Patient Centered Healthcare Informatics

Christopher C. Yang

***Abstract*—The healthcare system is undergoing a transformation from reactive care to proactive and preventive care. Patients or health consumers are actively acquiring knowledge to manage their health and seeking supports from their peers in addition to receiving healthcare support from healthcare professionals. 74% of American adults use the internet, of which 80% have looked online for healthcare information [6]. With the popularity of social media, many health consumers are also exchanging informational and emotional support with peers who have similar health conditions or diseases. The large volume of health consumer contributed content provides valuable resources for healthcare informatics research. It is worth to note that the information in the health consumer contributed content is timelier than the traditional resources such as electronic health records, centralized reporting systems, and pharmaceutical databases because health consumers often discuss their concerns with peers before any of them are reported in the traditional resources [22]. In this article, we review a few important healthcare informatics research issues that are centered on the patient contributed content and concerns.**

***Index Terms* — Social media analytics, healthcare informatics, consumer health vocabulary, social support, drug safety signal detection, topic detection, recommendation systems.**

I. Introduction

HEALTH and wellbeing plays an important role in our societies. Improving the health and wellbeing of people is a main goal accomplished through both government and private healthcare organizations [21]. In the recent years, we have also observed the increasing effort of health consumers or patients in managing their own health proactively and preventively. Health consumers are actively educating themselves about health and wellness in order to maintain a healthy body or prevent diseases. Patients are going to Internet to acquire knowledge about their health conditions or treatments by identifying authoritative information from popular health web sites such as WebMD and PubMed. When the resources are limited, health consumers and patients are also going to social media sites such as MedHelp and PatientsLikeMe to seek and offer supports with their peers who have similar health conditions or diseases [23]. Many patients are sharing experiences with their peers and offering advice and opinions to support one another. In this article, we are focusing on five specific issues: (a) consumer health expressions, (b) social support, (c) community topic detection and recommendation systems, (d) drug safety signal detection, and (e) symptom profiling and clustering.

Christopher C. Yang is with in College of Computing and Informatics at Drexel University, Philadelphia, PA 19104, USA, Email: chris.yang@drexel.edu

II. Consumer Health Expressions

Despite the fact that patients and health consumers are actively seeking and exchanging healthcare information on the Internet, identifying the relevant and useful information is very challenging to most patients and health consumers. It is because health consumers and health professionals often use different vocabularies to express health related topics [10,29,30,31]. While health professionals are trained to use professional language, which can be easily identified from the healthcare professional ontologies such as UMLS and MeSH, to describe the health issues, health consumers use a variation of vocabularies to express their health concerns depending on their cultural, educational, social, and economic backgrounds. The language gap creates a huge barrier between the communications of health consumers and health professionals as well as between the communications of health consumers with substantially different backgrounds. For example, a patient experiencing nose bleeding may not be able to find relevant authoritative information when the scientific publications are using the professional term “epistaxis” to describe the symptom. Similarly, some patients may express the symptom as nose bleeding while some others may express it as bloody nose. The variation of expressions adopted by health consumers makes it difficult to communicate and search information online simply by keyword matching.

Many researchers have devoted to develop Consumer Health Vocabularies (CHVs) to capture the expressions used by health consumers and map these vocabularies to healthcare professional ontologies. Zeng et al. has developed the first generation of CHV [31]. However, a substantial amount of manual effort is required in their effort. In addition, the consumer health expressions are evolving from time to time. As a result, CHV needs to be maintained continuously in order to capture new expressions that have not been included in CHVs yet.

In the recent years, more efforts have been made to utilize the health consumer contributed content in social media to identify the new consumer health expression semi-automatically or automatically. Jiang and Yang have utilized co-occurrence analysis to extract consumer health expressions by expanding to the original CHV [7,8,9]. Co-occurrence analysis is used because two or more words that tend to occur in similar linguistic context tend to resemble each other in meaning. In the co-occurrence analysis, we find that most health consumer expressions are bi-grams. By using the expanded consumer health expressions, we can extract up to ten times more relevant online discussion threads on a particular health issue such as the adverse drug reaction heart disease. By mapping the expanded consumer health expressions to UMLS, health consumers can identify authoritative information more effectively or health professionals or researchers can extract the patient concerns from health consumer contributed content in a timely manner. The extracted content through using expanded consumer health expressions are also useful for many knowledge discovery applications such as drug safety signal detection, symptom analysis, topic detection and recommendation systems, which will be discussed in the later sections.

