shall properly represent the expected functionality.	D6	Yes
		D7	No
		D9	Yes
R16	UI elements shall be color coded.	D5	Yes
		D8	Yes
R17	The symptoms, questions and answers shall avoid complex medical terms.	C1	Out of scope
		C8	Out of scope
R18	The questions shall not be long.	C2	Out of scope
R19	Multiple-item questions shall be simplified and avoid leading to page scrolling.	C3	Out of scope

Id.	Requirement	Is- sue	Satis- fied
R20	Questions shall avoid compound statements with logical “AND” or “OR” conjunctions.	C4	Out of scope
R21	Algorithms shall avoid providing advice within three questions or fewer.	C5	Out of scope
R22	In case a multiple-choice question is provided, the application shall allow users to select an option explicitly.	C6	Out of scope
R23	To allow users to properly respond to all the questions, the application shall offer an “I do not know” or “I am not sure” option.	C7	Out of scope
R24	In case it is necessary to use a less-known medical term, explanation of the provided term shall be made available to users upon demand.	C9	Out of scope
R25	The language of the application shall not discourage or scare users.	C10	Out of scope

The revision of issues related to clinical contents and algorithms are out of scope of this research, and hence corresponding requirements are not satisfied. Three requirements (R14, R15, R16) corresponding to design issues are attempted to be satisfied in the following ways:

1. **R14:** Issues D1, D2, D3, D4, D10, and D11 are covered by this requirement. In order to resolve issue D1, another page (activity) was added to the application that first asks users about their common symptoms (see Figure 31a). To resolve the problems stated in D2 and D3, green bullet points were changed, and the scrollbar was adjusted on the side of the page (see Figure 32). D4, which is concerned with the visibility of the progress bar in the question page, was not changed as the main problem with the progress bar is the single-content issue mentioned in Table 8 (C7). This problem should be fixed first, so the progress bar could function properly, then remaining problems could be taken into consideration. The provision of a title for additional information should resolve D10. To resolve D11, the color of the “view more” link was changed to blue, so users could notice the link more clearly (see Figure 33).

2. **R15:** Issues D6, D7, and D9 are covered by this requirement. To resolve issue D6, the background color of the “Take me there” button was changed from red to white in order to soften the perceived urgency of the message. In addition, a Google Maps logo was added, as suggested by one of the users. D7 was about the “Call 911” button intimidating users when receiving it on the result page. Also, it was not clearly mentioned in the list of requirements (Table 5) that this feature should have been provided. Hence, this feature was eliminated from the application. Changes in the text size were made to resolve D9 (Figure 33b).

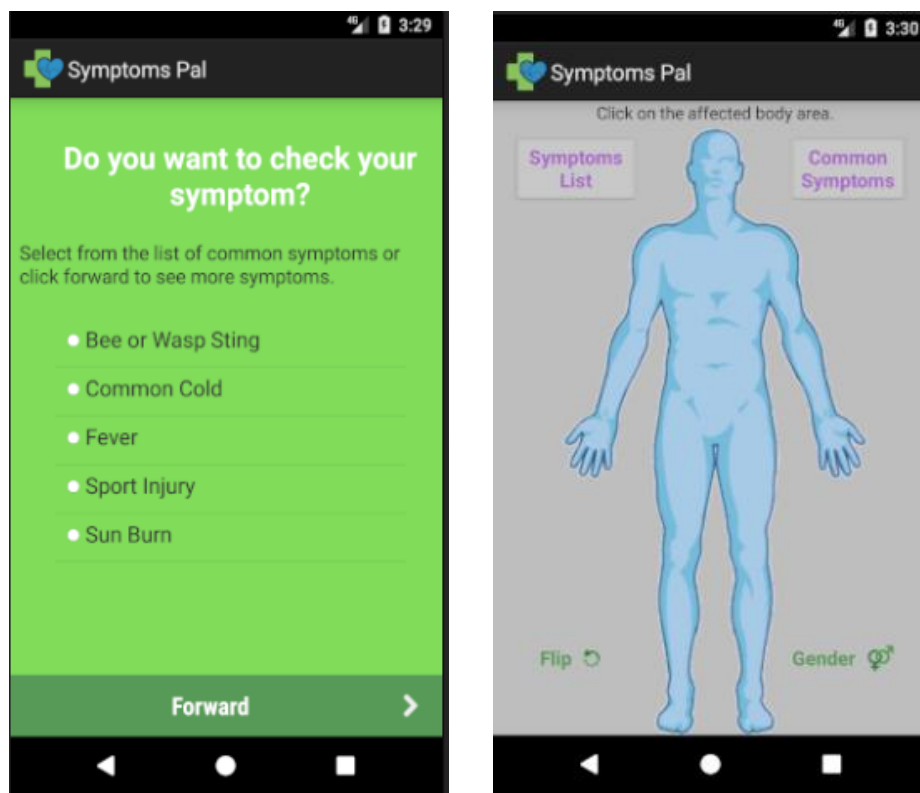


Figure 31 - a) Common symptoms interface, b) Modified start page

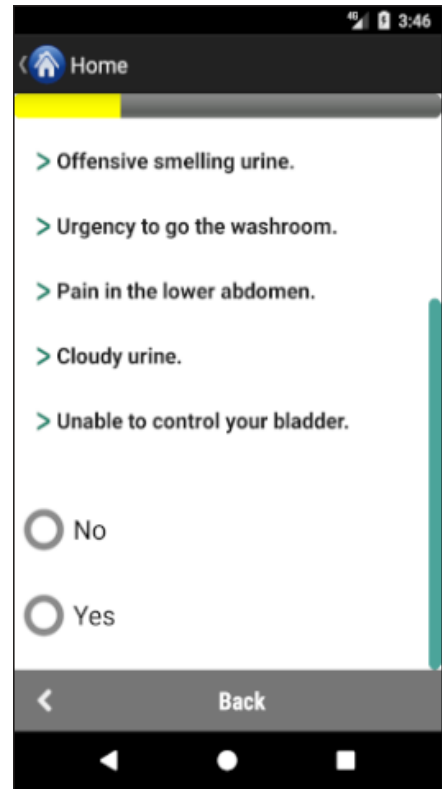
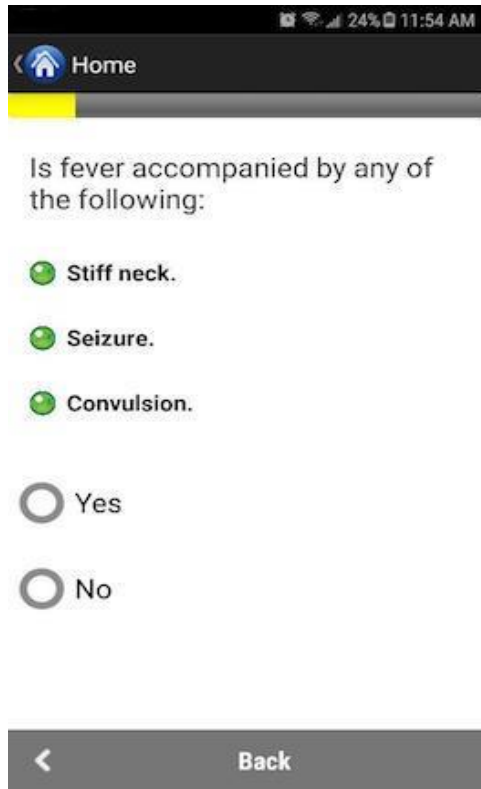


