WEB 2.0 BASED APPLICATIONS FOR ASTHMA/COPD PATIENTS; SOCIAL NETWORKING PLATFORM FOR SELF-MANAGEMENT VIA KNOWLEDGE EXCHANGE

by

Athra A. Alaujan

Submitted in partial fulfilment of the requirements
for the degree of Master of Health Informatics

at

Dalhousie University
Halifax, Nova Scotia
August 2012

DALHOUSIE UNIVERSITY

FACULTIES OF COMPUTER SCIENCE AND MEDICINE

The undersigned hereby certify that they have read and recommend to the Faculty of Graduate Studies for acceptance a thesis entitled "WEB 2.0 BASED APPLICATIONS FOR ASTHMA/COPD PATIENTS; SOCIAL NETWORKING PLATFORM FOR SELF-MANAGEMENT VIA KNOWLEDGE EXCHANGE" by Athra A. Alaujan in partial fulfilment of the requirements for the degree of Master of Health Informatics.

Dated: August 17, 2012

Supervisor: _______________________________

Readers: _______________________________

____

DALHOUSIE UNIVERSITY

Date: August 17, 2012

AUTHOR: Athra A. Alaujan

TITLE: WEB 2.0 BASED APPLICATIONS FOR ASTHMA/COPD PATIENTS; SOCIAL NETWORKING PLATFORM FOR SELF-MANAGEMENT VIA KNOWLEDGE EXCHANGE

DEPARTMENT OR SCHOOL: Faculties of Computer Science and Medicine

DEGREE MHI CONVOCATION: October YEAR: 2012

____

Signature of Author

DEDICATION

To the marvellous family I belong to; my parents, sisters, and brothers without your continuous love, support, and encouragement I could not stand where I am today. Also, I would like to dedicate this work to you King Abdullah of Saudi Arabia. I was lucky to be one of the students in your scholarship program and very honourable to represent my beloved country.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The world-wide-web has become a popular medium for health-related information exchange. New tools in Web 2.0 allow dynamic, disease-specific discussions. The impact on self-care and healthcare-related decisions merit academic focus and recognition. Our goal of this thesis is to assess the extent to which online social media are useful as information diffusion and knowledge exchange media, as guided by Asthma and COPD Action Plan recommendations. In this work, text data collected from online discussion forums were mined using syntax classification components of Symptoms and Actions. Coded text data were mapped to SNOMED-CT, UMLS, and Consumer Health Vocabulary to match clinical data terminology standards specified by the action plan. Those mapped data were converted into meta-matrix data for online group social networks and statistical analysis. The data analyses supported the hypothesis that unguided social media information exchange is congruent with two out of eight symptoms and actions. Results lead to the conclusion that health professional mediation is recommended

LIST OF ABBREVIATIONS USED

COPD	Chronic Obstructive pulmonary disease
AP	Action Plan
CHI	Consumer Health Informatics
DiG	Distributed Guidelines
CHV	Consumer Health Vocabularies
UMLS	Unified Medical Language System
SNOMED CT	Systematised Nomenclature of Medicine Clinical Concepts
PHR	Personal Health Record
ATR	Automated Term Recognition
CAU	Computer Assisted Update
ICTA	Internet Community Text Analyzer
SNA	Social Network Analysis

ACKNOWLEDGEMENTS

It was a valuable opportunity for me to do this thesis with you Dr. Swarna Weerasinghe. Thank you for your time, support, and thoughtful feedback. It was a very rewarding experience. To the two members of my advising committee, Dr. Christian Blouin and Dr.David Zitner, I would like to express my sincere appreciation for consenting to read my work. Thank you.

The contribution of Dr. Grace Paterson at the early developing stages of this thesis was valuable and highly appreciated. Thank you for your professional directions. Also, my thanks to respirology specialists at the Queen Elizabeth Health Sciences Centre, Dr.Paul Hernandez and Dr.Dennis Bowie. I highly appreciate your time and suggestions.

I would like to acknowledge the recommendations and suggestions of Dr. Anatoliy Gruzd as well: thank you for your clarifications and directions about the social media lab tools.

To the Dalhousie Computer Science professional community: thank you for your enthusiasm and sharing your expertise.

CHAPTER 1 INTRODUCTION

1.1 GENERAL OVERVIEW

The World Wide Web (Web 1.0) is becoming an increasingly common source of health information for patients and healthcare providers. One revolutionary aspect of this emerging technology is how healthcare information exchange helps patients take charge of their own health. Furthermore, with the emergence of Web 2.0 and the proliferation of its applications (e.g., online social groups), knowledge transfer and exchange are becoming virtually effortless, with considerable benefits accruing to patients, health care professionals, and even to biomedical researchers.

In addition to being a heavy burden on healthcare systems, respiratory medical conditions can be life-long afflictions that require adjustments at emotional, physical, professional, and social levels [1]. Canada ranks high in the list of the world's countries with a marked prevalence of Asthma and Chronic Obstructive Pulmonary Disease (COPD) [1]. Symptom control is one of the challenges in managing such cases. Along with emphasizing individualized care plans, Web 2.0 applications such as wikis, blogs, and podcasts are forms of social networking, peer learning, and information exchange tools that support patients' disease management [2]. Unlike Web 1.0 (the original Internet) and its application (the World Wide Web, which supports "read only" for online information), Web 2.0 provides more interactive functions for users and thus enables their input [2].

This study involved a patient-directed initiative of personal health information sharing and exchange among Asthma and COPD patients in two different websites. . Asthma/COPD patients interacted in these social networking media and discussion boards, asking questions and sharing opinions. The professional guideline document known as the Action Plan (AP) was used to explore the extent to which the pertinent health information exchanges are congruent with clinical recommendations. Health informatics and statistical methods are also be used in this research.

1.2 Background and Purpose

Disease conditions affecting the Respiratory System, such as Asthma and COPD, are a considerable burden to individual patients, communities, and healthcare governance sectors. Asthma is a chronic inflammatory respiratory condition characterised by airway hyper-responsiveness, which often leads to chest tightness and breathlessness. Treatment objectives of Asthma include controlling its symptoms and maintaining the individual's normal activities [3]. Because the severity and frequency of symptoms vary among patients, individualised disease management plans such as the AP could prove effective [4].

COPD is also of an area of interest in our research. This disease condition is manifested by progressive airflow reduction. Both COPD and Asthma have a significant impact on individuals' productivity and quality of life [1, 5].

Literature suggests that individualised APs have proven to empower patients' self-management and are seen to have better health outcomes when used appropriately [4]. As effective symptom control in COPD is considered a primary treatment objective [5], following an AP recommended by a General Practitioner has been shown to slow the worsening of diseases with a progressive nature [5]. There are several patient-empowering tools for self-management. Our primary goal in this research is to explore Web 2.0 social media applications for health-related information exchange.

This thesis research involves an exploration of social networking and peer interaction as a form or a medium of patients' knowledge gain. The goal was to evaluate this social interaction against clinical recommendations. Through these social media, patients could transfer and share their disease experience and knowledge with peers. In particular, I tested whether patients with Asthma/COPD gained knowledge on symptom control and effective self-management of these conditions, as specified in the AP (see Appendixes A and B).

For Asthma/COPD, an AP comprises written instructions within a specific context to guide patients when their symptoms become exacerbated or their condition deteriorates. The use of an AP supports and encourages self-management and individualized patient care [4, 5]. Well-explained APs for Asthma patients have been found to improve their disease conditions' and health outcomes [4]. It is noted in the literature that adhering to an AP for patients with COPD had positive effects on self-care and control of exacerbations [5].

Furthermore, patient education has been studied for Asthmatic and COPD patients as a source of empowerment, independence from health care providers, and enhancement of patients' quality of life [6]. Since Asthma is a chronic condition that affects individuals' daily lives, the psychological aspect would also have an influence on their clinical wellness measures and outcomes. Asthmatic patients who followed an AP experienced physical and psychological health benefits due to their ability to communicate and express their health-related issues [7].

At present, the implementation of an AP appears problematic. The process involves the following: upon hospital discharge, respiratory specialists instruct patients to take their AP to their family doctors for implementation. However, this step is not being fully implemented. In response, the respiratory specialists at Queen Elizabeth II Health Science Centre indicated an interest in testing other means of implementation. This thesis research follows one such option – the use of social media to provide AP instructions to patients.

As we become better adapted to our technologically advanced world, clinicians feel more comfortable advising patients to use Web 2.0 applications as a source of health-related information [8, 9]. Social networking is a core element of these applications. The use of these tools is advancing and proliferating among the general public; including among patients [2]. Social networking opens potential new professional communication channels that could improve clinical outcomes and help reduce health care costs. Unlike face-to-face communications, online discussions could overcome time and space constraints [8], and the numbers of users is expected to increase significantly in the future [10]. Any member of an online community could be a source of information for others requesting non-professional opinion [11].

Online social media tools and applications have already been used to empower patients with information for individualized self-care [12]. One example is “copdcanada.ca”, where a link to an AP is already part of the resources on this site. Although positive outcomes are yet to be proven, it will be difficult to research and study their direct impact on individual health without adopting further technological advancements.

Social network assessments and analyses have been deployed in several areas to measure and evaluate healthcare practices and interventions [13] or to assist in designing programs to adopt and promote new intervention [14, 15]. One dimension of this thesis work involves APs for Asthma and COPD, where our exploratory network analysis of the online discussion groups was assessed.

Online communities and social groups can be considered a source for future directions in the development and use of social capital. This could govern health promotion for patients with chronic conditions [16]. Considering the ecosystem of health care organizations and public health systems, and with current information systems advancements such as Electronic Health Records (EHR), the integration of social networking tools is recommended [17], either to meet community expectations or to reduce costs.

Social networking is considered a co-application in Medicine 2.0 and Web 2.0 as a form of mutual relationship that enables and encourages communication [18]. By using this technology, people with similar conditions share information and support through effortless and accessible modes of communication. Several online communities are being created, based on their interest in a particular kind of disease. An example is cancer patient groups [19]. Social networking is becoming particularly popular among the younger generations and with those with higher technical skills [20].

Although online disease-specific communities are becoming prevalent, there is a paucity of research that examines their effectiveness. Also, the methods that study these interactions are not yet well defined and can often be challenging [21]. In order to bridge the gap in knowledge, this thesis work includes data processing, as well as analytical frameworks and applications, with reference to Asthma and COPD online community discussions. Existing interactive websites for Asthma and COPD communities were analysed. Two websites for Asthma and COPD community networks were designed for the purpose of this thesis. Our aim was to assess whether patients' discussions are congruent with AP recommendations.

The use of an AP is the patient's responsibility: he/she obtains it from a respiratory clinician and communicates it to a family physician. In this study, an Internet discussion forum was created to understand the barriers and incentives of using an AP. This study also explores several possibilities of utilizing the Internet to increase awareness among Asthma and COPD patients about AP use.

This work is designed as a pilot study. It can be used for future research directions pertaining to analysing online social networking issues related to this category of patients.

1.3 Research Goals and Significance

This thesis is composed of several components: online social networking among Asthma and COPD patients and their care givers; the extent of the usability and usefulness of this medium to share knowledge; and the influence of online social networking on self-management and adherence to the AP. With the known benefits of using an AP for individuals with chronic respiratory conditions [4, 5], Web 2.0 social media applications could be adopted among targeted online communities as a promoting component in the contemporary technology era [11]. Users' engagement on Web 2.0

applications is a novel advantage. Also, online community forums for social interaction, networking and sharing health-related information among friends and family is recognised as the second source of health information after health professionals [11].

Our primary objective was to assess the extent to which online social networking is useful in knowledge exchange among Asthma/COPD patients and care givers. Text communicated to each group (i.e., groups of either Asthma or COPD patients) were our web-based platform to connect participants and have them share their posts. The text extracted from views, comments, questions, and/or any type of posts on an online discussion board were classified with respect to clinical data standards and terminology.

Social network analysis was performed on existing online community discussion forums. This was aimed at exploring the influence of social media on Asthma/COPD knowledge, self-management, and adherence to the AP. The research question was: Is online social media a viable option for knowledge exchange, self-management, and adherence in general and to an Action Plan in particular?

1.4 Terminology

Because this work incorporates clinical and technical perspectives, a few terms require explanation. The *Action Plan* (AP) is a set of recommendations, given by a clinician, for a patient to follow when their condition deteriorates and symptoms exacerbate. It is tailored and individualized based on the patient's disease severity and medications [4]. The AP format used in our work is based on the Canadian Respiratory Guidelines. An individualized AP is provided to patients by healthcare professionals at the Nova Scotia Capital District Health Authority (Appendixes A and B).

Web 2.0 is also called the "Social Web" due to its applications that allow for more user involvement [2]. It is more flexible than the original World Wide Web and permits more collaborative activities among users and their social networks than what was available earlier. *Medicine 2.0* is a reflection of Web 2.0 applications in healthcare and the personal health context. [18].

A *social network* is a mapping process of individuals and concepts that are somehow connected. Social network analysis is the analytical examination of those social connections among members of the network and their interactions with each other, with respect to several quantitative measures [23].

1.5 Structure

Following the introduction section, the next chapter includes: a literature review on Asthma/COPD patients' AP use; online social media (social networks) of health-related topics with respect to knowledge exchange; and social networks analyses. In the methods chapter, there are descriptions of the data collection process used, including: methods of social media site review; data collection, processing and management; social network analyses; and statistical analyses. Results of data processing, information management, and analyses are included in the results chapter, followed by a brief chapter on conclusions.

CHAPTER 2 LITERATURE REVIEW

2.1 Introduction

This chapter starts with a literature review of AP and its benefits for Asthma and COPD patients. The literature is critically reviewed to understand whether Web 2.0 tools, with their increasing popularity and applications among users, can be used as a medium and platform for health-related information exchange. We also investigate whether AP use and awareness for this targeted population could be promoted through social marketing tools in Web 2.0.

In the section entitled Consumer Health Vocabulary, [social media on Web 2.0 as] an approach to enable communication of health-related information in comprehensible format to patients as health consumers is presented and reviewed. In this literature review, particular attention was paid to health informatics and related journals.

2.2 Clinical Domain

Asthma and Chronic Obstructive Pulmonary Disease (COPD) are medical conditions affecting the respiratory system. These conditions are clinically well-defined [1, 3] and their aetiology and symptoms are intimately known to healthcare professionals and patients. However, individuals with Asthma and COPD often have substantially higher mortality and morbidity rates than the general population [1, 3]. In fact, resource utilization linked to care and services for COPD patients is soaring in healthcare facilities in Canada. COPD was found to be one of the most common reasons for hospital admissions [3]. For Asthma, there has been a 50% global increase in diagnosed cases every decade [1].

These conditions (COPD and Asthma) can impact the emotional and physical states of individuals and can result in social and professional restraints. Symptom control is central to treatment plans to mitigate these impacts.

In view of the presence of untreated and/or uncontrolled cases of Asthma, a global initiative suggests several objectives primarily directed toward more patient engagement [1]. Use of an AP, giving a patient individualized strategies through an AP to manage symptoms and condition exacerbation, emphasizes patient engagement through a formalized process [1].

In the literature, an AP is characterized by a set of recommendations and suggestions prescribed for patients to follow in cases where symptoms worsen or deteriorate [4]. Its recommendations can be broadly classified into two components as symptoms and actions. Symptoms include breathing patterns and cough characteristics. Actions constitute medications dosage change (with instructions on when and how to increase them) and some breathing techniques and instructions, aimed at controlling symptoms to decrease exacerbation. Asthmatic patients have been shown to have better health outcomes if they properly use their AP. Similarly, COPD patients' AP use has been linked to the successful management of acute exacerbation and reduced hospitalization rates [5]. In addition, the literature provides evidence for higher instances of exacerbations and use of treatments among AP non-users compared to users: AP users had accelerated recovery periods [24, 25].

An AP was formatted and prepared based on evidence-based clinical guidelines [26]. Results of use of this AP among patients, which were lower than expected, were tracked, and have endured, the recommendation-to-actual clinical practice gap [26]. AP utilization at the point of care was carried out through collaborative efforts of healthcare providers and patients, the gap in detainment of use and adaptation by patients is noted to be extensive and persistent in. Scholars assessed factors associated with this phenomenon for an Asthma AP and looked at patient and health professional perspectives [26]. Communication was identified as one of the main factors that hindered AP use and uptake. Patient knowledge and experience of the disease were communicated differently among themselves than among clinicians. Moreover, from a patient perspective, Asthma terminologies used in the AP were considered ambiguous when individuals described the condition. Consequently, the AP was found to be under-used because it was incompletely understood by patients [26].

It is timely to find out dimensions that would support and enhance patient understanding using technological advancements, since traditional knowledge acquisition patterns are not apt or suitable for current and future generations [27]. One particular means of knowledge acquisition is through social connections. This is highly relevant to the health context and individualized clinical conditions via social media connections. However, it could also be attained, facilitated and enabled through different Web 2.0 tools (e.g., online social communities), where knowledge and experiences are interchangeably shared. A literature review of this type of information gain, and its link to knowledge and patient understanding, will be presented in the following section.

2.3 Literature Review on Information Sharing and Exchange via Online Social Communities and Web 2.0 Tools

In our web-savvy age, individuals interact with online information in reshaped and changed means, from the first launch of the World Wide Web through to Web 2.0

applications. Some of the characteristics of Web 2.0 definition are as follows: 1) data sources become richer as more people use it; 2) long-term objectives are served and supported by providing self-service for customers; and 3) people capitalizing on collective intelligence [28]. Such characteristics extend to health and clinical contexts in the form of Medicine 2.0. The driving characteristics and main realms of Medicine 2.0, as seen in the literature, include user involvement, practice and collaboration influence, potentials for personalized healthcare advances, employment and applications in medical education, and correlated tools and methods [28]. Several components of Web 2.0 and Medicine 2.0 applications are interconnected to different extents with this work. Health social networks emerged, fulfilling health-care and self-management potential of Web 2.0.

Instead of offering passive involvement, these applications enable users to actively engage and participate in the creating and sharing of information. In other words, use of the Internet in an interactive manner was found to be more interesting and empowering for users [8]. Some interactive applications used on Web 2.0 platforms in health information exchanges are wikis, blogs, and podcasting.