III. Social Support

Social support has been proven to be important in healthcare intervention. Social media provides a platform for health consumers to make connection with others without time and geographical constraints. Hence the limitations of the traditional social support groups that meet regularly at dedicated locations can be removed to support a broader community. In our previous study, it is found that a substantial amount of informational support and nurturant support are found in healthcare social media such as MedHelp and QuitNet [1,2,3,4].

Informational support offered by health consumers provides information related to treatment or coping with diseases. The information includes advice to cope with situations, referrals to other resources, facts that reassessing situations, opinions on issues but not necessarily based on facts, and personal experiences. In a previous study, we found that most informational supports are found in online discussion forum setting [1]. Among all different types of informational support, personal experiences are the most popular, followed by advice and opinions [1]. Referrals and facts are relatively less popular.

Nurturant supports are expressions that show signs of listening, expressing sympathy or the importance of relationship. There are three major types nurturant supports, including (a) esteem support that gives positive comments to validate the recipient's self-concept and alleviate feelings, (b) emotional support that gives expressions to support the recipient's feeling or reciprocates emotion, and (c) networking support that focuses on connecting recipients to others with similar situation to broaden social networks. More nurturant supports are found in private settings such as the journal section of the personal profile pages, where health consumers often discuss their own health status [2]. Among all types of nurturant support, emotional support is the most popular followed by networking support and esteem support [2].

The interaction patterns in informational support and nurturant support are different [33,35]. Health consumers with health status in the later stages tend to offer informational support to other health consumers with health status in the earlier stage. For example, in the QuitNet forum, health consumers who have quitted smoking for a longer time tend to offer informational support to health consumers who have just started to quit or have quitted for a relatively shorter time. However, health consumers tend to offer nurturant support to other health consumers who have similar health status.

IV. Community Topic Detection and Recommendation Systems

The discussions in healthcare social media sites are valuable resources to discover the timely patient concerns. We have applied dynamic stochastic blockmodeling and temporal Dirichlet process to detect hidden communities [12]. Such detection model is able to detect the evolving communities since the discussion groups may expand, shrink, split, or merge as the discussions are going on. By monitoring the emerging topics and evolving communities, it is helpful to capture the issues raised in social media [15]; and hence, makes helpful recommendations [13], provides timely support to health consumers and identifies new research problems. ACTONNECT is a web-based search engine that aims to enable patients, clinicians, researchers, and others to conduct searches of health information gleaned from dozens of patient forums and social media sites and share their results graphically [18]. It has received the first place conceptual model in the PCORI Patient-Research Matching Challenge in 2013

Although there is a large number of users in healthcare social media sites, the online social networks are usually sparse. Each user may only interact with a limited number of peers while missing many other peers who have common interests or healthcare concerns. As a result, health consumers are often not connected to those peers who may offer them the best information or the nurturant support they need. Through understanding the user intent and the social support types the health consumers are involved (as discussed in the previous section), we are able to match health consumers with one another to enrich their interactions in healthcare social media sites. In light of this, we investigate an automatic process of classifying user intent and social support types with the human annotated content as the training data set [33,34]. In the classifier, we adopt content analysis and health status as features. The result is promising and it shows that the classification performance can be improved when health status are adopted.

By analyzing the interaction patterns, we have also proposed the UserRank algorithm to rank the user influence in healthcare social media sites [14]. The health consumers who are most active in social media are not necessarily the most influential [20]. Instead, the influence is a measure of how much impact a health consumer has made to the community. By identifying the influential users as well as the explicit and implicit relationships [19], we can utilize the healthcare social network to disseminate the timely and important information to the target users.

V. Drug Safety Signal Detection

Drug safety signal detection is important in postmarketing drug safety surveillance because many potential adverse drug reactions cannot be identified in premarketing review process. 5% of hospital admissions are attributed to adverse drug reactions and many deaths are reported every year. Current drug safety detection techniques relies heavily on resources such as centralized reporting systems, electronic health records, and pharmaceutical databases. However, there is a high under-reporting ratio in the centralized reporting system such as FDA Adverse Event Report System (FAERS) due to the nature of passiveness. Many adverse drug experiences reported by the health consumers are not necessarily recorded in electronic health records by the health professionals unless sophisticated evaluations are made. In the recent years, there is an increasing effort of detecting the drug safety signals using social media as the resources.