Figure 32 - a) Old question page, b) Modified question page

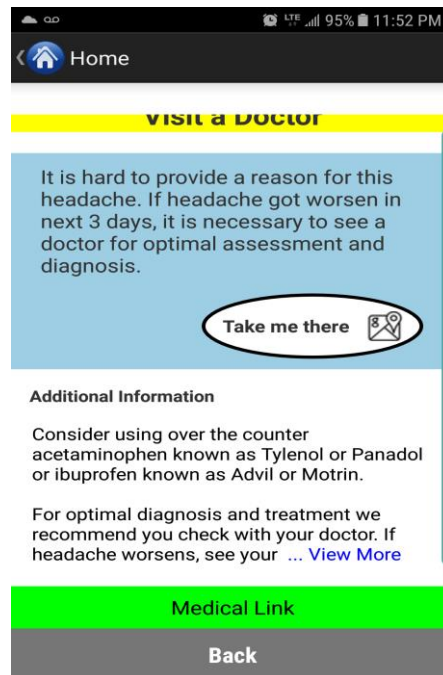
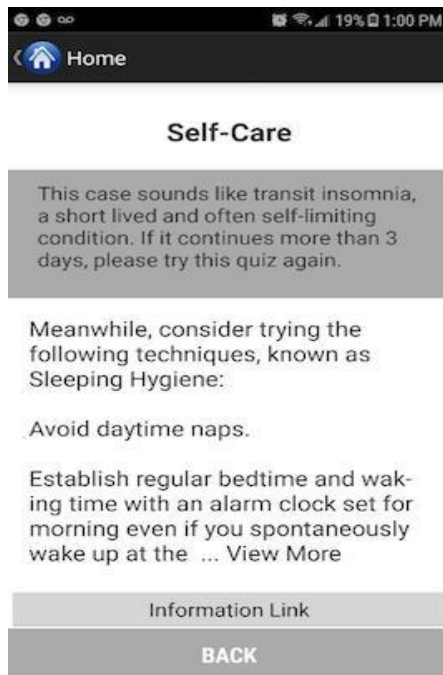


Figure 33 - a) Old result page, b) Modified result page

3. **R16:** Issues D5 and D8 are covered by this requirement. D5 states that a grey and white background is not satisfactory and that it would be better if brighter colors were used in the application. In order to resolve the D8 issue, the triage advices now use a background color that suggests the urgency of the situation (see Figure 34).

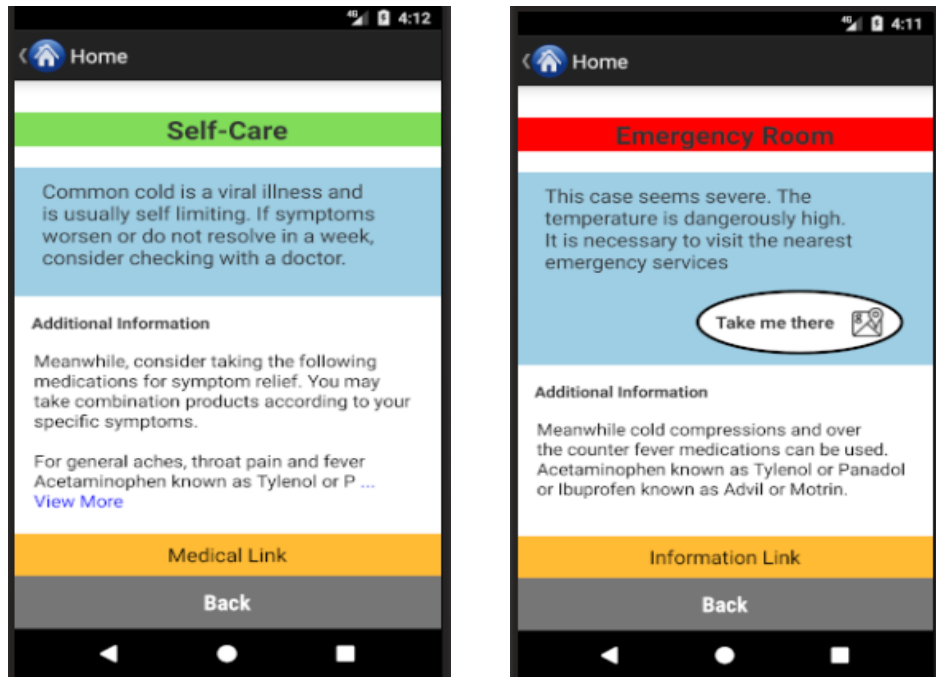


Figure 34 - a) Self-Care color coded in green, b) Emergency Room color coded in red

5.12 Chapter Summary

In this chapter, the evaluation of the Symptoms Pal application was described. Descriptive analysis was performed on tasks results, and the questionnaire results were thoroughly discussed. Design and content issues that exist in the application were identified and changes that were applied to the UI were described. It was concluded that participants perceived this application as useful with regards to the implementation of various features. Additional findings from users were covered about the number of questions they prefer to see in such an application, and about the organization of feedback presentation. It was concluded that users do not want to invest too much time in using such application and they prefer to use this application in order to receive reliable information, triage advice, and healthcare facility suggestions.