In this thesis, particular attention is paid to blogging. Blogging is a form of self-publishing about a particular subject and is one primary element in this work. Blogs are proliferating and are seen as a phenomenon of online health education and among communities of practice forums and groups [29]. Blogs attract people to share knowledge and produce debate about a particular health topic as a community or a group. Two main advantages of such applications are their ease of use and wide accessibility. Educational tools and knowledge exchange platforms for individuals in Web 2.0 applications are difficult to disregard, either from an academic or a research point of view [8].

A clinically-proven, meaningful approach for patients with chronic illness and/or experiencing a life-threatening situation occurs when patients write and express their feelings [7]. Noticeable improvements in health outcome indicators for Asthma groups were obvious in an experiment called Forced Expiratory Volume [FEV1], in which patients were asked to write and express their feelings. Healthcare providers realised possibilities for disease management, beyond symptom control and purely clinical treatment, that would take into consideration psychological and social factors While involving different implications, the type of writing and self-expression involved in this experiment could have similar potentials for better health outcomes in other areas [7].

2.3.1 Use of Web 2.0

In healthcare and healthcare education contexts, Web 2.0 applications and tools are seen as enablers with promising potential for the future [8]. More and more users (including patients) are signing in or just taking “free rides” from different blogs, wikis, podcasts, and social networks to obtain information. Such tools open new possibilities for professional communication and collective intelligence in community-oriented industries like healthcare. In blogs, various groups of patients with a particular clinical condition communicate their opinions and experience, whether for self-expression or for sharing/seeking information. The need to satisfy this type of information-sharing requires tools, the use of which will continue to grow. Thus, research is needed to test these applications and gauge their effective use in health care services [8].

Despite the paradox that healthcare professionals routinely question the credibility of some health- and disease-related information sources found on the Internet, patients will likely continue looking for this type of information outside their healthcare providers' offices [30]. Online communities and support groups have emerged through Web 2.0 applications. These often started as individual effort by and for patients (e.g., "PatientsLikeMe.com"). They have since been acknowledged and recognized as a robust source for supporting patient self-management. Today, there are even considerations to have such Web 2.0 applications integrated with health information systems like Personal Health Records.

When such technologies engage a community of people seeking and sharing the same kind of information, they often develop a whole-group understanding known as Collective Intelligence. Defined as "*the ability of a group or community to apply knowledge in a way that allows the group to adapt and thrive in its own environment*", collective intelligence within online health communities has been linked with better health outcomes [31]. This advantage is supported by more valuable and superior knowledge and data use by healthcare consumers. Patients can learn, from community members involved in a Web 2.0 application, gaining answers to their questions, and supporting health-related decisions and directions. Several observers and researchers have recognised the value of virtual communities and support groups in other contexts. These have shown increases in life expectancy and lessening of health costs [31].

People in many contexts have extensively used online social media and social groups to gather health-related information and knowledge. In the following section I will explore research and academic work done based on this medium.

2.3.2 Online Support Groups

Klemm and colleagues carried out extensive research and analysis on an online support group for cancer patients [32]. Rigorous analyses were done of the postings of persons in on-line support groups within several cancer subgroups (e.g., prostate and breast cancer). Klemm et al delineated their findings into the following four categories. First, individuals contribute to search for information and/or give it to others. Second, individuals posted comments to provide support and encouragement. Third, individuals posted personal views and opinions. Fourth, individuals communicated, with the group, their experience through these groups.

A similar study, on computer-mediated communication for individuals with disabilities, identified two main components of postings to a website by users and group members. [33]. One component was messages of support, personal feelings and expressions, making a socio-emotional contribution. A second component was task-oriented postings, where members asked for information and/or initiated problem-solving conversations.

In a more focused arena, Scanfeld and his colleagues acknowledged social media as a legitimate medium for information exchange [34]. In their observational study, they used an inductive approach. Within a large micro blog, they analyzed content of public messages in Twitter© postings in which the term “antibiotic(s)” was used , utilizing specific terminology co-occurrence in the search function. [34]. Sending and giving information, opinion, and advice was found to be the second most significant category analysed. This tends to confirm that the most common reasons for people to be members of online groups and communities is to participate with others and obtain information [35].

In a breast cancer group, the virtual sending and receiving of information was found to be the most common form of interaction [36]. Findings regarding communication-content classifications were similar to previously noted research (i.e., for social support, information trading and interchange, and individual empowerment). In addition, the authors suggested a potential for greater understanding and communication between patient and provider.

From the perspectives of consumer/patient empowerment and cost-effectiveness of healthcare services, shared communication among patients can minimizes the reliance on health professionals and provide health information and educational resources at minimal cost [37]. Members of health-care related groups using virtual communication also found a strong source of empowerment. These benefits are important for our target population of Asthma and COPD patients using a web-enabled AP.

The following section reflects some possible disadvantages for users, related to using health related information available online.

2.3.3 Online Health Related Information and Support Groups Drawbacks, Harmful Effects

This review identifies a number of healthcare information resources found in different online forms. To date, the quality of data in these sites is has not been certified, and systematic approaches to evaluate them are yet to be established [38, 39]. It is clear that more users and patients join peer-to-peer support groups, and that the number of these groups is rising. However, the potential benefits (or possible harm) of many of these sites on health outcomes remain to be determined. This makes it difficult to confirm advantages and disadvantages of the online interaction. Although these sites were used for research purposes they were not assessed for their effectiveness in delivering harmful or beneficial information. The sites were not assessed for their effectiveness in delivering harmful or beneficial information [38]. Rather, they have been only a component of larger-scope studies [38].

Additionally, access to online information can be limited due to user characteristics such as skill levels and overall disparities, available resources, and search difficulties [39]. Some scholars have named our technological age as an “information epidemic”, referencing the potential harm of online health-related information on users and patients [39]. Information reliability is a main concern here. The other two drawbacks are inadvertence to available clinical recommendations and accuracy issues that arise when reference to related health professional sources is not provided [39]. Assessing a site’s credibility can likewise be challenging for both researchers and patients/users [40]. Sites developed by individuals could look as good and “professional” as those created by professional or governmental organisations [40]. With this in mind, some scholars suggest that information access in Web 2.0 applications is best and most safely utilized as a complementary tool to a health professional’s advice and guidance [39-42].

Scholarly research on computer-mediated communication among different groups of patients addressed health education, social, and other perspectives. This will be presented next. The following section will focus on the health informatics perspective of this phenomenon.

2.4 Health Informatics Perspective and Consumer Health Vocabulary

Several health information technologies warrant review and improvement of quality and performance efficiency [43]. Access to care through information regarding healthcare service locations and points of care is increasingly being provided through health information technologies. Advanced applications will also support patient involvement and engagement as a consumer [43]. Beyond standards and regularity policies, health information exchange platforms are considered as incorporating both personal (individual patient) and family involvement. This is driven by growth in demand by users (patients) to be enrolled in the process of their health care plans.

Thus, health information is seen to advance in three dimensions simultaneously: longitudinal records related to each patient (including clinical observations and exams); scientific components supporting personalised medicine; and contextual facilitating information used by different healthcare providers, using health-centric social networking and personal health records [43]. Healthcare leaders have started to recognise consumer engagement as a form of individual empowerment and as a means to improve the quality of care [43]. Based on a co-evolution argument, the three-dimensional growth of such advancements is parallel [43]. I personally envision that consumer health informatics applications could facilitate a contextual aspect through communication.

2.4.1 Consumer Health Informatics / Vocabulary

In this thesis, patients' communication via online Asthma and COPD support groups are evaluated against clinical data standards as stipulated in a standard Asthma/COPD AP. It is my interest to match Consumer Health Vocabulary (CHV) terms with professional clinical data standards. Consumer Health Informatics (CHI) is described as an interface designed for consumers, to obtain disease-related information that promotes healthy behaviours; and support information exchange among peers within a network that embraces social support [44]. Such an interface is also designed to enable consumers to make informed health-related decisions [44]. By using advances in CHI communication and patient education the objective is to enable better health outcomes and lead to effectual health-related decisions [45, 46].

Derived from Medical Informatics (which deals with technological advancements in information, knowledge representation, and cogent developments among academics and healthcare providers [47]), CHI focuses mainly on patients [44]. As demonstrated by healthcare systems with interfaces designed specifically for consumers; CHI systems deploy innovative communication processes among and with patients, through information technologies [44]. Throughout the review, the objective is to patients with information [44]. Additional objectives include encouraging individual self-care, promoting better and healthy behaviours, supporting information exchange at peer level, embracing social support, and enabling patients to make informed decisions about their health [44].

With respect to the context of my research population (Asthma and COPD patients), and coupled with AP use, all CHI objectives addressed above do empower patients. By definition, CHI applications are patient-centred. CHI was derived and evolved originally from consumers seeking knowledge and information from the Internet [44]. CHI assists individuals in reaching and determining their own health-related decisions regarding their health in Web 2.0-enabled patient communications.

Individualised patient care is also possible through Web 2.0-enabled patient communications, Through CHI, better health outcomes have been reported [45, 46].

Information technology for health care providers is virtually guaranteed to grow. This is evident from the wide spread use of computerised medical/health records and the growing potential to integrate them with clinical decision support systems. Now, we are also seeing the growth of use by groups of health-care consumers. Figure 1 was extracted from the literature and explains the current focus of medical informatics [47]. This is seen through the following dimensions: 1) Having medical knowledge available for consumers; 2) having accessibility to electronic health records by patients, and 3) facilitating consumer choices through consumer decision support systems [47]. CHI focuses on supporting consumers to acquire and use health information through developing and evaluating approaches and applications.

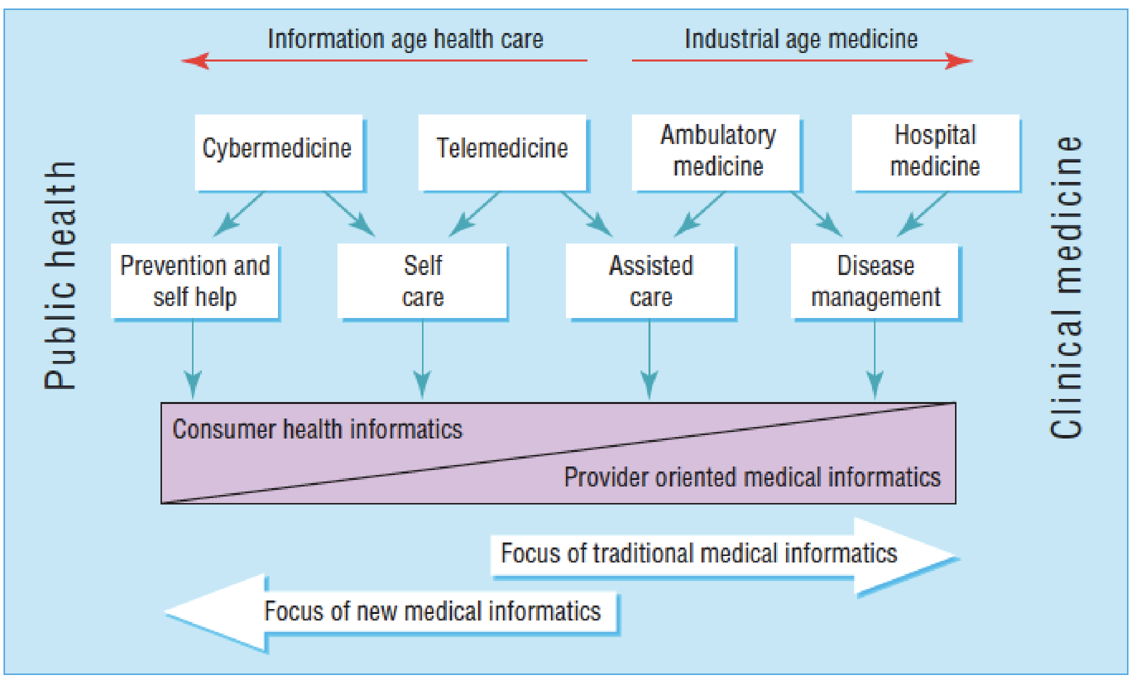


Figure 1 The focus of traditional medical informatics is shifting from health professionals to consumers [47, p. 1714]

CHI goals and objectives advance into examining and analysing health conditions to outline, apply, and maximise potency for consumers [47]. This offers potentials to lessen the knowledge gap among both patients/consumers and professionals in healthcare services. Nevertheless, individuals who are illiterate with regards to technological advances in information must also be considered. With consideration to the work presented in CHI literature and the focus shift presented above in Figure 1, the analytical framework developed in this thesis concentrates on “prevention and self-help” as part of “cyber medicine”. The following section will shed some light on a few examples of CHI applications.

2.4.2 Applications of CHI

In this thesis, principles of CHI will be applied to explore the extent to which the consumer health information exchange, via social media, adheres to professional standards, using Asthma and COPD AP documents as a guide. This section presents CHI applications and systems for health consumers’ empowerment [49, 46]. It also presents further evaluations and research implications on the publicly-accessed social networking platform, called PatientsLikeMe [49-52].

The promotion of teamwork in healthcare models, where professionals and patients share the process of decision making, is coming of age [48]. Most consumers have access to knowledge bases through the web, using the freely-accessible National Library of Medicine databases. Over and above that, knowledge representation has made it easier for patients/consumers to understand healthcare language, through advanced strategies in web-based applications [48]. Healthcare researchers and providers are approaching with confidence the use of tools to promote shared decision-making in healthcare models. With technological advancements and leaders recognising their potential, this model could be reversed, allowing patient engagement and involvement in their healthcare decision-making [48].

Distributed Guidelines (DiG), a CHI application, provides a framework for software that enables healthcare consumers to benefit from professional health-related information, despite knowledge variances [49]. Using clinical guidelines combined with public health record systems, deployed through semantic web and ontology modelling tools, DiG's main objective was personalised health. The aim was to exploit and extend professional knowledge so that such knowledge would be be available and intelligible for lay people. DiG provides an assessment tool for patient use through a web application based on data available on Personal Health Record (PHR).

DiG has two pivotal principles. The first is patients sharing their experience with a particular clinical condition. The second principle is being open through social web application like online groups and discussion forums. Using these principles, shared personal health information is executed through health information systems and has enabled valuable use of data from the PHR [49]. PatientsLikeMe is one example of a patient-centred platform. It represents health data provided by its members in a graphical display that can be viewed by other members within this community. By definition, social networks in a health context, are websites where individuals access health resources in any number of forms [51]. These forms vary from information on clinical trials to group exchanges for motional support and peer information exchange.

PatientsLikeMe is recognised as the largest health social network as well as the best-known one [50]. Comments within PatientsLikeMe were analysed, with the results showing that sharing data among members not only benefited other group members; it also benefited original information owners who improved their disease self-management [50]. Further, PatientsLikeMe has shown advantages in clinical management, clinical outcomes, symptom control, and treatment decisions [52] The platform is well recognised for clinical research [52]. On-line Groups involved in this analysis included (but were not limited to) Parkinson's disease, human immunodeficiency virus, multiple system atrophy, and mood disorders.

2.4.3 Components of Consumer Health Vocabulary

In this thesis, consumer health information that is exchanged via social media is categorized according to two AP (Action Plan) components: symptoms and actions. CHV (Consumer Health Vocabularies) is used as a guide to evaluate knowledge exchange among the virtual communities in relation to the clinical data terminology standards of Systematized Nomenclature of Medicine (SNOMED) and Unified Medical Language System (UMLS).

Scholars are increasingly studying to understand, structure and upgrade the process behind information search, to support both professionals and non-professionals wanting access to health-related information through web-based applications. As mentioned previously, language terminologies used by patients to search medically-related information was found to be an issue or dilemma [53]. Understanding users' behaviours to search for information could have potential impacts on supporting these systems' ability to retrieve and provide information to users. In an attempt to support better access to information by users of the National Library of Medicine database, queries were performed through this database [53]. One of the main issues that emerged in this analysis was related to concepts and terms used by individuals to search the database. Search terms and concepts were not completely mapped within standardised terminologies like the UMLS. The following section will provide insights about concepts and applications of CHV as initiatives for enhancing efficiency of information for patients as web users and healthcare services consumers.

Words and phrases used by consumers to describe their health-related problems, perform a research about these problems, or communicate to their healthcare professionals regularly vary from those used by health professionals [54]. Healthcare professionals share basic domain knowledge and education that is considered when developing terminology standards. On the other hand, consumers' language use in this domain will be modelled by different cultural and social backgrounds and influenced by formal or informal education. CHV applications could provide better public access to health-related information. Such applications could also improve individualised information according to needs and circumstances [54].

Since consumers are increasingly interested in seeking health-related information and in taking initiatives to be more responsible and involved in their personal health care, the vocabulary gap phenomenon will likely increase [54]. In view of this vocabulary gap, CHV initiatives relate to an array of attitudes, expressions, beliefs, and concepts used in health-related communications among members within consumer discussion groups (e.g., online discussion forums, consumer health websites users). With the intent to support two-way communication between patients (consumers) and healthcare information systems and applications, CHV is concerned with information retrieval [54]. Query term searches generated by consumers do not relate to clinical data terminology standards (e.g., UMLS). As a result, a consumer's search will not yield the same quality of information as it would with professional terms. The CHV initiative enables the mapping process of consumers who enter search terms, in order to standardise to the known professional terms; and then to automate these for future searches. The CHV used in this thesis operates as a larger-scope reference of language used by patients in their online communication, to complement mined concepts from Asthma and COPD groups.

2.4.4 Open Access and Collaborative Consumer Health Vocabulary Initiative and Online Social Networking

In this thesis, health information exchange is divided into threads, with each thread addressing a unique health issue. The concepts map for each thread, which includes both actions and symptoms, was developed using social network methods. For each thread name, network maps were constructed. This section of CHI literature review sheds more light on the development stages and validation process of CHV interface and integration of clinical data terminology standards (UMLS and SNOMED-CT) as well as mined concepts from health-related online social networks [56-60].

With the proliferation of Web 2.0 applications and their uptake by patients (e.g., online discussion groups and forums), existing consumer health vocabularies and language is exacerbated by the consumer-professional gap [55]. Community-centred characteristics of Web 2.0 generate and provoke user-created taxonomy recognised as *folksonomy*. This escalates content use over semantic architecture, having originated in a contrasting side from *taxonomy* (a branch of science concerned with classification). Folksonomy encourages and enables networking of interconnected concepts and interests (e.g., cancer or diabetes online groups), whereby users can become connected by sharing their views and experiences. Informal patient consultations thus transpose into networking, and language users and generators prompt health information system developments and improvements [55].