Yang et al. have adopted the expanded health consumer expressions (as discussed in Section II) to discover the discussions on adverse drug reactions on social media sites [24]. By extending the previous effort, Yang et al. have developed association mining [25] and heterogeneous network mining techniques [27] to detect the adverse drug reactions of particular drugs [17] and to detect the drug-drug interactions of any given two drugs [28]. Not only social media data is promising in detecting drug safety signals, they have also conducted temporal analysis and found that the techniques can detect the adverse drug reactions earlier than FDA alters by several years [26]. It can be explained by the fact that health consumers are actively discussing the adverse drug experiences on social media sites before any traditional resources have records of such adverse drug reactions. The heterogeneous network mining techniques also indicate the meaningful paths that involve users, drugs, adverse drug reactions and diseases, which are helpful to present the relationships of drug-drug interactions. For example, some drug-drug interactions cannot be observed by their direct relationships but the interactions can be detected when two drugs are prescribed to patients who have multiple diseases. The heterogeneous network mining is also potential for investigating drug repositioning or off-label use of drugs.

VI. Symptom Profiling and Clustering

There have been many clinical longitudinal studies trying to understand how symptoms are developing over time and how symptoms are correlated. In particular, in cancer treatments, a symptom cluster is defined as three or more concurrent and related symptoms frequently found in patients. Symptom clustering is drawing attention in the recent years. It is because co-existing symptoms may share a common underlying etiology [5,10]. For example, biomarkers such as serum cortisol, melatonin, and serotonin are all related to a cluster of symptoms including fatigue, sleep, and depressive moods during chemotherapy. Examining co-existing symptoms is more efficient and effective than coping with symptoms one by one. It is found that understanding the co-variation in symptoms is helpful in the discovery of physiological mechanisms that lead to the manifestation of disease and side effects of treatment. Previous studies also suggest that intervention improves multiple symptoms concurrently. As a result, there are both clinical and physiological interests in studying symptom clustering.

Clinical studies usually require a lot of effort in recruiting subjects and the same group of subjects may not always be available for a longitudinal study due to the time and geographical constraints. Social media data is an alternative source for symptom clustering. Yang et al. have recently conducted a comparative study of symptom clustering on clinical and social media data for breast cancer [16]. In the study, it is found that there is a substantial agreement between the results derived from the social media data and from the clinical study data. However, there are also some significant discrepancies. In general, we find that there are a couple of clusters with a large number of symptoms and there are also clusters with only one single symptom when the clinical data is used. It can be explained by the fact how the data is collected. In the clinical study, each subject was given a long list of symptoms and was asked to check the symptoms that each had experienced. In such case, the subjects were able to examine the symptoms one by one and checked all those that they had experienced regardless if they had serious concerns on the checked symptoms. On the other hand, the users in healthcare social media sites were voluntarily discussing the symptoms that they concerned. As a result, general symptoms were not discussed as frequent in social media. Clusters of symptoms can be easily identified and the symptoms are more evenly distributed to the clusters when social media is used. The highly correlated symptoms are grouped into the same clusters. In the future, we are also interested to investigate the symptom profiles of patients and how they are correlated.

VII. Conclusion

In this article, we have discussed five emerging issues of patient-centered healthcare informatics research by harnessing healthcare social media. Health consumer expressions are essential to understand the concepts that health consumers are concerning. It is a continuous effort in expanding the health consumer expressions as the vocabularies used in social media are evolving. Social support, which helps to engage users interactions, plays an important role in healthcare social media sites. The underlying social network are useful in understanding the interaction patterns and identifying the influential users; and hence valuable for disseminating timely and important healthcare information. Not only the health consumer expressions are evolving, the hidden communities in healthcare social media sites are evolving. By capturing the dynamic communities, we are able to understand the evolving issues raised by health consumers. Recommendations can also be made effectively to the health consumers in order to enrich the user interactions. Social media data is also useful in knowledge discovery applications such as drug safety signal detection and symptom profiling and clustering. It can supplement the traditional resources such as centralized reporting system, electronic health records, and clinical and pharmaceutical databases. There are also many opportunities of harnessing the social media platforms in patient-centered healthcare informatics research that have not been explored yet. By integrating healthcare sensor data and mobile applications with social media data, a large volume of healthcare data can be collected and more sophisticated analysis can be developed to understand the impact of medical treatments and medications on patients' health conditions. Patient-centered healthcare management system can also be developed to support health consumers in managing their own health and wellbeing. Health consumers are becoming more proactive and preventive. They want to be equipped with knowledge and personalized data analytics to make their own healthcare decisions. As we continue in these efforts, a smart and connected health era may not be too far away.