The next chapter compares the proposed mobile application with closely related approaches and reviews the discussion that was done in this chapter about various features. Limitations and potential threats to the validity of such application and the study itself are also presented.

Chapter 6 Discussion

In this chapter, the application is compared with similar applications that meet most of the measurement criteria presented in section 3.3. In addition, the limitations and threats to the validity of this research are discussed.

6.1 Technical Comparison

This section provides a comparison of Symptoms Pal (SP) with other competitors presented in section 3.3, where different existing applications were selected based on inclusion/exclusion criteria. These applications were then compared with regards to measurement criteria in Table 3. Table 11 contains the applications that satisfy most of the measurement criteria. A comparison is done of those applications and the mobile application introduced in this thesis (SP).

As shown in Table 11, most of the applications are attempting to provide reliable information about different symptoms on multiple platforms. All of the applications here do not allow users to have access to information without a connection to the Internet. Also, three applications are not customized to locations. Four applications are supporting multiple languages and, except for S12, no evidence was found about the usability and functionality evaluation of applications in peer-reviewed literature.

The application introduced in this thesis allows users to have access to information even without a connection to the Internet. Its content (symptom-based triage algorithms) is designed by a group of doctors at Montfort Hospital and is being validated. It is customized to location and can recommend the closest healthcare facilities that could take care of a user problem.

Multi-platform access is partially covered in this application. The application is working on all the Android platforms (mobile, tablet). However, the iOS version of Symptoms Pal is not completed yet. A thorough usability study was conducted in order to help us discover flaws in this application, and also to assess the usability of this application. As

reported in section 5.10, the majority of the participants in the study perceived the application as useful for various health concerns. As discussed in the literature review, the provision of diagnoses and user data collection are features that could affect users negatively and add to their concerns. Hence, to avoid any conflict, these two features are not provided in Symptoms Pal. One side effect is that Symptoms Pal does not learn or adapt based on its users' inputs (i.e., it is static).

Table 11 - Comparison with other applications.

Application	Offline Accessibility	Multi-Platform Access	Medical Information Supervision	Customized to Location	Multi Language Support	Usability and Functionality Validation	Data Collection	Availability of Self Diagnoses	Type
S5	No	Yes	Yes	No	Yes	No evidence	No	Yes	Static
S7	No	Yes	Yes	Yes	No	No evidence	+/-	Yes	Static
S8	No	Yes	Yes	Yes	No	No evidence	No	No	Static
S11	No	Yes	Yes	Yes	No	No evidence	+/-	Yes	Dynamic
S12	No	Yes	Yes	No	Yes	[44]	+/-	Yes	Dynamic
S14	No	Yes	Yes	Yes	No	No evidence	+/-	Yes	Dynamic
S15	No	Yes	Yes	Yes	Yes	No evidence	No	Yes	Static
S16	No	Yes	Yes	No	Yes	No evidence	No	Yes	Dynamic
SP	Yes	+/-	Yes	Yes	Yes	Yes	No	No	Static

6.2 Limitations and Threats to Validity

As proposed by Perry et al. [52], three types of validity threats are discussed here:

- Construct Validity: verifies the hypothesis/systems claims based on tests performed.
- Internal Validity: examines any bias in performing the research.

- External Validity: determines whether the extension of experimental results to other cases or situation is possible or not.

6.2.1 Construct Validity

The major threats to construct validity are:

- The clinical content of the Symptoms Pal application was developed by a group of physicians and its accuracy was checked by them. However, the presentation of the symptoms, questions, and answers is still too oriented towards medical practitioners and not sufficiently towards lay users. How accurate the triage results are also remains under investigation. These issues may threaten the perception of some users.
- The number of symptoms that the application supports is limited. The selected symptoms may not reflect the most frequent reasons for visiting ED and hence might affect the satisfaction of one goal of this application (reducing the number of inappropriate visits to ED).

6.2.2 Internal Validity

The major threats to internal validity are:

- The reviews of the literature and of existing applications, as well as the usability study, were mainly done by one person (the thesis author). This might have introduced some bias.
- The usability study conducted is solely reporting Symptoms Pal's features and other applications described in Chapter 3 were simply evaluated based on the thesis author's experience with these applications.

6.2.3 External Validity

The major threats to external validity are:

- Most participants in this study were inexperienced in working with symptom checkers. Only two participants had worked with such applications. Hence, the comments collected could be solely based on the users' experience with Sym-

toms Pal. For instance, most of the users approved our pattern for the presentation of results, however, their opinion could have differed should they have experienced another symptom checker with a different pattern for the presentation of results.

- In this study, healthy university students participated to the evaluation (convenience sampling). Whether the conclusions apply to people who are experiencing sickness, who are older than the current participants, or who have a lower level of literacy or education remains a question.
- Participants were performing tasks based on predefined scenarios, not on their real-life situations.

6.3 Chapter Summary

This chapter discussed how the self-triage mobile application compares to commercial applications, and several benefits were highlighted. Threats to the validity of this research were identified and summarized. The liability and presentation of clinical content are major concerns that need to be addressed. The next chapter provides concluding remarks about this thesis and discusses future work items.

Chapter 7 Conclusion and Future Work

This chapter recalls the main contributions of the thesis and discusses important future work opportunities.

7.1 Conclusion

Closely-related academic research and commercial solutions were reviewed to understand the domain and find gaps in existing work. Based on meetings with stakeholders and the review results, requirements for a new self-triage application were defined. The architecture of the application then was created to satisfy the requirements. After a fully-functional solution (Symptoms Pal) was developed, a usability study was conducted to evaluate the features of the application and to understand the perception of users towards the application. The proposed application was also compared with competing applications against a set of measurement criteria.

This thesis contributes the requirements, architecture and usability assessment of a self-triage mobile application developed for Android devices. The Symptoms Pal application provides users with means to self-triage in order to prevent them from visiting an ED inappropriately, and also to encourage them to visit a healthcare facility in case of warranted medical needs. The main research objectives are satisfied in the following way:

- The main research objective here was to determine the usability of Symptoms Pal. A study involving 34 participants suggests that the application is somewhat usable. The results of the usability study showed that most of the features of this application are approved by the participants. Users assessed the proposed solution to be learnable, efficient, easy to use, and intuitive (section 5.8). However, some features provided in the application did not completely satisfy users, because of different design and content issues existing within the application (section 5.9).
- The secondary objective of this application was to improve the understanding of features and functionalities that users would like to see in general in self-triage ap-

plications. The results of the usability study allowed us to provide many suggestions (in the form of requirements) to developers about the presentation of questions and medical suggestions (section 5.10.2).