During data exchange and processing, matching symptom terms used by patients and clinical data terminology standards (e.g. UMLS) were carried out. Unmatched terms were identified, with reasons for lack-of-match analyzed. Some of the identified reasons for unmatched terms were fragmentation, expressions that nested multiple clinical concepts, and indistinct terms. However, social networking supports and encourages knowledge construction that could subsequently induce new classification, coding directives, and vocabularies of healthcare concepts [55]. The next section will represent a larger scale vocabulary development with the cooperation and involvement of patients as healthcare consumers.

Consumer Health Vocabularies (CHVs) initially were developed by assessing consumers' readability of, and familiarity with, health-related concepts and terms [56-58]. In this preliminary pilot study and in congruence with the general objective of Consumer Health Informatics (CHI) to facilitate and ease consumers' comprehension of health-related information, a model was developed to assess consumers' familiarity with technical terms at two level:1) the lexical form of the term, testing "surface-level term familiarity"; and 2) the basic meaning of the concept, testing "concept-level term familiarity" [58]. A predictive model for enhanced consumer understanding through CHV term acquaintance was achieved through a practical approach in evaluating each level of terms. .

Because users of blogs change and progress in their use/understanding of terms, as described in research studies and analysis with CHV, a systematic method for identifying candidate terms has been integrated in consumer-targeted educational websites and other health information system applications [59]. The two principal stages in developing vocabulary are: 1) strings identification, considered to be a candidate within a particular domain; and 2) deciding on identified strings to be included as valid terms in the vocabulary (“termhood determination”). There is presently no general solidarity in this step with known health vocabularies like SNOMED and ICD-9. Exploring and identifying CHV can be challenging due to the extensive range of health consumers as well as their different backgrounds and experiences [59]. Narrowing the scope to an individual population of a specific task was found to be useful for codifying the health language of lay individuals.

An Automated Term Recognition (ATR) approach has been deployed to assist in selecting candidate terms for human review. Two methods from this approach (ATR) used in the open access CHV initiative are described. In the biomedical domain, ATR is concerned with general term search systems that would search terms against named entities. These named entities were usually recognised if they matched an ontology or directory of terms and had a meaning in a context outside the one being searched [60]. The two ATR mechanisms used here were an Information theory-based approach (t-test) and a syntax-based approach (C-values). Models were established for each method and were applied successfully in “termhood recognition” [59]. Multiple stages to identify terms were recognized. Reviewers distinguished and identified approximately 1890 terms to be part of the Open Access Collaborative-Consumer Health Vocabulary (OAC-CHV) project. Reviewers collaboratively reached termhood criteria derived from multiple disciplines like health informatics and computer science [59]. This project will be referenced further in this work, along with the CHV list.

Terms used by healthcare professionals continuously migrate into their patients' realm. These new terms are then taken up into popular terminology [60]. As language evolves, it is important that the CHV be updated to assimilate and reflect language-use changes among patients/consumers, and to maintain the effectiveness of its performance in consumer health application systems [60]. A Computer-Assisted Update (CAU) system was developed for the OAC-CHV to have both collection and production of changes automated, and subsequently to support its validation [60].

The architecture of the CAU system incorporated three processing stages to extract terms from the assigned website, apply several UMLS and CHV filters, and yield results evaluated by humans [60]. Specifications of each stage are beyond the scope of this section. However, a system methodology, supporting vocabulary maintenance, was shown to be necessary and proven. Mining online social networks is an integral part of the CAU system [60]. Using websites like PatientsLikeMe provides a prosperous medium for CHV extraction. Compared to earlier work done with CHV and terms identification and extraction, PatientsLikeMe yielded a higher quantity of candidate terms than MEDLINEPlus [60].

2.5 Summary

User-generated (i.e., patient-generated) data and information gleaned through web-based applications is recognised as an information system. It can encompass online support groups and disease specific discussion forums. However, terms and concepts used by professionals in health care and communicated to patients (whether via AP or patient-accessible health record language) are understood, utilized and comprehended at different levels. Analysing patient-generated information systems prompts health/medical informatics researchers to explore the facilitation of health information technologies that empower consumers to become more involved in their healthcare decision-making process. CHV initiatives have been presented here as both an example of research interest and as supportive data for the next chapter on methodology.

CHAPTER 3 METHODS: DATA COLLECTION AND ANALYSIS

3.1 Chapter Outline

In this chapter, systematic methodologies and approaches of social media data collection and analyses are presented. The first section includes details of the two data collection processes using existing social network sites as well as the development of a research-specific social media site. Next, the methodological framework, based on health informatics principles, applied in data collection and analyses is explained. An overview of the methodological framework developed by Dr. Weerasinghe and applied in this research is provided in detail elsewhere [61, 62].

From a Health Informatics theoretical standpoint, this chapter explores communications standards that patients, as non-clinicians, use for social networks, against communications standards and methodologies used by clinicians . It also highlights applications of patient network data mining, management and analyses.

3.2 Data Collection

The principal source of data collection is via Internet communications through blogs or weblogs. A weblog, or "blog" is a web-based presentation that facilitates two-way communications, and uses a Web 2.0 Internet tool. In a blog, participants enter dated comments on a particular topic. These are generally shown in reverse chronological order, with recent posts listed at the top or beginning. The following two subsections describe the process I developed for choosing existing Internet communication sites related to health issues that are objective-driven and research-specific; and for the development of two new sites.

3.2.1 Selecting an Existing Internet Communication Source

While there are well-established guidelines for literature reviews of scientific papers, there are no accepted protocols for reviews of electronic media for health information exchange data collection. The following steps were undertaken to select electronic health information exchange sites for data collection.

Sites were chosen where the target population of the electronic health information exchange site is Canadians.

Blogs were selected over Facebook© and Twitter©, since blogs may feature posts in the form of Question and Answer, subject-guided, and/or free posts. As well, links are given to useful resources or publications. Users have the option to add new posts, read existing ones, add comments, or just read the posts and comments. The free, user-friendly, open and engaging environment of these blogs is also available for patients, but the number of such blogs is limited at the present time.

Asthma- and COPD-related organizations and institutions were approached to obtain details of social media and interaction platforms. An environmental scan was conducted of the existing blogs identified by these organizations. Patient-centred discussion and communication forums were the primary objectives of my analysis, regardless of credibility and/or validity of the information provided. Having a community of patients or their caregivers connected in these blogs was considered highly advantageous. Having professionals involved would have been a robust additional advantage to these blogs, but was not a determining criterion.

The characteristics that were used for selection of messages were first: individual postings that involved simply typing a message and posting it for others, and second: that the site maintained an archive of previous postings [63]. The following steps describe the search and screening criteria that I developed for this thesis research. The findings are summarized in Appendix C.

1. Determining the type of network exchange; Blogs versus Facebook© and Twitter©: Asthma and COPD blogs were found to be superior to Facebook© and Twitter© for our research objective. The blogs contained the features pertaining to this research, such as having posts in the form of Question and Answer, subject-guided, and/or free posts, as well as links to useful resources or publications. Users had the option to add new posts, read existing ones, add comments, or just read the posts and comments. Thus, blogs provided and supported tacit knowledge and experiential knowledge transfer and sharing. Facebook© and Twitter© links were also available in these blogs and served as additional media for users' comments and communication . As well, disease-specific online groups that were chosen for this thesis research in the form of a blog gained considerable attention from healthcare researchers [31-33]. Thus, the principles of health information communication via secure transmission of health-related data are fulfilled.
2. Determining search engines: General and blog-specific search engines (Google and BlogScope) were used with a combination of the following search terms: "Asthma", "COPD", "health", "online", "groups", "discussion forum",

"discussion board", and "communities". The search resulted in identification of 25 related websites, as listed in Appendix C.

3. Blog assessment: Further evaluation of these websites was made according to subject, participants, demographics, Facebook© and Twitter© incorporation, identification of Asthma/COPD as the subject matter. Once a social network site was selected, the subjects of discussions and posts types were assessed as to whether they have: (a) guided subjects posts, (b) free posts, and (c) questions and answers. Participant types of each site were assessed as (a) patients only, (b) professionals only, and (c) both patients and professionals. Of the 25 websites that used the above-mentioned criteria, 10 supported blogging and discussion board features, 13 included Facebook© and Twitter© features, and 6 included Asthma or COPD as either main or subcategories of the website.

As our target population was Canada, only two websites made the final screening stage: *copdcanada* [64] and the Asthma Society of Canada [65]. These two sites covered our specific context of Asthma and COPD solely, while another online patient community, *PatientsLikeMe*, included Asthma among other disease groups [66]. The next few paragraphs provide further details about these three candidate blogs selected for data collection.

The COPD Canada patient network constituted the first source of my online peer-to-peer communications and discussions analysis [64]. The majority of its members are COPD patients, related health professionals and related caregivers. As of the date of this writing, there are 718 members involved in this message board, with 7,068 different topics and 67,000 posts in total. Although it is a Canadian group, members come from around the world, including Europe and the United States. I collected my datasets from the “General Discussion” forum, a subsection of the “General” forum. In this blog, one member will start a topic thread to which other members will respond. Message thread topics and subjects are heterogeneous but typically involve medication use, symptom control strategies, and general or specific disease inquiries.

The second candidate blog was the national asthma patient alliance [65]. It advocates providing evidence-based, Asthma-related information to support self-management, disease prevention, and health care research. At the time of writing, the discussion board covered 48 different topics, and viewer numbers reached more than one thousand in some topics.

The third blog was “PatientsLikeMe ”, which was established by a team of professionals to provide patient support [66]. At the outset of each visit, the site asks the user a few questions to direct him/her to the appropriate group with similar conditions. It also provides different visual charts for every participating member, including scales for Quality of Life and Distress Components. Patient discussion boards are grouped according to clinical conditions, with Asthma and COPD tags being listed in the “Lungs and Respiratory” room. This example discussion board provided features that were beyond what is considered basic and offered more advanced support to users.

3.2.2 Data Collection from Three Candidate Weblogs

The sampling approach used was to collect all text information exchanged between September 2010 and August 2011. Details of the sampling plans for social network analyses are explained elsewhere [61, 62]. The unit of analysis is a blog, and each blog was categorized into a thread that contained Asthma/COPD-related health information exchange.

The text data collected from the selected websites weblogs are described below. Each topic or conversation with different replies was recognised as a thread, with thread content being collected within a one-year timeframe aiming to cover the peak seasons of COPD (winter and spring). Since the aim of the research was to analyse the text data from blogs for their relevance to AP (Action Plans), the threads were not classified or refined according to subject or topic title at the primary data collection stage. Instead, all thread discussions within this timeframe were considered suitable for further analysis.

3.2.2.1 Profile of COPD Canada Network Data

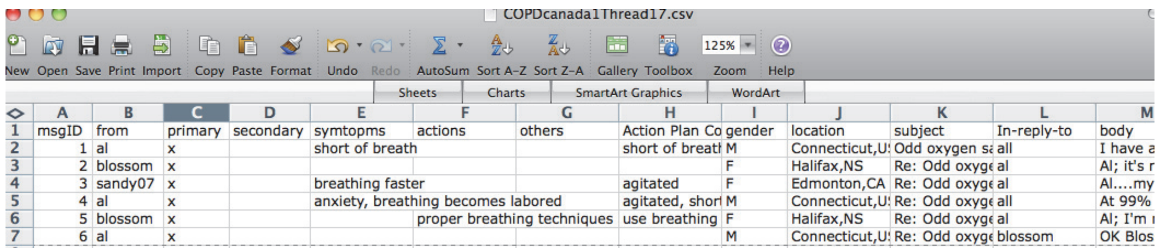
Altogether, 106 thread messages were identified. Topics varied from symptom-related (e.g., “shortness of breath and anxiety” and “coughing up phlegm”) to medication-related (e.g., “antibiotics” and “spiriva am vs pm”). Topics referenced to context-specific procedures were also present (e.g., “bronchoscopy procedure” and “oxygen at the work place”). Further posts involved direct questions such as “when do you go to the ER?” and “how do you use inhalers?” The number of viewers or readers of each of these 106 topics was 563 on average, with a minimum of 37 viewers and a maximum of 670 per topic. Contributions and responses for each thread topic ranged from zero responses (i.e., the thread consisted solely of the original poster’s post) to 53 members discussing the same thread topic, with 12 as the average number of readers per thread.

3.2.2.2 Data Cleaning

The next step involved data cleaning, the details of which are explained elsewhere [61]. Data cleaning includes deletion of unrelated information. Hence, the thread contents relevant to my research question were manually assessed for clinical contents and appropriateness to the research questions. The purpose of the manual assessment was to filter out threads and messages that transcended the domain knowledge exchange. Terms and phrases that were considered relevant to the domain (i.e., COPD Action Plan contents) included “emphysema”, “antibiotics”, “shortness of breath”, “cough”, and “flare-up”. Sixty-four potential threads (five of which were found to be directly related to the Action Plan use for COPD) underwent additional rigorous analysis according to the methodological approach addressed further in this chapter, while 42 irrelevant topics, such as “repairing body tissues”, “chlorine and our lungs”, and “research fraud”, were excluded.

3.2.2.3 Creation of Text Database

After carrying out the data cleaning for each of the 64 threads selected, I created, in Excel 2008, an independent dataset consisting of twelve main attributes [67]. These attributes are *Subject, MessageID, Body, in-reply-to, Location, Gender, Primary, Secondary, Symptoms, Actions, Others,* and *Action Plan Concepts*. One example is presented below in Figure 2. *Subject* and *MessageID* are intended for thread identification in general and for the automated text analysis. The *Body* attribute contains the entire text content of the thread message. *From* and *in-reply-to* consist of tag names that members used to identify themselves within the analysis. *Location* and *Gender* attributes provide further details about members. *Primary* and *secondary* are indicators of member node positions (whether the person started the blog or responded to the query) within the network (this will be addressed in detail in the social network analysis in the Method section). *Symptoms, Actions*, and *Others* contain words or terms used by patients to describe their symptoms or actions that I considered to be relevant to Action Plan components. Finally, in the *Action Plan Concepts*, I entered and matched corresponding Action Plan terminology to that used by patients in *Symptoms* and *Actions.*



	A	B	C	D	E	F	G	H	I	J	K	L	M
1	msgID	from	primary	secondary	symtopms	actions	others	Action Plan Co	gender	location	subject	In-reply-to	body
2	1	al	x		short of breath			short of breatl	M	Connecticut,U!	Odd oxygen sa	all	I have a
3	2	blossom	x						F	Halifax,NS	Re: Odd oxyge	al	Al; it's r
4	3	sandy07	x		breathing faster			agitated	F	Edmonton,CA	Re: Odd oxyge	al	Al....my
5	4	al	x		anxiety, breathing becomes labored			agitated, short	M	Connecticut,U!	Re: Odd oxyge	all	At 99%
6	5	blossom	x			proper breathing techniques		use breathing	F	Halifax,NS	Re: Odd oxyge	al	Al; I'm
7	6	al	x						M	Connecticut,U!	Re: Odd oxyge	blossom	OK Blos

Figure 2 Example Dataset

The above example dataset incorporated data from selected candidate weblogs. One source involved data that was directly imported and extracted from the discussion board (e.g., *Subject, Body, From, in-reply-to, Symptoms, Actions,* and *Location)*, and the other was entered manually as part of text analysis and 'network of information exchange' creation and analysis. These datasets were converted into different formats as required by the analytical software. The next subsection presents resources for data collection in addition to the selected candidate sites described above.

3.2.3 Data Collection: Developing Research-Specific Sources

Data was collected from a newly designed web portal on adherence to COPD and Asthma action plans, specifically addressing this research objective. Using Drupal, an open source contents management system, I developed both *copd.patientnetwork.ca* and *asthma.patientnetwork.ca* with its basic modules [68,69]. The website layout design includes four main forums, as shown in Figures 3 and 4 below. The basic features of a blog: free posts and comments are provided; and my professional involvement as an experienced Respiratory Therapist is referenced . While the aim is to serve a target patient population living in Nova Scotia, Canada, and diagnosed with Asthma/COPD, membership is open to anyone with access to a computer. Both websites were publicized and advertised through the Canadian Lung Association nationally, and through existing online patient groups.

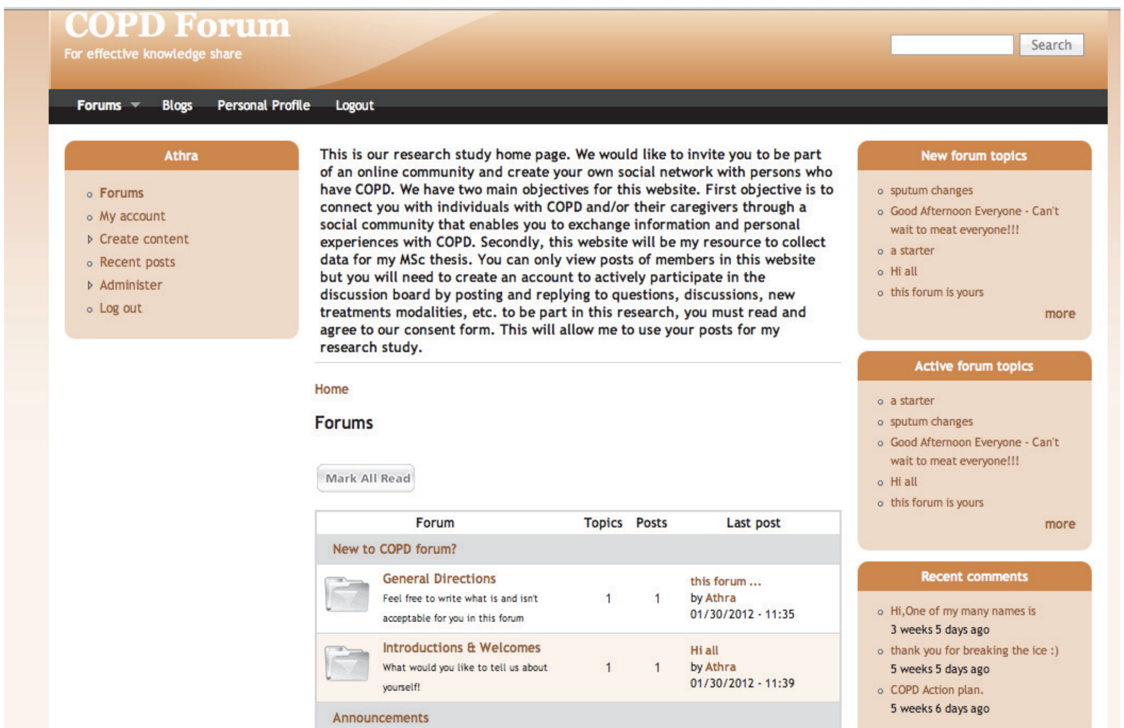


Figure 3 COPD Forum Homepage

Figure 4 Asthma Forum Homepage

Individuals interested in being part of the study and signed into the website as a patient and/or caregiver were asked to answer a few questions as part of the research's demographic data. Requests for membership were reviewed prior to approval. Requests for membership that did not meet the criteria were denied.