REFERENCES

[1] K. Chuang and C. C. Yang, “Informational Support Exchanges on Different Computer-Mediated Communication Formats in a Social Media Community of Alcoholism,” *Journal of the American Society for Information Science and Technology*, vol.65, no.1, 2014, pp.37-52.

[2] K. Chuang and C. C. Yang, “Interaction Patterns of Nurturant Support Exchanged in Online Health Social Networking,” *Journal of Medical Internet Research*, vol. 14, no.3, 2012.

[3] K. Chuang and C. C. Yang, “A Study of Social Positions in an Online Alcoholism Community,” *Proceedings of International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction*, Washington, D.C., April 2 – 5, 2013.

[4] K. Chuang and C. C. Yang, “A Study of Informational Support Exchanges in MedHelp Alcoholism Community,” *Proceedings of International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction*, College Park, MD, April 3-5, 2012.

[5] M. Dodd, S. Janson, N. Facione et al.m “Advancing the Science of Symptom Management,” *J Adv irs*. 33, 668-676, 2001.

[6] S. Fox. (2011, October 17). *The Social Life of Health Information*. Available: http://pewinternet.org/Reports/2011/Social-Life-of-Health-Info/Summary-of-Findings.aspx

[7] L. Jiang and C. C. Yang, “Expanding Consumer Health Vocabularies by Learning Consumer Health Expressions from Online Health Social Media,” *Proceedings of International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction 2015*, Washington, D. C., April 1 - 3, 2015.

[8] L. Jiang and C. C. Yang, “Using Co-occurrence Analysis to Expand Consumer Health Vocabularies from Social Media Data,” *Proceedings of IEEE International Conference on Healthcare Informatics*, Philadelphia, PA, September 8 - 11, 2013.

[9] L. Jiang, C. C. Yang, J. Li, “Discovering Consumer Health Expressions from Consumer-Contributed Content,” *Proceedings of International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction*, Washington, D.C., April 2 – 5, 2013.

[10] T. B. Patrick, H. K. Monga, M. E. Sievert, J. Houston Hall, and D. R. Longo, "Evaluation of controlled vocabulary resources for development of a consumer entry vocabulary for diabetes," Journal of medical Internet research, vol. 3, p. E24, 2001.

[11] J. K. Payne, B. F. Piper, I. Rabinowitz, M. B. Zimmerman, “Biomarkers, Fatigue, Sleep, and Depressive Symptoms in Women With Breast Cancer: a Pilot Study,” *Oncology Nursing Forum*, 33, 775–783, 2006.

[12] X. Tang and C. C. Yang, “Detecting Social Media Hidden Communities using Dynamic Stochastic Blockmodel with Temporal Dirichlet Process,” *ACM Transactions on Intelligent Systems and Technology*, accepted for publication.

[13] X. Tang, M. Zhang, and C. C. Yang, “Leveraging User Interest to Improve Thread Recommendation in Online Forum,” *Proceedings of International Conference on Social Intelligence and Technology*, State College, PA, May 8 – 9, 2013.

[14] X. Tang and C. C. Yang, “Ranking User Influence in Healthcare Social Media,” *ACM Transactions on Intelligent Systems and Technology*, vol.3, no.4, 2012.

[15] X. Tang, M. Zhang, and C. C. Yang, “User Interest and Topic Detection for Personalized Recommendation,” *Proceedings of IEEE/WIC/ACM International Conference on Web Intelligence*, Macau, December 4-7, 2012.

[16] C. C. Yang, E. Ip, N. Avis, Q. Ping, and L. Jiang, “A Comparative Study of Symptom Clustering on Clinical and Social Media Data,” *Proceedings of International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction 2015*, Washington, D. C., April 1 - 3, 2015

[17] C. C. Yang, H. Yang, and L. Jiang, “Postmarketing Drug Safety Surveillance using Publicly Available Health Consumer Contributed Content in Social Media,” *ACM Transactions on Management Information Systems*, vol.5, no.1, April, 2014.