The main thesis goals are hence satisfied in the following ways:

- 1- Twenty-five technical and reusable requirements were identified and are presented in Table 5 and Table 10.
- 2- A sample architecture was designed and implemented as discussed in Chapter 4.
- 3- A usability study for self-triage applications was described in section 5.2.
- 4- The evaluation of the results from 34 users against three scenarios and two user interfaces (graphical and list) was performed, and conclusions for self-triage applications were drawn and presented in Chapter 5.

The main features of the Symptoms Pal self-triage mobile application are as follows:

- This application provides many features to improve the overall user experience and to minimize the risks of misleading users when doing self-triage. The application allows users to work with the application without acquiring any personal data. It allows offline accessibility. It also, provides self-treatment suggestions only for clear non-acute cases in order to minimize the risk of misleading users. The application supports providing information in different languages and on multiple devices. Additional reliable information is also available for users within the application and via links to informative medical websites. The preferred start interface for users was also understood in the usability study, and a graphical interface is provided for users to easily select their main symptom in the application.
- To reduce the number of inappropriate visits and to expedite the warranted visits to emergency departments, this application provides reliable information about multiple symptoms validated by doctors from Montfort Hospital. The application also suggests the most suitable and closest healthcare facilities that could take care of a user's conditions based on the triage advice.

The technical comparison of this solution with similar applications included in this study showed that our proposed prototype covers more features than other solutions. The three research questions of the usability study presented in the introduction are answered in the following way:

- 1- *How effective is the self-triage mobile application in helping users select appropriate symptoms and make a decision?* As user responses showed, the graphical interface was favoured over the list interface for the selection of symptoms within the application. The names of the symptoms were generally understandable to users, with a few exceptions. The majority of the participants of this study (21 out of 34) perceived this application as useful for a variety of health conditions. Twelve users stated that they would use the application for non-urgent or less urgent situations, seven mentioned they would use the application for any unknown medical condition, and six replied that they would use the application if they thought that their medical problem was urgent. Hence, the application should be able to provide medical suggestions for different types of health conditions effectively.

- 2- *What usability aspects of the self-triage mobile application require improvement?* The results of the questionnaire, observational findings, and user comments show the presentation of symptoms, questions and answers needs to be re-evaluated and improved. Also, the functionality of the progress bar needs to be reconsidered. The level of cognitive load of this application is not considered minimal yet, which directly affects the usability of this application and needs to be taken into consideration by applying changes on the presentation of the medical content and resolving the design issues.

- 3- *What additional functional features would users like to see in a self-triage mobile application?* Many suggestions were provided by users that need to be taken into consideration. One feature that was also determined as a requirement (see R24 in Table 10) and was suggested by many users was to provide the descriptions for the

necessary medical terms mentioned in the application. Also, as results of the questionnaire showed (90% satisfaction score) the healthcare facility suggestion received the best score from users among the features provided by this application, which shows that this feature is considered as useful and necessary by users.

7.2 Future Work

A number of research opportunities are identified based on the information presented in previous chapters and considering the validity threats identified in section 6.2. To address some of these limitations, there are future opportunities we foresee:

- 1- Applying the remaining modifications highlighted in section 5.9. Many changes should be applied on the medical content presentation, in terms of the number of questions, the wording of the questions, and the presentation of answers to the questions.
- 2- Increasing the number of symptoms included in the application to increase the credibility of the application.
- 3- Performing the clinically-oriented functionality study (as discussed at the end of Section 5.2) with various real patients (i.e., users who already went to the ED and users who are experiencing symptoms at home, with ages and levels of education more aligned with the target population) to validate the accuracy of the decision algorithms, medical suggestions, and triage advice.
- 4- Developing the application for the iOS platforms.

References

- [1] Abdullah, N., Annamalai, M., Rani, A., Amry, M.K.: Paging vs Scrolling: Navigation style for self-triage of epidemic diseases. *Journal of Theoretical & Applied Information Technology*, 88(2), pp. 262-272 (2016)
- [2] Ada™. Retrieved on April 27, 2018 from <https://ada.com>
- [3] Affleck, A., Parks, P., Drummond, A., Rowe B.H., Ovens, H.J.: Emergency department overcrowding and access block. *Canadian Journal of Emergency Medicine*, 15(6), pp. 359-370 (2013)
- [4] Afilalo, J., Marinovich, A., Afilalo, M., Colacone, A., Leger, R., Unger, B.: Giguere, C. Nonurgent emergency department patient characteristics and barriers to primary care. *Academic Emergency Medicine*, 11(12), pp. 1302-1310 (2004)
- [5] American Academy of Family Physicians. Facts about: family practice. American Academy of Family Physicians, (1996)
- [6] Boyer, C., Baujard, V., Scherrer, J.R.: HONcode: a standard to improve the quality of medical/health information on the internet and HON's 5th survey on the use of internet for medical and health purposes. In 6th Internet World Congress for Biomedical Sciences (INABIS 2000), (1999)
- [7] Berner, E.S.: Clinical decision support systems. Springer Science + Business Media, LLC (2007)
- [8] Bhavnani, S.P, Jagat, N., Partho P.S.: Mobile technology and the digitization of healthcare. *European heart Journal*, 37(18), pp. 1428-1438 (2016)
- [9] Bisson, L.J., Komm, J.T., Bernas, G.A., Fineberg, M.S., Marzo, J.M., Rauh, M.A., Smolinski, R.J., Wind, W.M.: How accurate are patients at diagnosing the cause of their Knee Pain with the help of a web-based Symptom Checker? *Orthopaedic Journal of Sports Medicine*, 4(2), doi:10.1177/2325967115S00150 (2016)
- [10] Brahler, S.: Analysis of the android architecture. Karlsruhe institute for technology. (2010)
- [11] buoy, Retrieved on April 27, 2018 from <https://buoyhealth.com/symptom-checker/>
- [12] Carret, M.L., Fassa, A.C., Domingues, M.R.: Inappropriate use of emergency services: a systematic review of prevalence and associated factors. *Cadernos de Saude Publica*, 25(1), pp. 7-28 (2009)