Through these websites, I initiated discussions about AP use and asked additional related questions, such as: 1) Do you have an Asthma/COPD AP? If not, why not, and if yes, do you use it? 2) In your opinion/experience, what are the benefits of using an AP? 3) What are the barriers to using one? I also posted additional questions that were related directly to AP symptom-and-action recommendations, aiming to lead specific discussions. In addition, users were granted freedom to post views, opinions, and/or questions.

Both sites went live in January, 2012, after approval from the Health Sciences Human Research Ethics Board.

The next section provides details of the developed framework.

3.3 Data Analyses

Understanding knowledge shared in a virtual community, and the relationship of such knowledge to specific clinical guidelines (i.e., AP components in this research) are not well defined by the literature. Knowledge exchange models in virtual communities are influenced by a myriad of factors. One element is the willingness of members within these communities to share information and subsequently foster knowledge and collaboration. This attribute was found to be influential in defining the knowledge wealth of virtual communities [70]. An example model is shown below (Figure 5). This model, incorporating theories of Social Capital and Social Cognition, has been developed for knowledge sharing and exchange among virtual communities [69].

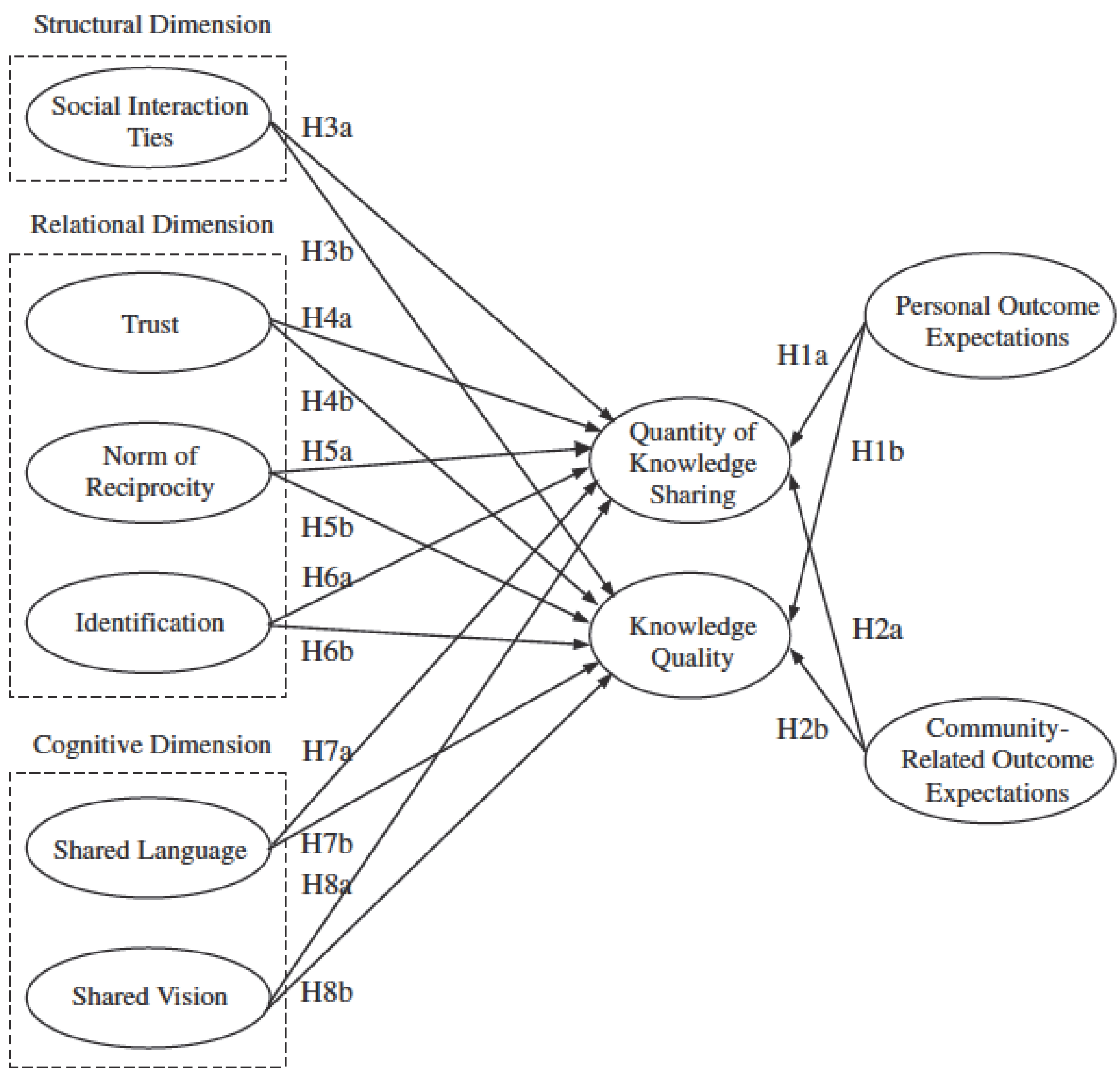


Figure 5 Online Communities Knowledge Exchange Model [70].
"H1a & H1b: member's personal outcome expectations are positively associated with the quantity (a) and quality (b) of knowledge shared, H2a & H2b: members' community-related outcome expectations are positively associated with the quantity (a) and quality (b) of knowledge shared among them, H3a & H3b: members' social interaction ties are positively associated with the quantity (a) and quality (b) of knowledge shared, H4a & H4b: trust is positively associated with the quantity (a) and quality (b) of knowledge shared among members, H5a & H5b: norm of reciprocity is positively associated with the quantity (a) and quality (b) of knowledge shared, H6a & H6b: identification is positively associated with the quantity (a) and quality (b) of knowledge sharing, respectively, H7a & H7b: shared language is positively associated with the quantity (a) and quality (b) of knowledge shared among members, H8a & H8b: shared

> vision is positively associated with the quantity (a) and quality (b) of knowledge shared among members"[63].

Using the framework of structural, relational, and cognitive dimensions modelled above (Figure 5), the following elements were tested for their effects on quality and quantity of knowledge exchanged among individuals of an online community: social interaction ties, norms of reciprocity, and shared language. Within our Asthma and COPD context: trust (H4), identification (H6), and shared vision (H8) were already established. Their effect on knowledge is beyond our research objectives.

The elements of social interaction ties (H3) and norms of reciprocity (H5) were examined for connection patterns in two kinds of exchanges: individual exchanges (among members) and network exchanges (within one thread). Social Networks Analysis (SNA), is addressed in section 3.3.4 below. Shared language (H7) was looked into extensively in the text analysis section (3.3.2).

This model provided a broader overview of factors that influence knowledge-sharing among virtual communities. This was accomplished by building on the general structural concepts of the medical informatics shift toward consumers (Figure 1 in Chapter 2) and integrating the knowledge exchange model with previous work found in the literature [46, 54, 61]. The following analytical framework (Figure 6) was developed. The diagram shows various analytical steps followed in this thesis, where health information classification systems are used for data management and analysis.

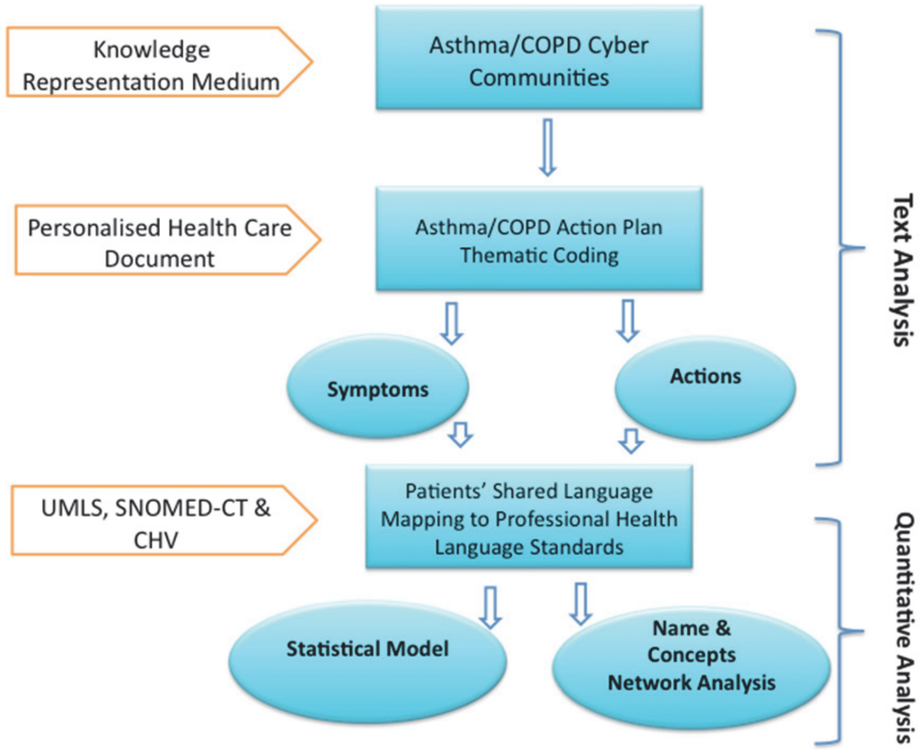


Figure 6 Analytical Framework

This analytical framework provides an approach for assessing Asthma and COPD information and knowledge exchange among online groups via blogs. It is constructed with reference to established methodologies in [61, 62] and integration of clinical data representations sources (UMLS and SNOMED-CT). The first of selection of knowledge representation medium was explained in sections 3.2.1-3.2.3. The second step of thematic coding is explained next.

The following subsections will address the methodical approach. Steps specific to the context of health information diffusion patterns through social networking are detailed in sections 3.3.2, 3.3.4, and 3.3.5 based on an already existing framework developed elsewhere [61].

3.3.1 Thematic Coding of Action Plan Elements

After careful review of the content of the Action Plan clinical document (Appendixes A and B), two main themes of symptoms and actions were developed. Clear naming with description of characteristics of these elements supported accuracy of coding [71]. Symptoms included those reported by patients, and actions were the recommended steps and measures for patients to take based on Canadian clinical standards. Tables 1 and 2 below illustrate COPD and Asthma AP classified elements. Guided by the contents of Tables 1 and 2, an inductive approach was used in subsequent text analyses. The AP elements established in the tables were also used to guide manual text analysis of thread messages, where manual analysis was done to circumvent limitations of automated text analysis.

Table 1 COPD Action Plan Components Thematic Classification using standardized terminology and coding

Theme	**Units**
Symptoms	Short of breath (S1)
	Coughing up sputum (S2)
	Regular cough (S3)
	Sputum color, volume, consistency (S4)
	Extremely short of breath (S5)
	Agitated (S6)

	Confused (S7)
	Drowsy (S8)
	Chest pain (S9)
Actions	Start medications for COPD flare up (A1)
	Avoid aggravating symptoms (A2)
	Use breathing, relaxing, body positions, and energy conservation techniques (A3)
	Increase oxygen (A4)
	Notify contact person (A5)
	Start antibiotics (A6)
	Start Prednisone (A7)
	Go to hospital emergency department (A8)
	Dial 911 (A9)

This was the first phase and preliminary step in my virtual community information and data exchange text analysis. The AP component classification defined the context to pilot a larger text mass analysis. The text analysis of an informal virtual community is generally considered challenging due to lack of specific regulations and the wide variety of educational and clinical backgrounds and experience of participants. This contrasts sharply to a professional online community that typically has a shared educational background as well as a basic formal knowledge foundation [58,72,73].

The critical steps of performing text analyses, determining the data processing, and setting data boundaries were set according to our pre-determined timeframe of data selection. Concepts aligned with AP units provided in Tables 1and 2 were identified, described and analyzed.

Where new codes emerged while exploring the data, these were included. Text analysis also included deciding on the inference level of concepts to be analyzed and drawing contextual and semantic inferences [59]. This will be addressed in detail in the following sections.

Table 2 Asthma Action Plan Components Thematic Classification using standardized terminology and coding

Theme	**Units**
Symptoms	Shortness of breath (S1)
	Coughing (S2)
	Wheezing, chest tightness (awaken from sleep more than once a week) (S3)
	Prolonged cough (with or without phlegm) that is usually not present (S4)
	Above symptoms more than twice a week (S5)
	Inhalers don't relief for the full 4 hrs (S6)
	Increased use of relief puffers (S7)
	Peak flow meters numbers drop (S8)
Actions	Remove and avoid symptoms triggers (A1)

	Increase your puffs (as prescribed) (A2)
	Add new puffers (as prescribed) (A3)
	Take prednisone/medrol (A4)
	Hold/continue inhaled steroids (A5)
	Call your doctor for further instructions (A6)

3.3.2 Online Discussions Text Mining, Text Analysis and Meta-Matrix Database Development

Text mining, or Internet Community Text Analyzer (ICTA), also called NETLYTIC-ICTA [74], is a web-based system for text analysis used in this research. Among other capacities, it facilitates text analysis in social media [74]. The visualization techniques, incorporated with data mining methods in NETLYTIC-ICTA, allow users to examine textual knowledge inherent in a given dataset. Users are given the choice to analyze their datasets in four major steps: 1) Import: where researchers upload their data into the three different formats of RSS feed, text files (.txt), and comma separated values (.csv); 2) Clean: which supports further elimination of potentially noisy (unrelated information) data after importing them into the system; 3) Text Analysis: which performs automated analysis on uploaded text files; and 4) Network Analysis: which provides network visualization of members who were involved in information exchange, using either the Chain network method or the Name network method, according to data type and research objectives.

The next paragraphs provide a more in-depth description of the analytical approach adopted in this thesis.

Adhering to the analytical framework that was originated in a health information diffusion context and developed by S. Weerasinghe, the following steps are carried out in the text analysis [61]. Text pre-processing is considered the initial step in text mining [61,62]. For the purpose of this study, collected textual posts were entered in Microsoft Excel tables, as detailed in the data collection section above (3.2.1). These were then converted into ICTA readable format as a file of Comma Separated Values (.csv).

Importing data constitutes the first step in text mining (analytic process described elsewhere [61]). Since I uploaded data collected manually, the Text Data Cleaning step was skipped (i.e., text cleaning and pre-processing were done prior to uploading data into NETLYTIC). The second and third steps were carried out with some modifications suitable for this research. Further text deletions and stemming were done retrospectively and compared to the AP contents, which were used as a guide to concept detection and categorization. Additionally, information diffusion evaluation was performed as outlined in this context [61].

Text Analyses: Once the data files were imported, datasets were created and made ready for analysis. In the NETLYTIC Text Analysis step, the "Analyze" command generated two outputs: "Results" and "Visualize" (Figure 7). "Results" showed the generated interactive tag cloud (grouping) of popular (or most frequent) nouns and noun phrases (Figure 8). The tag cloud generated by the NETLYTIC system provided the frequency of the presented nouns as they occurred in the dataset. This enabled comparing analysis output to AP units developed in Table 1 above; the focus of my evaluation. User defined categories that were considered context specific (Table 1 and 2 above) were used to reclassify AP coded units as "Categories" in NETLYTIC. Doing so supported text analysis regarding concepts pertinent to the objective oriented analyses; an additional focus of this study. This feature proved useful when I looked into analyzed post texts and compared them to AP concepts.

The "Visualize" option follows the "Analyze" step and plotted a stacked graph of terms (see Figure 9).