[18] C. C. Yang, S. Lin, M. C. Kim, L. Jiang, “ACTONNECT: A Platform to Support Patients and Researchers Collaboration,” *Proceedings of IEEE International Conference on Healthcare Informatics 2014 (ICHI 2014)*, Verona, Italy, September 15-17, 2014.

[19] C. C. Yang, X. Tang, H. Yang, and L. Jiang, “Identifying Implicit and Explicit Relationships in Social Commerce Activities,” *International Journal of Electronic Commerce*, vol.18, no.2, Winter, 2013-2014, pp.73-96.

[20] C. C. Yang and X. Tang, “Estimating User Influence in the MedHelp Social Network,” *IEEE Intelligent Systems*, vol.27, no.5, 2012, pp.44-50.

[21] C. C. Yang, G. Leroy, and S. Ananiadou, “Smart Health and Wellbeing,” *ACM Transactions on Management Information Systems*, vol.4, no.4, 2013.

[22] C. C. Yang, R. Chiu, S. Lin, and A. Kumar, “Patient-Centered Research and Social Media,” *IEEE Life Sciences Newsletter*, 2013.

[23] C. C. Yang, J. Yen, and J. Liu, “Social Intelligence and Technology,” *IEEE Intelligenct System*, vol.29, no.2, March-April, 2014.

[24] C. C. Yang, L. Jiang, H. Yang, and X. Tang, "Detecting Signals of Adverse Drug Reactions from Health Consumer Contributed Content in Social Media," *Proceedings of ACM SIGKDD Workshop on Health Informatics*, Beijing, 2012.

[25] C. C. Yang, H. Yang, L. Jiang, and M. Zhang, "Social media mining for drug safety signal detection," *Proceedings of the 2012 International Workshop on Smart Health and Wellbeing*, 2012, pp. 33-40.

[26] H. Yang and C. C. Yang, “Using Health Consumer Contributed Data to Detect Adverse Drug Reactions by Association Mining with Temporal Analysis,” *ACM Transactions on Intelligent Systems and Technology*, accepted for publication.

[27] H. Yang and C. C. Yang, “Drug-Drug Interactions Detection from Online Heterogeneous Healthcare Network,” *Proceedings of IEEE International Conference on Healthcare Informatics 2014 (ICHI 2014)*, Verona, Italy, September 15-17, 2014.

[28] H. Yang and C. C. Yang, “Harnessing Social Media for Drug-Drug Interactions Detection,” *Proceedings of IEEE International Conference on Healthcare Informatics*, Philadelphia, PA, September 8 - 11, 2013.

[29] Q. Zeng, S. Kogan, N. Ash, R. A. Greenes, and A. A. Boxwala, "Characteristics of consumer terminology for health information retrieval," Methods Inf Med, vol. 41, pp. 289-298, 2002.

[30] Q. T. Zeng and T. Tse, "Exploring and developing consumer health vocabularies," Journal of the American Medical Informatics Association : JAMIA, vol. 13, pp. 24-29, 2006.

[31] Q. T. Zeng, T. Tse, G. Divita, A. Keselman, J. Crowell, and A. C. Browne, "Exploring Lexical Forms: First-Generation Consumer Health Vocabularies," AMIA Annual Symposium Proceedings, pp. 1155-1155, 2006.

[32] M. Zhang and C. C. Yang, “Using Content and Network Analysis to Understand the Social Support Exchange Patterns and User Behaviors of Online Smoking Cessation Intervention Program” *Journal of the*

American Society for Information Science and Technology, accepted for publication.

[33] M. Zhang and C. C. Yang, "Classifying User Intention and Social Support Types in Online Healthcare Discussions," *Proceedings of IEEE International Conference on Healthcare Informatics 2014 (ICHI 2014)*, Verona, Italy, September 15-17, 2014.

[34] M. Zhang and C. C. Yang, "Classification of Online Health Discussions with Text and Health Features Sets," *Proceedings of AAAI International Workshop on the World Wide Web and Public Health Intelligence 2014 (W3PHI 2014)*, Quebec City, Canada, July 27, 2014.

[35] M. Zhang and C. C. Yang, "Social Support and Exchange Patterns in an Online Smoking Cessation Intervention Program," *Proceedings of IEEE International Conference on Healthcare Informatics*, Philadelphia, PA, September 8 - 11, 2013.