- [13] Derlet, R.W.: Overcrowding in emergency departments: effects on patients. *The California Journal of Emergency Medicine*, 1(1), pp. 2-3 (2000)
- [14] Doctor Diagnose, Retrieved on April 1, 2018 from <https://play.google.com/store/apps/details?id=com.appcolliders.doctordiagnose&hl=en>
- [15] Drugs.com Symptom Checker, Retrieved on April 27, 2018, from <https://drugs.com/symptom-checker/>
- [16] Easy symptom checker, Retrieved on June 14, 2017 <https://apple.com/ca/easy-symptoms?src=globalnav>
- [17] FamilyDoctor Symptom Checker, Retrieved on April 27, 2018, from <https://familydoctor.org/your-health-resources/health-tools/symptom-checker/>
- [18] Farmer, S.E.J., Bernardotto, M., Singh, V.: How good is Internet self-diagnosis of ENT symptoms using Boots WebMD symptom checker? *Clinical Otolaryngology*, 36(5), pp. 517-518 (2011)
- [19] Fox, S., Duggan, M.: Health online 2013. *Health*, pp. 1-55 (2013)
- [20] Gann, B.: Giving patients choice and control: health informatics on the patient journey. *Yearbook of medical informatics*, 7(1), pp. 70-73 (2012)
- [21] Hageman, M.G., Anderson, J., Blok, R., Bossen, J.K., Ring, D.: Internet self-diagnosis in hand surgery. *Hand*, 10(3), pp. 565-569 (2015)
- [22] Hammond, T., Clatworthy, J., Horne, R.: Patients' use of GPs and community pharmacists in minor illness: a cross-sectional questionnaire-based study. *Family practice*, 21(2), pp.146-149 (2004)
- [23] healthdirect™, Retrieved on April 27, 2018 from <https://healthdirect.gov.au/symptom-checker/tool>
- [24] Healthwise®, Retrieved on April 27, 2018 from <https://healthwise.org>
- [25] Healthwise®, Retrieved on April 27, 2018 from <https://myhealth.aberta.ca>
- [26] Healthwise®, Retrieved on April 27, 2018 from <https://www.peacehealth.org/medical-topics/id/hwSxChk>
- [27] HealthNAV™, Retrieved on April 27, 2018 from <https://baycare.org/healthnav>
- [28] Hevner, A.R., March, A.T., Park, J., Ram, S.: Design science in information systems research. *MIS Quarterly*, Vol. 28, No 1, pp. 75-105 (2004)
- [29] Isabel® Healthcare, Retrieved on April 27, 2018 from <https://isabelhealthcare.com>
- [30] Isabel®, Retrieved on April 27, 2018 from <https://netdoctor.co.uk>
- [31] iTriage, Retrieved on December 31, 2017 from <https://itriagehealth.com>

- [32] Jääskeläinen, R.: Think-aloud protocol. Handbook of translation studies. (2010)
- [33] Johnson, G.I., Clegg, C.W., Ravden, S.J.: Towards practical method of user interface evaluation. Applied Ergonomics, Vol. 20, Issue 4, pp. 255-260 (1999)
- [34] Kitchenham, B., Pretorius, R., Budgan, D., Brerton, O.P., Turner, M., Niazi, M., Linkman, S.: Systematic Literature reviews in software engineering - a tertiary study. Information and Software Technology, 52(5), pp. 792-805 (2010)
- [35] Kitchenham, B.: Procedures for performing systematic reviews. Technical Report, Keele University and NICTA. (2004)
- [36] Kushniruk, A.W., Patel, V.L.: Cognitive and usability engineering methods for the evaluation of clinical information systems. Journal of Biomedical Informatics, 37(1), pp. 56-76 (2004)
- [37] Lanna, M., Amyot, D.: Spotting the Difference. Software: Practice & Experience, Vol. 41, Issue 6, pp. 607-626 (2011)
- [38] Li, H., Salah, H.: Comparing four online symptom checking tools: Preliminary Results. In IIE Annual Conference, pp. 1979-1984. Institute of Industrial and Systems Engineers (IISE), (2017)
- [39] Li, W.C.: Supporting Pathology Process Management with Real-time Business Intelligence. Master's thesis, Computer Science, University of Ottawa (2016)
- [40] Luger, T.M., Houston, T.K., Suls, J.: Older adult experience of online diagnosis: results from a scenario-based think-aloud protocol. Journal of medical Internet research, 16(1), (2014)
- [41] Lupton, D., Jutel, A.: It's like having a physician in your pocket! A critical analysis of self-diagnosis smartphone apps. Social Science & Medicine, pp.128-135 (2015)
- [42] Maalej, W., Nayebi, M., Johann, T., Ruhe, G.: Toward Data-Driven Requirements Engineering. IEEE Software, 33(1), pp. 48-54 (2016)
- [43] Marco-Ruiz, L., Bønes, E., de la Asuncion, E., Gabarron, E., Aviles-Solis, J.C., Lee, E., Traver, V., Sato, K., Bellika, J.G.: Combining multivariate statistics and the think-aloud protocol to assess Human-Computer Interaction barriers in symptom checkers. Journal of biomedical informatics, 1(74), pp. 104-122 (2017)
- [44] Moreno, E.B., Pueyo, I.F., Sanchez, M.S., Martin M.B., Masip, J.U.: A new artificial intelligence tool assessing symptoms in patients seeking emergency department care: the Mediktor applications. Emergencias: revista de la Sociedad Espanola de Medicina de Emergencias, 29(6), pp. 391-396 (2017)
- [45] Morita, T., Rahman, A., Hasegawa, T., Ozaki, A., Tanimoto, T.: The potential possibility of symptom checker. International journal of health policy and management, 6(10), pp. 615-616 (2017)