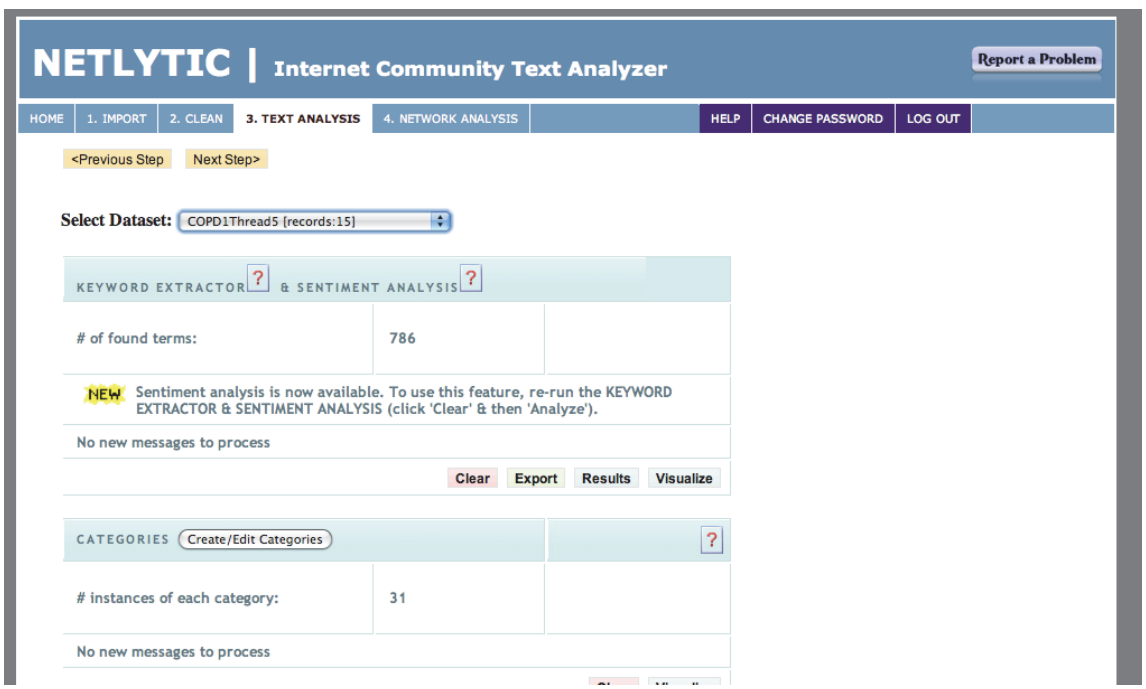


Figure 7 Text Analysis step outputs

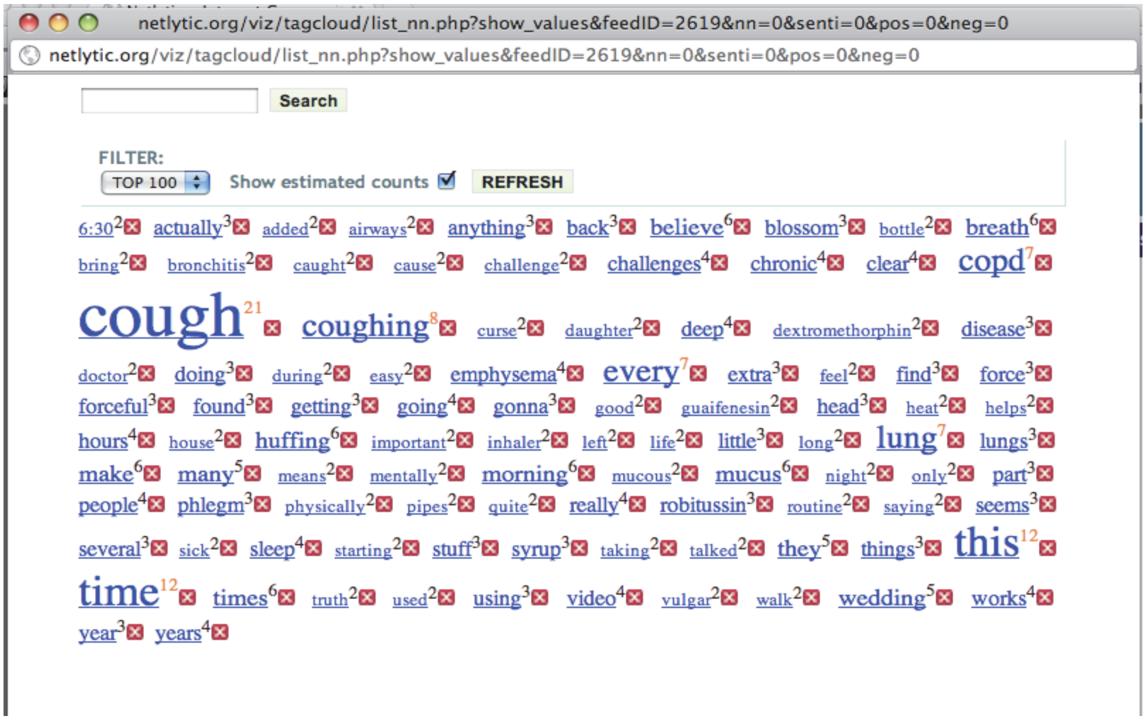


Figure 8 Interactive tag cloud generated in Text Analysis step "Results" before classification

<u>Development of Meta-Matrix data</u>: In this step, the text data were converted into meta-matrix data for quantitative analyses. After analyzing thread contents with the NETLYTIC steps above, the resulting terms were classified as "S" (symptoms) or "A" (actions). These terms are consistent with AP components. For example, explicit terms like "cough" and "short of breath" were classified as "S". A data matrix was developed according to the text categories of symptoms and actions and their order of existence (i.e., actions follow symptoms) in the thread. An example would be if a symptom (S) is discussed in one thread and followed by a suggestion of an action (A). A more detailed example is illustrated in Table 3 below.

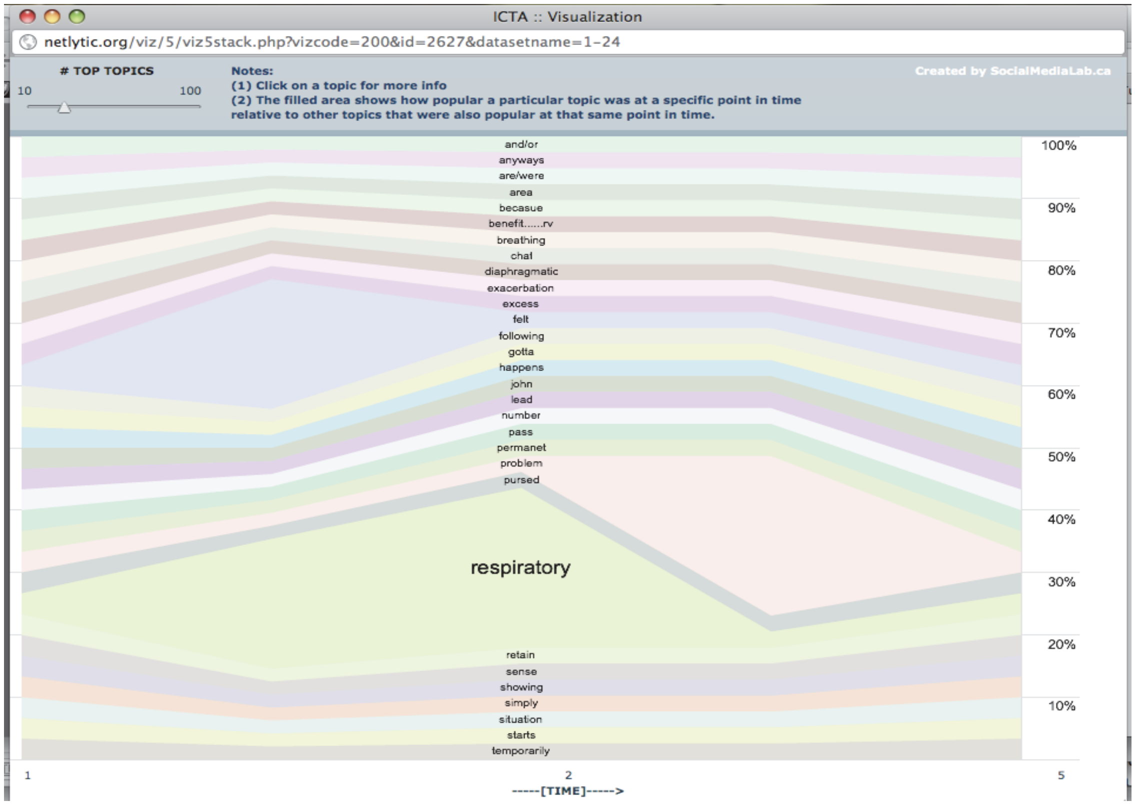


Figure 9 Graphical representation of terms resulting from the Text Analysis step "Visualize"

Summarized automated and manual text analysis findings for all datasets (one example presented above in Figure 2) were utilized for developing the results in Table 3 below.

Table 3 Data Matrix Sample

Thread ID	S1	S2	S..n	A1	A2	A..n
1-1	1	1	1	1	0	0
1-5	1	0	0	0	0	1
1-7	1	0	0	0	0	0

Note: See table 1 and 2 for corresponding symptoms and actions

To ensure validity and establish an accuracy level that would support further analysis, all analyzed terms and concepts were linked to corresponding standardized terminology. Units of AP were also linked to professional terminology standards, such as SNOMED-CT and UMLS. SNOMED CT, or The Systematised Nomenclature of Medicine – Clinical Concepts is considered the world's most extensive healthcare clinical terminology reference and is widely used in clinical data retrieval and analysis [75]. Similarly, the Unified Medical Language System (UMLS) is "the largest thesaurus in the biomedical domain that provides a representation of biomedical knowledge consisting of concepts classified by semantic type and both hierarchical and non-hierarchical relationships among the concepts" [76]. Both SNOMED-CT and UMLS are integrated into health information systems and software such as electronic health records.

In this thesis, SNOMED-CT and UMLS open-source web interfaces were used for classification and use of standardized terminology coding. Nouns and language in patient thread messages were correlated and compared to the Consumer Health Vocabulary (CHV). This process will be discussed further in the next section.

3.3.3 Concepts Mapping – from Consumer Health Vocabulary to Standardized Clinical Terminology

The clinical concepts emulation and matching process was done by using both automated and manual analysis. Shown below (Figure 10) are the principal steps in this process. In the biomedical domain, UMLS and SNOMED CT are the largest knowledge representation resources for clinical data. Programs facilitating text mapping to these biomedical knowledge sources, such as data mining projects, have been successful [77]. Hierarchical and non-hierarchical semantic connections between clinical concepts are well represented in the UMLS Metathesaurus, and applications of automated tools connected to the Metathesaurus have been pivotal in decision support systems, management of patient records, and information retrieval. However, these are projects of a much larger scope than is required for this Master level thesis research.

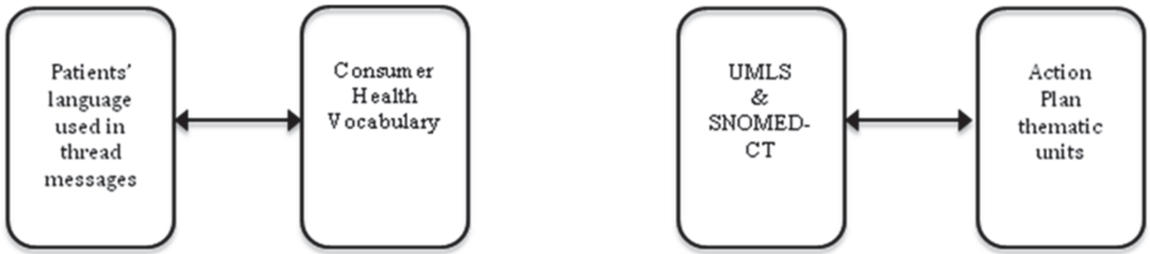


Figure 10 Concept mapping approach

In light of the benefits of involving health professionals in the analysis addressed above, and in consideration of the target population of this study (persons with Asthma and COPD), I combined manual and automated steps when performing concepts mapping. I incorporated SNOMED-CT and UMLS interfaces along with AP thematic units. Following the automated process in text analysis with ICTA (Internet Community Text Analyzer), I employed the main steps illustrated in Figure 11 above. The application of my domain knowledge and my professional experience as a Respiratory Therapist was primarily deployed in coordinating consumer/patient language with CHV.

The next subsection details network assessments of analyzed text with respect to development of name and concept networks.

3.3.4 Development of Social Networks of Health Information Exchange

The “name network” focused on who responded to whom in the network, while the “concept network” assessed the flow of symptoms and actions as a network. In this section, the network structures of Asthma and COPD online communities are described according to identified network literature terminologies discussed below (details of network components are explained elsewhere [61,62]).

Network measures used to analyze social network data are explained next. For each thread network, the *primary* node was set to be the individual who started the thread or initiated the discussion, and the *secondary* nodes were the other members who replied to and commented on the original message posted by the *primary* node. The size of the network was the number of all nodes in a given network, including *primary* and *secondary* nodes, but counting each person once. The *degree* quantified connections between the *primary* node and all other *secondary* nodes within the same network and provided an indication of network strength. Repeated interactions between two nodes were also quantified and were part of the degree. The *size* of the network quantified how frequent and how strong discussions referenced to AP units were. These measures within each thread assessed individual and network compliance (in terms of adherence) with AP units and components. Hence, a quantitative measure of strength of adherence to AP components was provided.

The *closeness* centrality assessed the distance from one actor (or group member) in the network to all others within the network [78]. *Closeness* measures how close one person to others within one group or network with respect to information exchange and transfer. Keeping within our context, we measured the distance among Asthma or COPD group members. The *closeness* centrality in our data represented an information path from one participant to another within a thread conversation. *Betweenness* was another representation of distance within the network. It identified nodes positioned between nodes [78] and also recognized members that act as “brokers” or “gatekeepers” in terms of information diffusion. This measure identified members within conversations in which one participant had to go through another to reach someone else within the same group. The software and programming tools used for network analysis are described next.

One of the features in ICTA (as discussed above and used for text analysis) is that it maintains relationships between members of individual thread messages to enable further relational network analysis. Thread message databases are exported from ICTA into UCINET (open source software used to calculate network measures and delineate networks) format [79].

3.3.5 Statistical Analysis

A data matrix developed following text mining and analysis was used to test our logistic regression model, and to test adherence to AP. The hypothesis we aimed to test regarded the actions and symptoms of the AP document used in this research. Within a network (or thread discussion) in which symptoms were discussed, we wanted to know whether there would be suggestions or references to actions that were part of the AP. Those actions and references that were also part of the AP were coded as one; and any others were coded as zero. The model is as follows:

$$P = e^{\beta o+\beta 1X1+\beta 2X2+....+\beta kXk+\varepsilon}/1 + e^{\beta o+\beta 1X1+\beta 2X2+....+\beta kXk+\varepsilon}$$

where P is the probability of adherence to the action plan, and $X_{1\ldots k}$ are symptoms in an AP document regarded as independent variables. These actions and symptoms are presented in Tables 1 and 2 in section 3.3.1 above. Models were tested for all actions, as permitted by the data collected, and provided a valid model fit [80].

3.4 Summary

This chapter presented a discussion on the screening of available online health-related discussion groups specific to Asthma and COPD. It also presented the development of our research-specific data collection from blogs intended for Asthma and COPD individuals. Various methodology sections discussed the analytical steps taken in this research. Although studies in this context are currently developing, the analytical framework, as illustrated in Figure 6, applied to and referenced the frameworks developed by Weerasinghe [62].

CHAPTER 4 Results and Analysis

4.1 Chapter Outline

In this chapter, results of data processing, text mining and statistical analyses conducted with patient health information are presented according to the methodological strategies discussed in the previous chapter. General findings are addressed with regards to selection of subject blogs related to Asthma and COPD; and processes for recruitment of on-line community members. These observations are in keeping with observations in previous research studies on communications through social networking tools used in health information exchange and found in consumer health informatics literature. This is followed by a section on text mining, and network analysis and results. The chapter's final section presents results of logistic regression statistical analyses.

4.2 General Overview of Asthma and COPD Online Communities and their Social Networking Experience

The research explored potential data sources for this thesis work by recruiting online community members and analyzing observations, as discussed below. The scarcity of published research on this subject merit reporting of these observations, as does the fact that online community data collection methodologies are still evolving.

4.2.1 Subjects Recruitment

As outlined in section 3.2.3 (chapter 3), two separate research-specific web-based discussion boards were developed and launched for Asthma and COPD participants. Neither forum met the intended number of members for discussion participation within the anticipated short period of time for thesis work completion. As the web administrator, I received several requests for membership validations, but the majority of these requests were deemed inadequate for research participants due to lack of qualifications. Many requests were received from non-real users and advertisers. Moreover, discussions among already-approved users were not up to our research expectations. Participant numbers were too small and therefore the motivation for discussions was not as strong as anticipated, compared to existing patient discussion forums.

This could be due to several factors associated with persons involved in virtual communication. In virtual communities, mutual trust, interaction, and reciprocal communication are challenging [80]. There are highly motivated online communities with an established and well-defined knowledge-sharing medium and established social capital built up over time. Our timeframe, and possibly other factors, were insufficient to create the conditions of trust and the habits of communication for us to achieve our objective. However, locating alternative publicly-accessed online Asthma and COPD discussion forums to pursue our research and collect additional data was feasible, through sites that had already built on the increasing popularity of this type of cyber communications on the web, . The next section presents results of content analysis of shared language in threaded conversations from some of these alternative sites.

4.2.2 Preliminary Results of Health Information Content Analysis

Health information dissemination through social media applications and tools has been well defined in the literature [34]. Content analysis processes and classification of main categories in posted messages and comments have been systematically described [34]. These categories include “advice and information”, “to support others”, and “to let others know that I have gone through it, too” [34, 35]. Although these studies did not include blogs from our targeted communities, their categories are similar to those for virtual groups for Asthma and COPD. Medications, inhalers and oxygen administration methods along with advice about breathing techniques can be categorized in the "advice and information" category.

One example of good advice was one in which members suggested that a newly-diagnosed young member quit smoking. In one conversation thread examined, members supported each other on holding to the importance of quitting smoking. Those members who had already quit smoking for some time praised new quitters and assured them that withdrawal symptoms would be for a transitory period of time.

Other examples in the "advice and information" category noted in the dataset included recommendations related to breathing exercise to overcome such problems as panic attacks, shortness of breath at night, and losing weight (particularly in the abdominal region) to reduce muscular effort when breathing. Although the last two examples were suggested without reference to any clinical recommendations (and might even be considered unfavourable for some individuals), members suggested them because they were of benefit to themselves.

There was some potentially harmful advice as well. One example involved advice about inhalers, medications and devices. The advice given was to extend their use beyond the expiry date, which could impose clinical side effects. Another example related to a patient's own invented techniques of using inhalers with "spacer devices".

Further knowledge distribution and sharing elements are discussed in the following sections.

Exploratory assessment of our sample groups' communication showed some challenges for doing text analysis of members' shared language. Following are several examples: (1) In attempting to assess symptom severity and place it on a scale, examples of text used for describing severity included "worst exacerbation" and "extreme worsening". It can be challenging to apply a scale to such vague language. (2) There was some lack of assertion and clarity when expressing symptoms (e.g. "breathing problem" and "asthma problem"). (3) Indications that there was unclear understanding of some of the terms used, such as "exacerbation". These kinds of challenges created difficulty with content analyses.

4.3 Text Mining Results

Using conversations that occurred in the social networks developed by Asthma and COPD online groups, initial thread text mining and assessment were done for all sample data collected. Of these data, 62 % of analysed samples were topics and message threads with relevance to Action Plan (AP) components. This included discussions about various symptoms and recommendations for their management and control as addressed in APs, as well as topics concerned with AP use in particular. An example was: "*What is AP, what are its uses and benefits, and how do I follow its recommendations*?" Another example of an AP topic in the COPD group was "*I have had an action plan for quite a long time… It provides me with a number of scenarios of directions…which empowers me*". The remaining 38% of the messages were topics and threads outside the range of our measures of interest. This included posted links of a published study about COPD or Asthma patients and queries about specific equipment or drug providers. Overall, the majority of patient communications were related to clinical standards specified by AP contents.

4.3.1 Results of Text Mining in Relation to AP Guidelines: Quantity versus Quality

One of the two objectives in this thesis work is to evaluate and assess the extent of knowledge diffusion and exchange within the context of Asthma and COPD APs. In the text evaluation, 60 % (64 out of 106) of the conversations and discussion threads were within the context of AP symptoms and actions. These were discussions about questions, recommendations, and suggestions related to the control and management of disease symptoms. Following data cleaning, relevant messages were retained for further analysis. This analysis included mapping and comparisons with the Consumer Health Vocabularies (CHV) list, as well as with clinical terminology professional standards. The results of these analyses are discussed in the next section.