- [46] Nayak, B.K.: Why learn research methodology. *Indian journal of ophthalmology*, 57(3), pp.173-174 (2009)
- [47] NHS choices, Retrieved on May 20, 2018 from <https://www.nhs.uk/>
- [48] NHS Direct, Retrieved on April 27, 2018 from <https://nhsdirect.wales.nhs.uk>
- [49] Nielsen, J.: 10 usability heuristics for user interface design. Nielsen Norman Group, (1995)
- [50] North, F., Ward, W.J., Varkey, P., Tulledge-Scheitel, S.M.: Should you search the internet for information about your acute symptom? *Telemedicine and e-Health*, 18(3), pp. 213-8 (2012)
- [51] Patient Symptom Checker, Retrieved on April 27, 2018 from <https://patient.info/symptom-checker>
- [52] Perry, D.E., Porter, A.A., Votta, L.G., Empirical studies of software engineering: a roadmap. In: *Conf. on The Future of Software Engineering, ICSE'2000*, ACM, pp.345-355 (2000)
- [53] Powley, L., McIlroy, G., Simons, G., Raza, K.: Are online symptoms checkers useful for patients with inflammatory arthritis? *BMC musculoskeletal disorders*, 17(1), p. 362 (2016)
- [54] Robertson, N., Polonsky, M., McQuilken, L.: Are my symptoms serious Dr. Google? A resource- based typology of value co-destruction in online self-diagnosis. *Australasian Marketing Journal (AMJ)*, 22(3), pp. 246-256 (2014)
- [55] Rotmensch, M., Halpern, Y., Tlimat, A., Horng, S., Sontag, D.: Learning a health knowledge graph from electronic medical records. *Scientific reports*, 7(1), p. 5994 (2017)
- [56] Sanaeifar, A., Faraahi, A., Tara, M.: SEPHYRES 1: A Symptom Checker based on Semantic Pain Descriptors and Weight Spreading. *Applied Medical Informatics*, 38(3/4), p. 105 (2016)
- [57] Saunders, M.N.: *Research methods for business students*. Pearson Education India, Fifth Edition (2011)
- [58] Semigran, H.L., Levine, D.M., Nundy, S., Mehrotra., A.: Comparison of physician and computer diagnostic accuracy. *JAMA internal medicine*, 176(12), pp. 1860-1861 (2016)
- [59] Semigran, H.L, Linder, J.A, Gidengil, C., Mehrotra, A.: Evaluation of symptom checkers for self diagnosis and triage: audit study. *BMJ*, doi:10.1136/bmj.h3489 (2015)

- [60] Silva, B.M., Rodrigues, J.J, de la Torre, Díez, I., López-Coronado, M., Saleem, K.: Mobile-health: A review of current state in 2015. *Journal of biomedical informatics*, 1(56), pp.265-72. (2015)
- [61] Speicher, M., Both, A., Gaedke, M.: Ensuring web interface quality through usability-based split testing. In *International Conference on Web Engineering*, pp. 93-110. Springer, Cham (2014)
- [62] Stone, D., Jarrett, C., Woodroffe, M., Minocha, S.: *User Interface Design and Evaluation*. Elsevier (2005)
- [63] Symcat, Retrieved on April 27, 2018 from <https://symcat.com>
- [64] Symptify[®], Retrieved on April 27, 2018 from <https://symptify.com>
- [65] Symptomate, Retrieved on April 27, 2018 from <https://symptomate.com>
- [66] Symptoms Checker Android Application, Retrieved on July 12, 2017 from <https://play.google.com/store/apps/details?id=com.mediahape.symptomschecker>
- [67] The Human Diagnosis Project, Retrieved on April 27, 2018 from <https://human-dx.org>
- [68] Van den Heede, K., Van de Voorde, C.: Interventions to reduce emergency department utilisation: A review of reviews. *Health Policy*, 120(12), pp.1337-49 (2016)
- [69] Wang, W., Chan, S., He, H.G.: Developing and testing a mobile application programme to support self-management in patients with stable angina: a feasibility study protocol. In *Nursing Informatics*, pp. 241-248 (2014)
- [70] WebMD Symptom Checker, Retrieved on April 27, 2018 <https://symptoms.webmd.com/#/info>
- [71] Your.MD Symptom checker, Retrieved on April 27, 2018 <https://your.md>
- [72] Zagorecki, A., Orzechowski, P., Hołownia, K.: Online diagnostic system based on Bayesian networks. In *Conference on Artificial Intelligence in Medicine in Europe*, Springer, Berlin, Heidelberg, pp. 145-149 (2013)

Appendix A: Email to Participants

Dear University of Ottawa Student/Faculty/Employee,

As part of my Master's thesis research in Systems Science under the supervision of Prof. Daniel Amyot and Prof. Wojtek Michalowski, I (Mir Kamyar Ziabari) am conducting a study that aims to evaluate certain features of a Symptom Checker Mobile (SCM) application developed to help people having health issues to assess their condition using reliable and accessible information.

I am looking for a feedback on this SCM application from graduate/undergraduate students and employees at the University of Ottawa. This feedback will be used to assess a proof of concept design and the usability of the application.

Symptom checkers are systems that allow a lay person to use his/her health data to assess presence of potential health concerns. Using diagnostic algorithms developed on the basis of medical evidence (clinical guidelines or pathways), symptom checkers require users to answer a series of short questions about their symptoms or to input some data about their health status. These applications usually provide users with two types of feedback. The first type is to help user to conduct a self-diagnosis, and the second one is to assist them with self-triage regarding what to do next. The self-triage option aims to inform whether user should seek immediate medical attention such as visit to an Emergency Department, or should visit a family doctor (walk-in clinic), or do self-care. This SCM application is concerned solely with self-triage.

The 45-minute session I am asking you to participate in, will include a demo of the SCM, a list of tasks to be completed, and a short system usability questionnaire. Your answers will help to evaluate the SCM usability and identify further potential improvements. The session will take place on the University of Ottawa campus (DMS or SITE), at a time convenient for you, including evenings and weekends.

If you agree to participate, or have any questions or comments about the study, please contact **me at the email below or, alternatively, one of my supervisors, Professor Daniel Amyot or Professor Wojtek Michalowski.**

Thank you for your consideration.

Best regards,

Mir Kamyar Ziabari
Master of Systems Science
School of Electrical Engineering and
Computer Science, University of Ottawa
xxxx@xxxx.xxx

Daniel Amyot, Ph.D., P.Eng.
Full Professor
School of Electrical Engineering and
Computer Science, University of Ottawa
xxxx@xxxx.xxx

Wojtek Michalowski, Ph.D.
Full Professor
Telfer School of Management
University of Ottawa
xxxx@xxxxx.xxx

Appendix B: Consent Form

Title of the study: Usability study of a mobile self-triage application.

Name of researcher:

Mir Kamyar Ziabari, graduate student, School of Electrical Engineering and Computer Science,
cell: xxx-xxx-xxxx, e-mail: xxxxx@xxxxx.xxx

Supervisors:

Prof. Daniel Amyot, School of Electrical Engineering and Computer Science, phone xxx-xxx-xxxx, e-mail: xxxx@xxxxx.xxx;

Prof. Wojtek Michalowski, Telfer School of Management, phone xxxx-xxx-xxxx ext. xxxx, e-mail: xxxx@xxxxx.xxx

Invitation to Participate: I am invited to participate in the research on mobile symptom checker application conducted by **Master's** student Mir Kamyar Ziabari from the School of Electrical Engineering and Computer Science. Mir Kamyar Ziabari is under the supervision of Prof. Daniel Amyot from the School of Electrical Engineering and Computer Science and Prof. Wojtek Michalowski from the Telfer School of Management.

Purpose of the Study: The goal of this study is to evaluate user experience with a Symptom checker mobile application. Symptom checkers are systems that allow a lay person to use his/her health data to assess the presence of potential health concerns. Using diagnostic algorithms developed on the basis of medical evidence (clinical guidelines or pathways), symptom checkers require users to answer a series of short questions about their symptoms or to input some data about their health status. These applications usually provide users with two types of feedback. The first type is to help user to conduct a self-diagnosis, and the second one is to assist them with self-triage

regarding what to do next. The self-triage option aims to inform whether user should seek immediate medical attention such as visit to an Emergency Department, or should visit a family doctor (walk-in clinic), or do self-care. This SCM application is concerned solely with self-triage.

Participation: My participation will consist of a single 45-minutes session during which the researcher will first explain how the application is to be used in the context of the experiment. I will then be asked to use the application to answer the pre-defined set of questions, and finally to answer an anonymous usability questionnaire with multiple-choice and short-answer questions about my impressions regarding the application.

Risks: Beyond the time needed to participate in this experiment, there are no known risks associated with my participation in this experiment greater than those I might encounter in everyday life. I have received assurance from the researcher that every effort will be made to minimize these risks.

Benefits: My participation in this study will provide the researcher with experimental data to evaluate and propose improvements to the Symptom Checker Mobile Application.

Confidentiality and anonymity: I have received assurance from the researcher that the information I will share will remain strictly confidential by communicating the results of the experiment in an aggregated manner (e.g., “half of the participants found that the tool was easy to learn”). I understand that the answers will be used only to evaluate the symptom checker mobile application only.

Anonymity: My anonymity will be protected by not recording my name or any identifiable information. If needed, data may be tagged with non-traceable numeric Identifiers.

Conservation of data: The data collected in the hard-copy questionnaire will be kept in a secure manner that will be accessed only by the researchers. The data will be kept by one of the supervisors for a period of 5 years in case of an audit.

Voluntary Participation: I am under no obligation to participate and if I choose to participate, I can withdraw from the study at any time and/or refuse to answer any questions, without suffering any negative consequences. If I choose to withdraw, all data gathered until the time of withdrawal will be removed and paper questionnaires will be destroyed.

Acceptance: I, (Participant's Name) _____,
having read and understood the above text, and having had the possibility to ask and receive complementary information on the study, agree to participate in this study.

If I have any questions about the study, I may contact the researcher or his supervisors.

If I have any questions regarding the ethical conduct of this study, I may contact the Protocol Officer for Ethics in Research, University of Ottawa, Tabaret Hall, 550 Cumberland Street, Room 154, Ottawa, ON, K1N 6N5. Tel.: (613) 562-5387. Email: ethics@uottawa.ca

There are two copies of the consent form, one of which is mine to keep.

Participant's signature:

Date:

Researcher's signature:

Date:

Appendix C: Pre-Test Interview

- 1- Are you a) Student, or b) Employee?
- 2- Age group: [18, 25], [26, 35], [36,45], [46, 55], [55,65], [Over 65]
- 3- Do you own a smartphone/mobile device? Yes No
- 4- Have you ever used a Symptom Checker Mobile Application? Yes No
- 5- Have you ever been to an Emergency Room? Yes No

Appendix D: Study Protocol

Reminder on the use of this protocol: The activities in this protocol are to be followed in the order described below. The time estimated for each activity may change from one participant to the other; this should be allowed and participants should not be rushed to finish any activity.

1. Consent form (5 minutes). The participant is presented with the consent form and given the opportunity to ask any question or clarification before consenting.
 - 1.1. If the participant declines to sign the consent form, thank him/her.
 - 1.2. If the participant signs the consent form, continue the experiment.
2. Pre-Test Interview (5 Minutes). Before introducing the application, participant is given a “Pre-Test Questionnaire” document and will be asked to answer questions provided in the document. The participant can refuse to answer any question with which s/he is uncomfortable.
 - 2.1. The researcher should answer any question that the participant may have about the questionnaire as a whole or about individual questions.
3. Introduction (5 minutes). Introduce symptoms checker applications to the participant, and show the basic functionalities of the application. No preparation is necessary for this task.
4. Tasks (10 minutes). The participant is given the “Participant tasks” document (see Appendix E – Participant tasks), where he/she is asked to perform a sequence of four tasks and think aloud throughout the process.
 - 4.1. The researcher should verbally help the participant if asked to do so, or if he observes that the participant is having difficulties completing a task.

- 4.2. The researcher should make note whether help was required to complete the task, as well as any emerging issues that the participant may have while using the application.
5. Questionnaire (15 minutes). The participant is given the “Questionnaire” document (Appendix F) and asked to answer the questions. Remind the participant that s/he can refuse to answer any question with which s/he is uncomfortable.
 - 5.1. The researcher should answer any query that the participant may have about the questionnaire or its individual questions.
6. Conclusion (1 minute). The researcher collects the “Participant tasks” and “Questionnaire” documents from the participant, thanks the participant for his/her time.

Appendix E: Task List

Participant # _____

Task 1- A list of scenarios is provided to you, read the first scenario and try to fully understand it. Researcher will ask you to select the symptom in the version of the application that is provided to you.

Task 2- Answer the questions. You should answer follow up questions provided by application based on the selected scenario.