Using an inductive approach guided by AP-coded theme units (Tables 1 and 2 in Chapter 3), further text processing was carried out. The results are summarized in Table 4 below with ICTA categorizations [74]. The number of instances provided in Table 4 is derived according to the two main AP categories: “symptoms” and “actions”; and the use of “Action Plan” as a phrase. Instances counting as symptoms include *breath, cough, wheeze, pain,* and *anxiety*, but these are only a representative list of a broader collection of terms used by group participants. For a complete list of categorized symptom, see Appendix E. Analysis of derivative strings of instances using *breath* and *cough* resulted the highest count within the category of symptoms reviewed, with counts of 204 and 66 instances, respectively. Actions instances included *increase (medication), use (inhalers), stay away from (triggers),* and *take (medications)*. However, the categorization of "actions" and the process calculation from the text were more complex than that used for "symptoms". Consequently, "actions" did not provide a direct match with AP terms. As a result, "symptoms" constituted the majority of the terms and concepts within Asthma and COPD group discussions.

The precision rate in the "symptoms" category is higher than those in the "actions" category, which is in part due to the language used by patients. The last category listed in Table 4 under the Action Plan is calculated as one phrase and not as a separate category. This item was added as an indictor of references to an AP within both Asthma and COPD groups. These instances did include its applications and AP usages.

Table 4 Frequency Counts in Text Mining

Category	Number of instances
Symptoms	367
Actions	38
Action Plan (as a phrase, not a category)	7

The quality of communication exchange for "symptoms" and "actions", which subsequently reflected knowledge diffused among group members, was evaluated and matched to professional clinical terminology standards (UMLS and SNOMED-CT). CHV was used to validate and qualify semantic components of shared language between participants, and AP-coded units were compared to professional language standards (UMLS and SNOMED-CT). Table 5 below illustrates a representative sample of the process results. Two units only of AP themes ("Short of breath" and "Cough up sputum") are presented in the table below. Their corresponding matches and detailed results are included in Appendix E.

Table 5 Terms mapping to clinical data standards

AP Themes Units	SNOMED-CT	UMLS	CHV	Patients' vocabulary
Short of breath	Breathless	Dyspnea	Shortness of breath	Breath more often
	Mild exertion	Shortness of breath	Breathing difficulty	Insufficient breathing
	Cannot breath deeply enough		Difficulty breathing at night	Shortness of breath
	Moderate exertion			Breathing getting worse
				Breathing becomes laboured
Cough up sputum	Able to cough up sputum	Coughing	Cough	Cough
	Cough			Coughing up any phlegm
	Does cough up sputum			Coughing up any phlegm
	Productive cough			

Prominent precision levels (1.0) and legitimate accuracy rates of concept and phrase comparisons were evident in this process. These rates were present between CHV and patient communication text in words like "confusion". They were also present in AP-themed units and clinical terminology standards like "sputum colour", and among all four dimensions of comparisons (AP themes units, SNOMED-CT/UMLS, CHV, and patients' language). However, further along in the analysis, there was evidence of terms and phrases that were not explicitly matched with terminology standards. These were categorised according to the inferred meaning of their semantics. For example, phrases like "major problem with breathing" and "difficult to breath" were classified in the symptoms category "short of breath". One particularly challenging term was "exacerbation". As discussed above, this term was one of the shared language issues we discovered. It was suitable for classification in multiple units of AP themes like "short of breath", "extremely short of breath", and "agitated", and was highly contextual to the person and the situation discussed.

The standards of health consumer language, used in communications within virtual communities, varied according to the educational backgrounds and individualised disease experiences. However, CHV and controlled clinical terminology compatibility establishment in relation to Asthma and COPD patients' language is beyond the scope of this thesis research. Terminological assessments of web-assisted messages from patients have led to the conclusion that "*the notion of a paradigmatic consumer using vocabulary specific to their consumer status may be ill-founded*" [81]. The text mining of patients' language, and comparisons of that language to CHV and clinical terminology, are guided by our aim to appraise knowledge exchanged in those communities. The process was intended to support evaluation of AP-related knowledge exchanged among those groups.

In conclusion, compared to UMLS, SNOMED-CT was shown to be more comprehensive and detailed when matched to AP units (Appendix E results). For example, the AP unit "short of breath" corresponded to the following results in SNOMED-CT: "breathless, mild exertion, moderate exertion, and cannot breath deeply enough", whereas UMLS results were only "dyspnea" and "shortness of breath". In Figure 1-1, the matching of SNOMED-CT to AP units was measured at 64%, while UMLS achieved only 41%.

In the following sections, results about the dynamic activities and connections between group members that could facilitate and influence information sharing are presented.

4.4 Network Analysis Results

One guiding objective in this thesis is to measure the influence of Web 2.0 social media applications on Asthma and COPD patients, through examination and analysis of consumers/patients' participation in online discussion forums and support groups. Social network analysis of Asthma and COPD blogs is used to further investigate this objective. Quantitative measures are used in this section to examine the strengths and weaknesses of the connections within the community. These measures will also explore and evaluate indicators of how strong and/or weak connections are among members. The community's ability to disseminate information and share their knowledge will likewise be investigated. The measures are discussed in the "name networks" subsection below, followed by a concepts network analysis that focuses on information flow with respect to the AP classified two themes of symptoms and actions.

4.4.1 Name Networks Results

By incorporating features of text analysis and network visualisation in ICTA [74] and the social network analysis program UCINET [79], name network assessment results are presented here. Name network assessment consists of two main components: 1) node discovery (i.e., names and/or members), and 2) tie discovery, which further quantifies the nature of the connections between nodes. An ICTA-executed algorithm identifies names referenced in the messages and supports basic visualisation of the network. UCINET is used for detailed visualisation of these networks as well as for their descriptive measurements (e.g., degree centrality). Network structures and measurements results are presented next.

4.4.1.1 Name Networks Structure

One of threads, “named network information”, is shown in Figure 12 below. In this illustration, members are presented as nodes and the connections among them are presented as directional arrows. This data type is identified as *1-Mode Complete.* Generally, the group as a whole demonstrated cohesive connections, with few isolates spotted outside the community. Furthermore, the individual network structure for one thread conversation consisted of *primary* and *secondary* nodes. As shown in Figure 14 below, a *primary* node represents an individual who posted a message and initiated a conversation thread , while *secondary* nodes represent those who replied to that individual. In general, virtual communities examined illustrated well-connected networks and supporting structure for information exchange. The next section consists of further detailed results of quantitative measures of node activities, associated connections within the network, and impacts on information and knowledge flow.

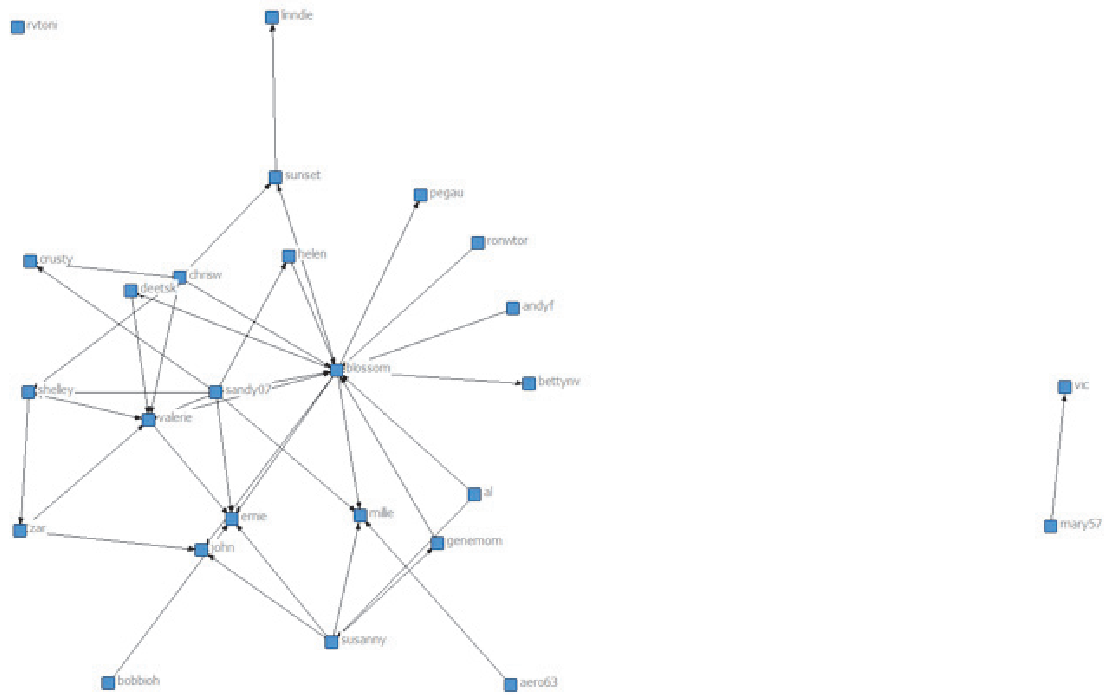


Figure 11 Network data type structure

4.4.1.2 Name Networks Quantitative Measures

Table 6 below contains results of network measures for data collected from both Asthma and COPD groups. The network size assesses the number of people in a given network (conversation thread). Disregarding individuals who are viewing and reading the posted messages without actively participating, a network size reflects the engagement level in the network and the community in topics related to AP components. The average size of a network within these groups is 4, which represents 4 nodes. Thus, within our analysed groups, there were on average four members in any given discussion (an example is shown in Figure 13 below). Conversations that involve one member making a post to which no replies ensue were recorded as a network size of 1. In contrast, highly popular discussions (network size 10) were extreme and their counts were insignificant (less than 15% of overall population samples). As online group interactions are unpredictable, conclusive methodologies examining online groups and their unique characteristics are still a developing area of research [9].

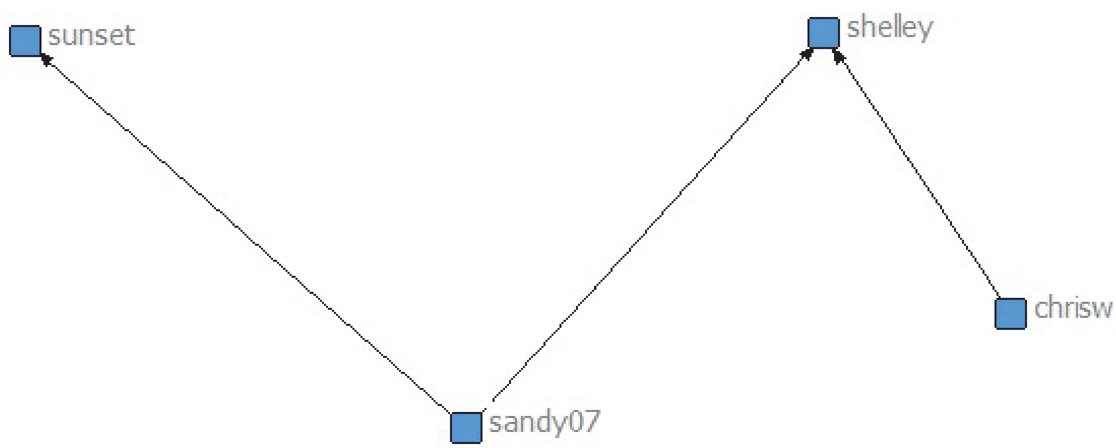


Figure 12 Example of a network of size 4

The degree centrality (C_D) was obtained at two levels: first, between participants in one conversation thread, and second, among all participants contributing in the group as a whole. Degree centrality of an average of ($C_D = 1.2$) is reported from node connections within one discussion network. For degree centrality, a higher number indicates a stronger network and more solid density of node connectivity. At an individual (single node) level, higher values indicated the superiority of that member within the network and stronger connections with all other members. Further, measures greater than 1 suggest the reciprocal nature of reconnections; for every posted message from members, there will be at least one participant in the group that responds to or comments on that message. The second level of degree centrality is among all members of the group. An average of (C_D =3) for each node was found. One candidate blog demonstrated adequate connectivity among its members; 16% of the entire nodes (members) recorded ($C_D \geq 5$). Figure 14 below illustrates an example network with a central node of C_D = 5. Those individuals with higher degree centrality hold superior control over information flow and exchange. Altogether, these two levels of degree centrality results indicated an adequate bidirectional and multidirectional communications between at two or more members within one group. Also, nodes with higher centralities found less dominant and less individual control over information flow within the community.

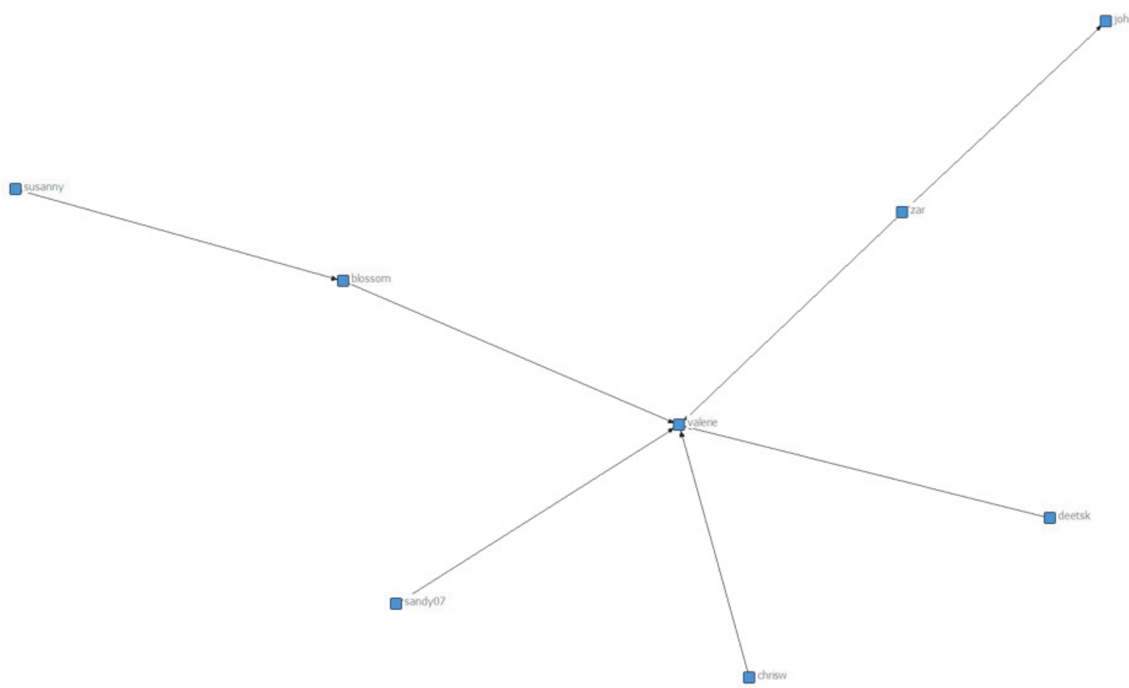


Figure 13 An example of a node with degree centrality C_D =5, C_C =4, and C_B = 19.

Closeness (C_C) is calculated by measuring the extent of closeness between nodes (members) [83]. Of the COPD and Asthma sample collected, C_C = 1.5 was the average closeness measure documented for the majority (64.5%). Nodes with higher C_C values suggest closer connectivity to adjacent nodes, and vice versa. Twenty-six percent of the results were $C_C \geq 2$ and only 9.5 % of the networks were $C_C = 0$. The central node in Figure 14 above represents a high C_C value of 4, suggesting that this node is closest to all other nodes within this particular network. In contrast, Figure 15 below provides an extreme example of a low C_C network where nodes are isolated and not close to each other by any means of communication. The closeness assessment of Asthma and COPD blogs suggested that information and knowledge flow among group members are well facilitated within an average distance (calculated as 2 if the third member responds to the second one etc) for the majority of the networks. Results of alienated networks with low C_C were minimal and their influence on information sharing among the community as a whole can be regarded as insignificant.

The fourth measure of social network analysis examined is betweenness (C_B). More than half of the samples collected (53%) were associated with $C_B = 0$. Since betweenness represents another distance measure within the network that identifies nodes positioned between other nodes in the community, lower C_B values are favourable for unrestricted information and knowledge distribution among members. This can supports two members in the network to communicate directly without a third one in the middle who could exert some influence or control over information flow. Figure 15 below represents the concept of betweenneess in a network with a central node of $C_B = 1$. This value is considered relatively low; as shown in figure 15 below, a node in one end can reach the other end only by going through a single central node only. Information diffusion within a network could be monopolised by members of high C_B values. This was witnessed among only 18% of networks examined with $C_B \geq 5$. Members within those networks did demonstrate a high $C_B = 19$ (the central node in Figure 14 above).

Figure 14 Betweennness representation in a network with central node $C_B = 1$

However, this counted for an insignificant number of members and their impact could be negligible on the overall information and knowledge exchange among members of the whole community.

Collectively, all four measures of social network analysis addressed above suggested that examined blogs were a viable medium for information and knowledge dissemination. Network size clearly showed reciprocal contributions among members within communities. Also, degree assessments indicated low individual influence and control on information flow.

The next section will present results of Asthma and COPD AP component distribution within networks.

Table 6 Social network analysis measures

Measure	**Result**
Degree Centrality (C_D)	1.2
Betweeness Centrality (C_B)	3.1
Closeness Centrality (C_C)	1.5
Network size	4

4.4.2 Concept Networks Results

Guided by AP units, the evaluation of "symptoms" and "actions" (Tables 1 and 2 in Chapter 3) is based on information flow in a conversation thread (network) from symptoms to actions, as directed by AP recommendations. In addition to basic text mining analysis of a single term, the aim is to further validate the viability of knowledge exchange via online social media sites. In the course of the groups' discussions, when the existence of a symptom was reviewed, the intent was to ascertain whether or not it was followed by an action.