Task 3- Results will be generated by the application, based on the suggested care option(s). Choose the best healthcare setting recommend by the application.

Task 4- In this task the version of the application would be modified by researcher. You should read the second scenario and follow the procedure this time, in the second version. It is optional to choose the healthcare setting again.

Task 5- Researcher will describe a scenario to you, he asks you to listen carefully, and follow, what he is describing. You can select your preferred version of the application, and do all the steps one more time, based on researcher description of symptom.

Task 6- You can select any of the symptoms, for example symptoms that you recently had, or a family member problem, using your desired version of the application and follow the process.

Note: Remember to think out loud through the process and state your comments if there are any.

Appendix F: Questionnaire

Participant # _____

All questions below should be answered based on the features that were discussed and/or showed during the experiment. Please, do **NOT** base your answers on your previous knowledge or expected features of the application.

For each of the questions below, *circle* the answer that best matches your opinion.

Question 1

Learnability is a measure of how easy it is to learn to use a software product. Given this definition, the Symptom Checker Application is easy to learn.

Strongly Agree Agree Neutral Disagree Strongly Disagree

Question 2

Ease of use is a measure of how easy it is to use a software product after its use has been *learned*. Given this definition, the Symptom Checker Application is easy to use.

Strongly Agree Agree Neutral Disagree Strongly Disagree

Question 3

Efficiency is a measure of how quickly tasks can be performed with a software product after its use has been mastered. Given this definition, the Symptom Checker Application allows users to perform tasks efficiently.

Strongly Agree Agree Neutral Disagree Strongly Disagree

Question 4

Intuitiveness is a measure of how easy it is to understand the output or the interface of a software product. Given this definition, the Symptom Checker Application is intuitive to use.

Strongly Agree Agree Neutral Disagree Strongly Disagree

Question 5

Cognitive Load refers to the total amount of mental effort being used in the working memory in order to understand a context. Minimized cognitive load, prevent users from becoming overwhelmed while using the application. Given this definition, cognitive load was minimal.

Strongly Agree Agree Neutral Disagree Strongly Disagree

Question 6

It is easy to locate the desired symptom in the Symptom Checker Application from *Symptoms List*.

Strongly Agree Agree Neutral Disagree Strongly Disagree

Question 7

It is easy to locate the desired symptom in the Symptom Checker Application from *Graphical Interface version*.

Strongly Agree Agree Neutral Disagree Strongly Disagree

Question 8

The reason you were asked to find a symptom provided by researcher was to simulate a situation where you need to check symptoms for someone else. Given this explanation, it is easy to locate the required symptom in the Symptom Checker Application from *Symptoms List*.

Strongly Agree Agree Neutral Disagree Strongly Disagree

Question 9

The reason you were asked to find a symptom provided by researcher was to simulate a situation where you need to check symptoms for someone else. Given this explanation, it is easy to locate the required symptom in the Symptom Checker Application from *Graphical Interface version*.

Strongly Agree Agree Neutral Disagree Strongly Disagree

Question 10

Names of the symptoms that were provided in the Symptom Checker Application (e.g., headache, fever) are clear and transparent and could easily be interpreted and understood.

Strongly Agree Agree Neutral Disagree Strongly Disagree

Question 11

After selecting the desired symptom, a series of questions are asked so the Symptom Checker Application can assess health concern. The presentation of questions is appealing (font, size, colors, and overall look).

Strongly Agree Agree Neutral Disagree Strongly Disagree

Question 12

There is a progress bar located above the questions in order to help the user to know how far s/he is in the process of answering the questions. Having the progress bar is helpful.

Strongly Agree Agree Neutral Disagree Strongly Disagree

Question 13

After answering all the questions, the medical recommendation is provided. The presentation of recommendations is clear.

Strongly Agree Agree Neutral Disagree Strongly Disagree

Question 14

After looking at the care suggestions, you can use “Take me there” Button to access the list of healthcare facilities. You believe that this feature is useful.

Strongly Agree Agree Neutral Disagree Strongly Disagree

Question 15

Is there a feature that you think is not useful? If Yes, describe why?

Question 16

Which types of health concern decisions, if any, do you think that the Symptom Checker Application could help you make?

Question 17

Would you use the Symptom Checker Application in its current form when having some health concern? If Yes, based on what features?

Yes

No

Question 18

What feature would you suggest to add to the Symptom Checker Application

Question 19

Please suggest other potential improvements to the Symptom Checker Application.

Question 20

Choose your level of preference for the type of view (Graphical Interface or Symptoms List)?

Definitely Graphical Somewhat Graphical Both Somewhat List Definitely List

THANK YOU FOR YOUR PARTICIPATION!

Appendix G: User Scenarios

Scenario #1

Peter, 20 years old male has not felt well for past couple of days. He experienced stomach pains, nausea, and general discomfort. This Saturday morning he woke up with a feeling of fullness and started to vomit. He noticed a strange dark color both in his stool and in the vomit.

Scenario #2

Jeanette, 28 years old doctoral student is concerned because she has had flu-like symptoms for the past couple days and she really cannot concentrate to study for her comprehensive exam. She has fever, headache, and a bit of chills. She stays in bed, takes some aspirin and hopes to get better. Jeanette woke up this morning feeling miserable – her fever is really high and she is getting throbbing headache. She did not sleep well last night so her neck feels a bit stiff. Jeanette is not sure what to do and if she should go to the pharmacy to buy over the counter cold medication.

Scenario #3

This scenario will be described to the user by the researcher.

(Hidden from participant)

The Scenario

Mary, 20 years old student athlete likes spicy food. With couple of her friends she went to Thai restaurant close to the campus and since that visit (3 days ago) she does not feel well. It is not stomach that bothers her but her recurring pain in the chest area is worrisome. The pain feels like a sharp burning pain of short duration in chest area close to heart that worsen after a meal. She also feels a bit bloated, and experiences burning while drinking hot drinks.

Expected outcome for scenario #1: The symptoms are heavily bleeding ulcer and Peter should go to the ED.

Expected outcome for scenario #2: While initial symptoms were consistent with common cold, jump in temperature combined with perceived stiffness in neck point towards possible meningitis that asks for the ED visit.

Expected outcome for scenario #3: The symptoms are consistent with acid reflux and considering that Mary is young and athletic it is not a heart condition; Mary should be visiting her family physician for further evaluation for suspected reflux.