Excluding conversation threads/networks that were considered irrelevant to AP symptoms and actions, 71% of the eligible threads from both the Asthma and COPD groups contained at least one symptom and one action. Table 7 below summarized quantities of symptoms (S) and actions (A) found to be in sequence within networks. *Short of breath* (S1) and *use breathing techniques* (A3) were the most frequent "symptom" and "action" in the discussions. They were likewise prominent in their association as well. Specifically, S1 was followed by A3 in 13 discussions (20% of the overall discussions and 92% of A3 instances). The second most prevalent combinations of "symptoms" and "actions" were *Short of breath* (S1), *start medications for flare up* (A1), and *increase oxygen* (A4). This was evident in 7% of the networks examined. The third "symptoms"/"actions" correlation, which constituted between 6 and 3 %, included *cough up sputum* (S2) and *regular cough* (S3); followed by *start medications for flare up* (A1) and *Short of breath* (S1), accompanied by *start antibiotics* (A6).

Further minor evidence suggested that "action" instances like *start prednisone* (A7) and *go to hospital emergency department* (A8) were related to instances with "symptoms" like *Short of breath* (S1), *extremely short of breath* (S5), and *agitated* (S6). Moreover, concept network flow diagrams showed instances of "symptoms" only that were not followed by "actions". Evidently, knowledge exchanged was mostly on "symptoms", with very few suggestions provided for "actions". Indeed, "actions" like *notify contact person* (A5) and (A9) *dial 911* were never suggested. Therefore, it is reasonable to conclude that information exchanged within Asthma and COPD blogs was limited to some common "actions" like *use breathing techniques* (A3) and *start medications for flare up* (A1).

Table 7 Concepts mapping results

AP Themes Unit (Symptoms)	Total	AP Themes Unit (Actions)	Total
S1	38	A1	7
S2	19	A2	4
S3	8	A3	14
S4	11	A4	6
S5	12	A5	0
S6	17	A6	4
S7	2	A7	1
S8	3	A8	2
S9	3	A9	0

4.5 Statistical Analysis

In this analysis, multiple logistic regression models are established to examine relationships between the Asthma and COPD AP components, "symptoms" and "actions". The models are also used to examine the existence of correlations of social capital within computer-mediated communications of the Asthma and COPD groups. Concept networks, discussed above, were used to assess this type of association manually within individual conversations, where statistical inference was used to quantify the significance of the associations.

Multiple logistic regressions were aimed at testing the association between all "symptoms" (S1-S9) and "actions" (A1-A9) separately. Each "action" is modelled as a dependent variable and tested with all "symptoms" as independent variables.

Sample data collected as a result of previous analytical steps in this thesis work was processed through Statistical Analysis System (SAS). Asthma and COPD data was separated and assessed for their adherence to AP using logistic regression. However, findings from these logistic regressins were not statistically significant and therefore the two datasets were combined and results are as follows

Table 8 below summarizes significant findings for models examined. Models tested for A5-9 are not listed in the results table because they did not qualify with any valid regression results due to insufficient observations. Results for A2,3 were not significant. Model (A1) *start medications for flare up* and (A4) *increase oxygen* showed significant results, indicating correlation of their existence with particular symptoms. A1 model's likelihood test statistics $\chi^2 = 13.5$, with 9 df and associated $P < 0.027$, suggest that at least one regression coefficient is significant. This coefficient was associated with (S2) *cough up sputum* and $P = 0.047$. Therefore, discussion messages (threads) among members in our sampled groups with references to "symptoms" related to "cough" (S2) were significantly coupled with "actions" related to "flare up medications"(A1). Similarly, A4 model's likelihood test statistics $\chi^2 = 4.04$, with 1 df and associated P <

0.044, indicated significant correlations of symptoms to A4. Specifically, this connection counted for S3 *regular cough* with P = 0.039; S3 was notably linked to conversations about "higher oxygen requirements" A4.

To conclude, the statistical analysis of informational flow within Asthma and COPD virtual communities resulted in some significant findings. "Cough"-related "symptoms" instances are joined with "actions" concerning medication start and higher oxygen requirements.

A few limitations and challenges were encountered in this statistical analysis, one of which was insufficient sample size.

Table 8 Regression models results

Model	**Odds Ratio Estimates**	**P- Value**
A1	8.1	S2 (0.047)
A4	7	S3 (0.039)

4.6 Summary

The results presented in this chapter focused on contents and their connections in Asthma and COPD blogs and discussion forums. Analyses were carried out according to the framework presented in Chapter 3. Cyber discussions and knowledge exchange were processed using text mining approaches, social network analysis, patient communication content matching with clinical data standards using SNOMED-CT and CVH, and statistical analysis. Consumer health information exchange (patient vocabulary and terms) is different from standard terminology and needs to be mapped using expert knowledge. Network mapping and assessment of connections among Asthma and COPD group members resulted in supportive measures for unrestricted information flow within the community (social network analysis results). This is supportive of online social media application uses for knowledge exchange and self-management. Statistical analyses showed significant connections between symptoms and actions. Therefore, we conclude that unsupervised patient communication via Internet social media can be used as an effective patient information exchange site, when expert advice and guidance is also provided.

CHAPTER 5 Conclusion

5.1 Conclusion

In this thesis, an extensive review of social media tools revealed that blogs are a viable medium for knowledge exchange between Asthma and COPD patients. Although the published literature on literature review methods is abundant, there is scarcity of methods to review web-based information. Therefore, the methodological framework developed in this thesis work for website review (Chapter 3) can be tested and used in other similar research endeavours.

Furthermore, this work integrated core elements of health informatics, including data mining and clinical standards (Action Plan) concept mapping. Results of text mining of message threads, using NETLYTIC, were mapped to CHV and compared to SNOMED CT and UMLS. In addition, health informatics applications of knowledge representation and management strategies were combined in the development of new uses of existing analytical frameworks within this thesis context (Figure 6 in Chapter 3).

While consumer health informatics focuses on empowering users as healthcare consumers through its applications, its primary aim is to lessen the language gap between healthcare professionals and patients. Consumer Health Vocabulary (CHV) is an application of consumer health informatics, which is a component of medical informatics. In this research, a CHV interface, when combined with concept mapping of mined text data with SNOWMED and UMLS, provided a better mapping protocol than using one method or tool alone. Integration occurred after a rigorous assessment of CHV advancement, testing and validation, and applications. A significant portion of the reviewed literature was devoted to a review of these aspects of CHV. Although still under development, the CHV web interface tool added significant value to findings related to the consumer health vocabulary developed in this application.

The findings supported the hypothesis explored in this thesis, that online social media can be considered a viable option for knowledge exchange and self-management among Asthma and COPD patients, using an Action Plan as a benchmark measure in two of the AP components. The network dynamics and connections parameters indicated that knowledge diffusion via blogs is possible. In other words, weblogs provided an efficient mechanism of knowledge exchange among well-connected group members among whom were high reciprocal connections. Also, several AP components provided a strong match with patient/consumer action suggestions, with further details given below. The low congruence could be due to the fact that the patient social media sites were not coordinated by health professionals.

The examined terms and phrases reflected the quality of the use of language by patients/consumers, when compared with CHV and clinical terminology standards. Also, our findings supported significant relationships between the “symptoms” and “actions”, in correspondence with the recommendations given by the Action Plan. Nevertheless, some challenges that could still arise can be addressed with a professional moderator.

5.2 Challenges, Biases and Limitations

One example of language challenges that was found to be a barrier to adherence to AP within the studied online social media applications, is the use of terms and phrases communicated by users. Consumer health vocabulary (CHV) that expressed the extent or degree of symptom severity was challenging, as was mapping with explicit classifications as being either a “symptom” or an “action” . One example of these terms was “exacerbation”. Established medical terminologies recognised this term to effectively describe chronic diseases progression. However, in Asthma and COPD patients’ communications, the scalability of this term was difficult and did not easily lead to an effective following “action”.

One limitation encountered in this work is related to subject recruitment for the research-developed forums. It is possible that time constraints limited our ability to engage required number of participants for each group. Another reason may have been the existing patient networks that expressed conflict of interest in the development of new research-related networks. Another barrier that could hinder the process of implementing AP through publicly-accessed social media was involvement of users with commercial interests. Members were registering to be participants within groups with the intent of advertising commercial goods that were not relevant to our population. This could be recognised as a harmful interaction.

5.3 Study Contribution

The results of this thesis research provide a protocol for the application and testing of existing analytical frameworks that can be adopted for evaluating and examining information and knowledge exchange among disease-specific groups via social networking in Web 2.0. In this work, the focus was on Asthma and COPD Action Plans, driven by a combination of health informatics and consumer informatics. To address the scarcity of analytical frameworks in web searches, we identified enabling elements in social Web 2.0 applications for Asthma and COPD patients to empower symptom control, self-management, and the making of effective health-related decisions.

The study protocol used in this thesis can be applied to research similar social media forums for individuals with other clinical conditions.

5.4 Future Directions

Whether from a clinical, practical, or technical perspective, this work speaks to a wide variety of domains. Moreover, although Asthma and COPD online social media users were the specific focus here, the analysis process can be applied in numerous fields. The results of this study invite future research in consumer health informatics and consumer health vocabulary studies to conquer, reduce, and bridge the information gap between health professionals and patients. This can be accomplished by analysing communicated language among users and patients, describing their symptoms, and discussing their disease management process. The enabling elements suggested by network analysis findings within groups foster the promotion of new techniques and treatment strategies for Asthma and COPD patients, bolstered by robust health professional involvement and participation.

BIBLIOGRAPHY

[1] Chapman KR, Bourbeau J, Rance L. The burden of COPD in Canada: results from the Confronting COPD survey. *Respiratory medicine*. 2003;97 Suppl C(2003):S23-31. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12647940.

[2] Kamel Boulos MN, Wheeler S. The emerging Web 2.0 social software: an enabling suite of sociable technologies in health and health care education. *Health information and libraries journal*. 2007;24(1):2-23. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17331140.

[3] Braman S, Global Burden of Asthma. *Chest.* 2006;130:4S-12S.

[4] Gibson PG. Written action plans for asthma: an evidence-based review of the key components. *Thorax*. 2004;59(2):94-99. Available at: http://thorax.bmj.com/cgi/doi/10.1136/thorax.2003.011858 [Accessed June 20, 2011].

[5] Wood-baker R, Mcglone S, Venn A, Walters EH. Written action plans in chronic obstructive pulmonary disease increase appropriate treatment for acute exacerbations. *Respirology*. 2006:619-626.

[6] Gallefoss F, Bakke PS. Impact of patient education and self-management on morbidity in asthmatics and patients with chronic obstructive pulmonary disease. *Respiratory medicine*. 2000;94(3):279-87. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10783940 [Accessed April 21, 2011].

[7] Smyth JM, Stone a a, Hurewitz a, Kaell a. Effects of writing about stressful experiences on symptom reduction in patients with asthma or rheumatoid arthritis: a randomized trial. *JAMA : the journal of the American Medical Association*. 1999;281(14):1304-9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10208146.

[8] McLean R, Richards BH, Wardman JI. The effect of Web 2.0 on the future of medical practice and education: Darwikinian evolution or folksonomic revolution? *The Medical journal of Australia*. 2007;187(3):174-7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17680746.

[9] White M, Dorman SM. Receiving social support online: implications for health education. *Health education research*. 2001;16(6):693-707. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11780708.

[10] Gavrin JR. Living in the electronic age: musings on nearly two decades in Cyberspace. *Journal of pain & palliative care pharmacotherapy*. 2009;23(2):141-

4. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19492214 [Accessed June 23, 2011].

[11] Thackeray R, Neiger BL, Hanson CL, McKenzie JF. Enhancing promotional strategies within social marketing programs: use of Web 2.0 social media. *Health promotion practice*. 2008;9(4):338-43. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18936268 [Accessed June 13, 2011].

[12] SA. Khan, DJ. McFarlane, J. Li, JS. Ancker , C. Hutchinson, A. Cohall,R. Kukafka. Healthy Harlem: empowering health consumers through social networking, tailoring and web 2.0 technologies. *AMIA Annual Symposium proceedings / AMIA Symposium. AMIA Symposium*. 2007;16:1007. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18694106.

[13] JD. Allen, AM. Stoddard, G.Sorensen . Do social network characteristics predict mammography screening practices? *Health education & behavior : the official publication of the Society for Public Health Education*. 2008;35(6):763-76. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2859725&tool=pmcentrez&rendertype=abstract [Accessed December 6, 2010].

[14] M. Krauss, N. Mueller, D. Luke. Interorganizational relationships within state tobacco control networks: a social network analysis. *Preventing chronic disease*. 2004;1(4):A08. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1277948&tool=pmcentrez&rendertype=abstract.

[15] DO. Erwin, VA. Johnson, M. Trevino, K. Duke, L. Feliciano, L. Jandorf. A comparison of African American and Latina social networks as indicators for culturally tailoring a breast and cervical cancer education intervention. *Cancer*. 2007;109(2 Suppl):368-77. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17173279 [Accessed June 23, 2011].

[16] B. Bolam, C. McLean, A. Pennington, P. Gillies. Using new media to build social capital for health: a qualitative process evaluation study of participation in the CityNet project. *Journal of health psychology*. 2006;11(2):297-308. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16464926 [Accessed October 9, 2010].

[17] A.Shachak, AR. Jadad. Electronic Health Records in the Age of Social Networks and Global Telecommunications. *JAMA*. 2010;303(5):452-453.

[18] G. Eysenbach. (2008). Medicine 2.0: Social Networking, Collaboration, Participation, Apomediation, and Openness. *Journal of Medical Internet Research*.

[19] J. Preece & D. Maloney-Krichmar. (2006). Online Communities: Design, Theory, and Practice. *Computer Mediated Communication , 10* (4).

[20] J.L. Bender, M C. Jimenez-Marroquin & A.R. Jadad. (2011). Seeking Support on Facebook: A Content Analysis of Breast Cancer Groups. *Journal of Medical Internet Research , 13*.

[21] GG. Bennett, RE. Glasgow. The delivery of public health interventions via the Internet: actualizing their potential. *Annual review of public health*. 2009;30:273-92. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19296777.

[22] LM. Rasmussen, K. Phanareth, H. Nolte, V. Backer. Internet-based monitoring of asthma: A long-term, randomized clinical study of 300 asthmatic subjects. *Asthma diagnosis and treatment.* 2005; 1137-1142.

[23] Weerasinghe,S. "Social Networks of Health Information Exchange" PowerPoint presentation. Health Informatics Seminar,Feb.28,2011.

[24] JC. Trappenburg, EM. Monninkhof, J. Bourbeau, T. Troosters, AJ. Schrijvers, TJ. Verheij, JW. Lammers. Effect of an action plan with ongoing support by a case manager on exacerbation-related outcome in patients with COPD: a multicentre randomised controlled trial. *Thorax* (2011): 977-984.

[25] EW. Bischoff, DH. Hamd, M. Sedeno, A. Benedetti, TR. Schermer, S. Bernard, F. Maltais, J. Bourbeau. Effects of written action plan adherence on COPD exacerbation recovery. *Thorax* (2011): 26-31.

[26] N. Ring, R. Jepson, G. Hoskins, C. Wilson, H. Pinnock, A. Sheikh, S. Wyke. Understanding what helps of hinders asthma action plan use: asystematic review and synthesis of the qualitative literature. *patient education and counsling* (2011): e131-e143.

[27] Robin Masion & Frank Rennie. E-Learning and Social Networking Handbook Resources for higher eduaction. Taylor & Francis e-Library, 2008.

[28] Hughes, Benjamin, Indra Joshi and Jonathan Wareham. Health 2.0 and Medicine 2.0: Tensions and Controversies in the Field. *Journal of Medical Internet Research* 10.3 (2008).

[29] Kamel Boulos, Maged N, Inocencio Maramba and Steve Wheeler. Wikis, blogs and podcasts: a new generation of Web-based tools for virtual collaborative clinical practice and education. *BMC Medical Education* (2006).

[30] Hesse, Bradford W., Nelson DE, Kreps GL, Croyle RT, Arora NK, Rimer BK, Viswanath K. Trust and Sources of Health Information The Impact of the Internet and Its Implications for Health Care Providers: Findings From the First

Health Information National Trends Survey. *Archives of Internal Medicine* (2005): 2618-2624.

[31] Hesse, Bradford W., O'Connell M, Augustson EM, Chou WY, Shaikh AR, Rutten LJ. Realizing the Promise of Web 2.0: Engaging Community Intelligence. *J Health Commun*. (2011): 10-31.

[32] P.Klemm, M. Hurst, SL. Dearholt, SR. Trone. Cyber Solace: Gender Differences on Internet Cancer Support Groups . *Computers in Nursing* (1999): 65-72.

[33] Finn, Jerry. Computer-based self-help groups: a new resource to supplement support groups. *Health and Social Work* (1995): 109-117.

[34] D. Scanfeld, V. Scanfeld, E. L. Larson. Dissemination of health information through social networks: Twitter and antibiotics. *the Association for Professionals in Infection Control and Epidemiology* 38 (2010): 182-188.

[35] C. Ridings, D. Grefen. Virtual Community Attraction: Why People Hang Out Online. *Journal of Computer-Mediated communication* (2004).

[36] Sharf, Barbara F. Communicating Breast Cancer On-Line: Support and Empowerment on the Internet. *Women and Health* (1997): 65-84.

[37] White, Marsha and Steve M. Dorman. Receiving social support online: implications for health education. *Health Education Research* (2001): 693-707.

[38] Eysenbach, G., Powell, J., Englesakis, M., Rizo, C., & Stern, A. (2004). Health related virtual communities and electronic support groups: systematic review of the effects of online peer to peer interactions. *BMJ , 328*.

[39] Cline, R., & Haynes, K. (2001). Health related virtual communities and electronic support groups: systematic review of the effects of online peer to peer interactions. *Health education research , 16* (6), 671-692.

[40] Eysenbach, Gunther. "Credibility of Health Information and Digital Media: New Perspectives and Implications for Youth." Digital Media, Youth, and Credibility. Edited by Miriam J. Metzger and Andrew J. Flanagin. The John D. and Catherine T. MacArthur Foundation Series on Digital Media and Learning. Cambridge, MA: The MIT Press, 2008. 123–154. doi: 10.1162/dmal.9780262562324.123

[41] MORAHAN-MARTIN, J. M. (2004). How Internet Users Find, Evaluate, and Use Online Health Information: ACross-Cultural Review. *CyberPsychology & Behavior , 7* (5), 497-510.

[42] Sillencea, E., Briggsa, P., Harrisb, P. R., & Fishwicka, L. (2007). How do patients evaluate and make use of online health information? *Social Science & Medicine , 64*, 1853-1862.

[43] Ciriello, James N. and Nalin Kulatilaka. Smart Health Community: The Hidden Value of Health Information Exchange . *American Journal of Managed Care* 16 (2010): 31-36.

[44] Houston, Thomas K. and Heidi E. Ehrenberger. The Potential of Consumer Health Informatics. *Seminars in Oneology Nursing* (2001): 41-47.

[45] Lancaster, Tim and Lindsay F Stead. Self-help interventions for smoking cessation." Cochrane Database of Systematic Reviews 2 (2009).

[46] Bouhaddoul, Omar, Joseph G. Lambert and Suzanne Miller. Consumer Health Informatics: Knowledge Engineering and Evaluation Studies of Medical HouseCall. *AMIA* (1998): 612-616.

[47] Eysenbach, Gunther. "Recent advances Consumer health informatics." *BMJ* 320 (2000): 1713-1716.

[48] Eysenbach, Gunther and Alejandro R Jadad. "Evidence-based Patient Choice and Consumer health informatics in the Internet age." *Journal of Medical Internet Research* (2001).

[49] Yaqub, Edwin and Andre Barroso. " Distributed Guidelines (DiG): A Software Framework for Extending Automated Health Decision Support to the General Population ." *Perspectives in Health Information Management* (2010).

[50] Frost, Jeana H and Michael P Massagli. "Social Uses of Personal Health Information Within PatientsLikeMe, an Online Patient Community: What Can Happen When Patients Have Access to One Another's Data." *Journal of Medical Internet Research* (2008).

[51] Swan, Melanie. "Emerging Patient-Driven Health Care Models: An Examination of Health Social Networks, Consumer Personalized Medicine and Quantified Self-Tracking." *International Journal of Environmental Research and Public Health* (2009): 492-525.

[52] P.Wicks, M. Massagli, J. Frost, C. Brownstein, S. Okun, T. Vaughan, R. Bradley, J.Heywood. "Sharing Health Data for Better Outcomes on PatientsLikeMe." *Journal of Medical Internet Research* (2010).

[53] A. McCray, T., R.F. Loane, A. C. Browne, A. K. Bangalore. "Terminology Issues in User Access to Web-based Medical Information." *AMIA* (1999): 107-111.

[54] Q.Zeng and T. Tse. "Exploring and Developing Consumer Health Vocabularies." *JAMIA* (2006): 24-29.

[55] Smith, Catherine Arnott and Paul J. Wicks. "PatientsLikeMe: Consumer Health Vocabulary as a Folksonomy." AMIA 2008 Symposium Proceedings. *AMIA*, 2008. 682-686.

[56] Q. Zeng, E. Kim, J. Crowell, T. Tse . "A Text Corpora-Based Estimation of the Familiarity of Health Terminology." *Biological and Medical Data Analysis , COMPUTER SCIENCE* (2005): 184-192.

[57] Q. Zeng, T. Tse, J. Crowell, G. Divita, L. Roth, A. Browne. "Identifying Consumer-Friendly Display (CFD) Names for Health Concepts." *AMIA Annu Symp Proc*. 2005. 859-863.

[58] A. Keselman, T. Tse, J. Crowell, A. Browne, L. Ngo, Q. Zeng. "Assessing Consumer Health Vocabulary Familiarity: An Exploratory Study." *LMIR* 9 (2007).

[59] Q. Zeng, T. Tse, G. Divita, A. Keselman, J. Crowell, A. Browne, S. Goryachev, L. Ngo. "Term Identification Methods for Consumer Health Vocabulary Development." *JMIR* (2007).

[60] Doing-Harris, Kristina M and Qing Zeng-Treitler. "Computer-Assisted Update of a Consumer Health Vocabulary Through Mining of Social Network Data." *JMIR* 13 (2011).

[61] Weerasinghe, S. Social network pathways of health information diffusion: Concepts, models and mixed method analytical frameworks, paper submitted to Health Informatics Conference, annual meeting of NIH, Ottawa, Canada, April 2012.

[62] Weerasinghe**, S.** *Social Networks in Health: An Overview of Conceptual and Analytical Frameworks using Case Studies obtained from Atlantic Canadian Immigrants,*In "Networking and Immigrant Integration: Atlantic Canada and Beyond", UBC press, Accepted May 21, 2010

[63] Boulos MNK, Maramba I, Wheeler S. Wikis, blogs and podcasts: a new generation of Web-based tools for virtual collaborative clinical practice and education. *BMC medical education*. 2006;6:41. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1564136&tool=pmcentrez&rendertype=abstract [Accessed July 12, 2010].

[64] *COPD in Canada Message Board/Forum open to the public- for patients exchange info with your peers*. Retrieved Oct-Feb 2011-2012 from copdcanada.ca: http://philns.proboards.com/index.cgi?

[65] Canada, A. S. *National Asthma Patietn Alliance disscussion board*. Retrieved Oct/Aprl 2011/2012, from http://www.asthma.ca/napa/db/viewforum.php?f=9&sid=62c40cb84543d8c29d1fb5b81a1ce9b3

[66] Patientslikeme. *Forum/lungs and respiratory*. Retrieved Oct-Aprl 2011/2012, from http://www.patientslikeme.com/forum/lungs_respiratory/topics

[67] Microsoft. (2008). Microsoft Excel [computer software]. Redmond, Washington: Microsoft.

[68] A. Alaujan. (2012, January 15) COPD Forum [Online]. Available copd.patientnetwork.ca

[69] A. Alaujan. (2012, January 15) Asthma Forum [Online]. Available asthma.patientnetwork.ca

[70] Chiu, C.-M., Hsu, M.-H., & Wang, E. T. (2006). Understanding knowledge sharing in virtual communities: An integration of social capital and social cognitive theories. *ScienceDirect* , 1872-1888.

[71] Boyatzis, R. E. (1998). *Transforming Qualitative Information: Thematic Analysis and Code Development.* Thousand Oaks, California, United States of America: SAGE.

[72] Widman, L. E., & Tong, D. A. (1997). Requests for Medical Advice From Patients and Families to Health Care Providers Who Publish on the World Wide Web. *Arxh Intern Med , 157*, 209-212.

[73] Smith, C. A., Stavri, P. Z., & Chapman, W. W. In Their Own Words? A Terminological Analysis of E-mail to a Cancer Information Service. *AMIA 2002 Annual Symposium Proceedings* (pp. 697-701). *AMIA*.

[74] Gruzd, A. Exploring Virtual Communities with the Internet Community Text Analyzer (ICTA). In B. K. Daniel, *Methods and Techniques for studying Virtual Communties Paradigms and Phenomena* (Vol. I, pp. 209-217). New York, USA: Information Science Reference.

[75] Benson, T. (2010). SNOMED CT. In T. Benson, *Principles of Health Interoperability HL7 and SNOMED.* Springer London.

[76] Aronson, A. R. (2001). Effective Mapping of Biomedical Text to the UMLS Metathesaurus: The MetaMap Program. *AMIA* , 17-21.

[77] Brennana, P. F., & Aronson, A. R. (2003). Towards linking patients and clinical information: detecting UMLS concepts in e-mail. *Journal of Biomedical Informatics* , 334-341.

[78] Liu, E. Z., Shih, R. C., & Tsai, Y. L. (2011). Hyperlink network analysis of the educational blog. *British Journal of Educational Technology , 42*, E25-E29.

[79] Borgatti, S.P., Everett, M.G. and Freeman, L.C. 2002. *Ucinet for Windows: Software for Social Network Analysis*. Harvard, MA: Analytic Technologies.

[80] D' Agostino Sr., R. B., Sullivan, L. M., & Beiser, A. S. (2006). *Introductory Applied Biostatistics.* Belmont, CA, United States of America: BROOKS/COLE CENGAGE Learning.

[81] Poon, N., & Daniel, B. K. Challenges of Analysing Informal Virtual Communities. In B. K. Daniel, *Methods and Techniques for Studying Virtual Coomunities Paradigms and Phenomena* (Vol. II, pp. 589-592). New York, USA: Information Science Reference.

APPENDIX A COPD Action Plan template document

Canadian Respiratory Guidelines

PLAN OF ACTION FOR ____________________
(patient's name)

Patient's copy

I FEEL WELL

MY SYMPTOMS

- I feel short of breath: ____________________
- I cough up sputum daily. No Yes, colour: ____________
- I cough regularly. No Yes

I FEEL WORSE

MY SYMPTOMS

- **I have changes in my sputum (colour, volume, consistency),** not only in the morning
- **I have more shortness of breath than usual**

Note that these changes may happen after a cold or flu-like illness and/or sore throat.
Some people feel a change in mood, fatigue or low energy prior to a flare-up.

MY ACTIONS

- I use my **prescription for COPD flare up**
- I avoid things that make my symptoms worse
- I use my breathing, relaxation, body position and energy conservation techniques
- If I am already on Oxygen, I use it consistently and increase from ___ L/min to ___ L/min
- I notify my contact person _________ (Tel:____________) and/or see my doctor (Tel:____________)

PRESCRIPTION FOR COPD FLARE-UP

1) If your SPUTUM becomes yellowish/greenish

start Antibiotic ____________ **Dose:**___ **#pills:**___ **Frequency:**___ **#days:**___

If repeating antibiotics within 3 months, use the following antibiotic instead

start Antibiotic ____________ **Dose:**___ **#pills:**___ **Frequency:**___ **#days:**___

2) If you are more SHORT OF BREATH than usual, **take** ___ **puffs of** _________ **up to a maximum of** ___ **times per day, as necessary**

If your SHORTNESS OF BREATH DOES NOT IMPROVE,

start PREDNISONE ____________ **Dose:**___ **# pills:**___ **Frequency:**___ **# days:**___

____________	____________	____________	____________
Physician Name	Signature	License	Date

I FEEL MUCH WORSE OR IN DANGER

MY SYMPTOMS	MY ACTIONS
• My symptoms have worsened. • After 48 hours of treatment my symptoms are not better.	• I notify my contact person and/or see my doctor • After 5 pm or on the weekend, I go to the hospital emergency department (Tel:____________)
• I am extremely short of breath, agitated, confused and/or drowsy, and/or I have chest pain	• **I dial 911 for an ambulance to take me to the hospital emergency department.**

APPENDIX B   Asthma Action Plan template document

Respirology Clinic
Asthma Teaching Plan

Early Warning Signs Indicating
That Your Asthma is out of Control

---

1. Shortness of breath, coughing, wheezing or tightness in the chest that awakens you from sleep more than once a week.

2. Prolonged coughing, with or without phlegm that is usually not present. Can be due to nasal symptoms as well.

3. Asthma symptoms, cough, wheezing, chest tightness more than twice a week.
4. Inhalers do not provide relief for the full 4 hours.

5. You need to use your reliever puffer __________________more than _______ times per week / day,
or,

6. If you have a peak flow meter and your numbers drop by _______% or to _________ L/min.

Action Plan

When your asthma is out of control, DO:

1. **Look for and, if you can, remove yourself from the cause, triggers, e.g. animals, dust, moulds, grass, pollens, etc.).**

2. **a)** **Increase your ________________ to _____ puffs _____ times a day.**
   a. **Increase your ________________ to _____ puffs _____ times a day.**
   b. **ADD in your ________________ to _____ puffs _____ times a day.**

1.

If you **are not better** or **are worse** after _______ hours,

i. Take ______ tablets prednisone/medrol ______ times a day,

ii. Hold / continue your inhaled steroids e.g. Flovent, Pulmicort, Qvar etc

---

iii. Call Dr. _______________ for further instructions.

OR:

IF YOU ARE GETTING BETTER IN **24/48** HOURS, BUT NOT ALL BETTER

iv. Continue _______________ __ puffs __ times per day until your symptoms have disappeared and your peak flows have returned to 80 % of your best or _____L/min for _______ days in a row.

v. Then decrease your ______________ to ____ puffs ____ X per day for _______ days
Use your _____________ bronchodilator on an as needed basis only, hopefully only 1 – 3 times per week

vi. If still no symptoms for a further ____ days decrease your __________ to ____ puffs ____ X per day

vii. If still no symptoms decrease your _____________ to ____ puffs ________ X per day

viii. Stop / Decrease your prednisone on day _____ if no symptoms.
Start / continue your _______ inhaled steroid at ____ puffs ___ times per day (the day before you are to stop your prednisone).

ix. Decrease prednisone by ______ pill every ____ d

APPENDIX C Blogs Screening Results

	Subject			Participant			Demographic			Facebook©/ Twitter ©	Asthma/ COPD
	Guided subject posts	Free posts	Q & A	Only patients	Only professionals	Both patients & professionals	Canada	Global	Link		
1	X					X		X US	http://blog.patientslikeme.com/		
2		X				X		X	http://www.dailystrength.org/support-groups		X
3		X		X				X	http://network.patient.org.in/		X
4								X	http://www.ginasthma.com/index.asp?l1=1&l2=0	X	X
5									http://www.goldcopd.com/WCDIndex.asp?l1=1&l2=0		X
6									http://cnrchome.net/index.html		
7									http://www.lunghealthframework.ca/home		

8									http://www.nf.lung.ca/ physical groups not online		
9									http://www.on.lung.ca/ Facebook© wall for anouncements	F	
10									http://www.nb.lung.ca/ very few patient participated in F. wall	X	
11									http://www.pq.lung.ca/ very few patient's posts in F wall	F	
12									http://www.pei.lung.ca/ Both T&F walls for organization's announcements	X	
13									http://www.sk.lung.ca/ organization's announcements	X	
14									http://www.ab.lung.ca/ organization's announcements	X	
15									http://www.mb.lung.ca/ very few patient participated	X	

									in F. wall		
16									http://www.lung.ca/ home-accueil_e.php organization's announcements	X	
17		X		X			X		http://www.asthma.ca/adults/ founded below http://napa-blog.blogspot.com/ National Asthma Patient Alliance	T	
18									http://www.patient-experience.com /?s=asthma&submit.x=0&submit .y=0&submit=Search http://www.patient- experience.com/index.php /asthma-inhalers-blog/	X	
19		X		X				X UK	http://talk.nhs.uk/blogs /asthma/default.aspx http://talk.n	X	X

									hs.uk/blogs /copd/default.aspx people's opinions		
20		X		X	X			X US	http://e-patients.net/archives /category/patient-networks people's opinions	X	
21		X	X	X			X	X US	http://www.carepages .com/forums		
22		X			X			?	http://www.wsdha.org /category/health-care		
23		X		X	X			X UK	http://casesblog.blogspot. com/search/label/Pulmonology		
24									http://www.blogs.com/topten/top-10-best-clinical-medicine-blogs/ EDUCATIONAL FOR PROFESIONALS		
25		X		X			x		http://www.copdblog.com /default.aspx		X

APPENDIX D   Demographic Questions

(To be filled electronically by participant upon online web registration)

1. Have you been diagnosed with Asthma or COPD?
2. Are you taking care of someone who has been diagnosed with Asthma or COPD?
3. Do you consider the symptoms to be controlled or uncontrolled?
4. Age of participant (please choose one age group between 18-25, 26-30, 31-35, 36-40, 41-50, 51- 65 or more).
5. Gender of participant (please choose from male or female).
6. Country of residence of participant (please type in the name of your country).
7. Level of education of participant (please choose from: secondary, post-secondary, less than secondary).
8. Employment status of participant (please choose from: employed, health professional, unemployed, retired).
9. If a health professional, please provide more details about your speciality.

APPENDIX E Concepts Mapping Results

patietns' used terms	CHV	UMLS	SNOMED CT	AP-THEMES
breath more often	shortness of breath	dyspnea	breathless(F)	short of breath
insuffieicent breathing	breathing difficulty	shortness of breath	mild exertion	
flare up	difficulty breathing at night		moderate exertion	
breathing becomes labored			cannot breath deeply enough (F)	
exacerbation			difficulty breathing (F)	
breathing getting worse			difficulty taking deep breath (F)	
major problem with her her breathing			ineffective breathing pattern (F)	
exacerbations			irregular breathing (F)	
difficult to breathe			shallow breathing (F)	
cough	cough		able to cough up sputum (F)	cough up sputum
coughing up any phlegm			cough (F)	
cough up phlegm			does cough up sputum (F)	
			productive cough (F)	
			ability to cough up sputum(O)	
cough	cough	coughing	night cough present (S)	regular cough
			chronich cough (F)	
			cough (F)	
			dry cough (F)	
			non productive cough (F)	
caugh yellow	color of sputum	color of sputum	productive cough-clear sputum (F)	sputum color, volume, consistency
yellow phlegm			productive cough- green sputum (F)	

yucky mucus			productive cough- yellow sputum (F)	
lots and lots of mucus				
exacerbation			breathless- strenuous exertion (F)	extremely short of breath
would not get my breath back			catching breath (F)	
unable to catch my breathe			gasping for breath (F)	
impossible to breathe at al			increasing breathlessness (F)	
unbearable to breath			increasing breathlessness (F)	
bad exacerbation			labored breathing (F)	
extreme SOB			unable to take deep breaths (F)	
anxiety	agitation	agitation	feeling agitated (F)/	agitated
breathing faster				
exacerbations				
anxiety when I couldn't breath				
confusion	confusion	confusion	on examination- mentally confused (D)*	confused
	sleepiness	drowsness	drowsy (F)	drowsy
chest pain			chest pain on breathing (F)	have chest pain
pain ocassionally in my upper lung/chest				
fast acting INHALER helps to calm down the cough			start medications for COPD flare up	
Increase your puffs/take antibiotics for COPD flare up				
use my rescue inhaler				

stay away from sick people/				avoid symptom aggravation
pursed lip breathing	physical exercise	exercise	active cycle of breathing techniques(T)	use breathing, relaxing, body positions, energy conservation techniques
proper breathing techniques			breathing exercise,blow bottel (T)	
tak big deep breath…			breathing phonatory exercise (T)	
hold your breath.			deep breathing and coughing exercise (T)	
breath slower			deep breathing exercise (T)	
try blowing through a straw			diaphragmatic breathing exercise (T)	
			encouragement of deep breathing and coughing exercise(P)	
			relaxed breathing (T)	
			high concentration oxygen therapy (P)/home oxygen therapy(P)/oxygen therapy(P)/ increased oxygen (as F not P!!)	increase oxygen
				notify contcat person
take an antibiotic				start antibiotics
				start prednisone
				go to hospital emergency department

				dial 911