

## Comparing Information Needs of Diabetes Patients in Chinese and American Health Communities of Questions and Answers

Jing Shi<sup>a</sup>, Yuxing Qian<sup>a</sup>, Chenlu Li<sup>a</sup>, Liqin Zhou<sup>a</sup>, Bin Zhang<sup>b</sup>

<sup>a</sup> Center for Studies of Information Resources, Wuhan University, Wuhan, Hubei, China,

<sup>b</sup> Center of Traditional Chinese Cultural Studies, Wuhan University, Wuhan, Hubei, China

### Abstract

*This study selects the users generated content in diabetes groups of ManYouBang and DailyStrength which are two representative HCQA (Health Communities of Questions and Answers). Theme coding was applied to identify information needs while social network analysis was used to compare Chinese and American HCQA. In theory, we combine Social Network Analysis and content analysis to capture and contrast the pattern of Q&A in HCQA. On the one hand, the core needs of HCQA are "how to treat the disease" and the "diet" theme is closely related to other themes. On the other hand, the Chinese diabetes communities have "question and answer attributes" while the American diabetes communities have both "question and answer attributes" and "social support attributes." In practice, the founding provides enlightenment for the development of HCQA including improving content quality and strengthen emotional connection and disease treatment infrastructure.*

### Keywords:

Diabetes Mellitus; Diet; Social Support

### Introduction

Health information refers to information related to people's physical and mental diseases, nutrition, health care, etc. [1] Health Information Seeking Behavior (HISB) is described as "verbal or nonverbal behavior used to obtain, clarify or confirm knowledge or information about a specific event or situation" [2]. Recently, most HISB researches have adopted questionnaires and structured interview methods to conduct quantitative analysis [3]. Themes cover searching behaviors, motivations, purposes, influence factors, information quality assessment, impact on health practice, and so on [5]. Many people, either adolescents or older adults, reported that the Internet was an important channel to reach health information source available [4]. HISB can be affected by external and internal factors, such as uses and gratifications (U&G), Media System Dependency (MSD) [6], user's goals, motivations, emotions and so on [7]. Betsch [8] studied online vaccine information's impact on patient's vaccine receiving behaviors based on TPB theory. It appears information plays an important role in patient's vaccine receiving behaviors, especially for lack-of-knowledge individuals.

HCQA is different from users' online health information browsing. It is characterized by two main functions: informational support and social support. With the number of people with chronic diseases skyrocketing, more and more users joined HCQA to seek and share treatment experience and health knowledge. They offer views, conduct emotional communication, and seek emotional support from others in the

community. It is considered to be critical in patient's recovery [9]. Researches analyze more based on encoding topics including content analysis and quality evaluation [10,11]; user behaviors and psychological analysis [12, 13]; topic identification [14-16].

Kim et al. [10] and Oh et al. [11] studied the content of forums in Yahoo! Answers. The former analyzes linguistic properties of different types of questions about eating disorders while the latter investigates evaluation criteria people use about the online health information. Factor and regression analyses conducted among 231 breast cancer patients were used to examine the role of social support perception and emotional wellbeing on online information seeking within the context of CHES [12]. Ahadzadeh [13] used PLS-SEM to test the moderating effect of health consciousness on attitude towards internet usage in HISB. Coulson [14] and Arden [15] captured three main themes of alcohol discussion forums and guidance on public forums respectively by qualitatively thematic analysis. The current body of research has shed light on the HCQA related landscape. Notwithstanding, researches using hybrid methods on user online health information need are limited. Only a few studies focused on understanding information needs. Min [16] used both content and semantic network to analyze forums which revealed two dominant topics: breast cancer tests and treatment.

Few researchers have applied Social Network Analysis combined with content analysis to capture and contrast the pattern of Q&A in HCQA. To fill the gap, this study focuses on understanding the patterns and users' needs of HISB in Chinese and American communities, taking diabetes for example.

### Methods

Data were obtained from diabetes-patient-group in DailyStrength (<https://www.dailystrength.org/>) and Man Youbang (<http://www.manyoubang.com/>), one of the largest online communities in the USA and China. DailyStrength was established in 2006 comprised of 421 patients-groups and 1096 members in diabetes-group; while Man Youbang was established in 2013 comprised of 84 patients-groups and 7,995 members in diabetes-group. Although Man Youbang also raises money for patients, we only focus on its Q&A which is a relatively separate and important section in fact.

Users post messages and replies on a variety of topics referring to diabetes. A total of 3,553 replies associated with 777 questions from Man Youbang and 11,754 replies associated with 2,115 questions from DailyStrength related to diabetes posted in Jan. 1st, 2013 to Dec. 31st, 2017 were collected. Through manual review, meaningless entries are deleted. Then, we get 2,556 replies associated with 777 English questions fr

Table 1 – Topic Coding Scheme of Dataset

Level 3	Level 2	Level 1	Connotation
Disease specific information	Prevention	How to prevent diabetes	How to prevent diabetes from genetic, environmental, diet, exercise, and other factors.
	Diagnose	Symptom	Anything that accompanies diabetes and is regarded as an indication of diabetes.
		Cause	Events that lead to diabetes.
	State of an illness	Typing	Diabetes is divided into type 1 and type 2
		Describe the state of an illness.	Description of the patient's condition.
	Treatment	insulin therapy	Treatment with insulin injection.
	Complication	Other treatments	Other treatments, such as oral medication.
		Diabetic foot	A diabetic foot is a foot that exhibits any pathology that results directly from diabetes.
		Ketoacidosis	Acidosis with an accumulation of ketone bodies; occurs primarily in diabetes mellitus.
		Retinopathy	The most important manifestation of diabetic microangiopathy is specific changes in the fundus.
		Blood glucose monitoring	Patients observe and record their blood glucose by instruments, test paper, etc.
		Special cases	Pregnant women develop diabetes before or during pregnancy.
		Juvenile diabetes	Patients are diagnosed with diabetes between the ages of 13-25.
		Senile diabetes	Patients are diagnosed with diabetes after age 60.
Non-disease specific information	Drug's side effects	Others	Patients have other diseases besides diabetes.
		Side effects caused by drugs	Symptoms, such as nausea, vomiting, hair loss and fatigue, caused by drugs.
		Dietary precautions	Food that helps or deteriorates with treatment.
		Exercise to lose weight	Exercise during illness to lose weight.
	Insurance	Medical insurance	Obtain compensation or security for treatment.
		Other insurance	Other insurance.
	Finance	Financial concerns	Financial pressure during treatment.
	Emotional support	Spiritual support	Seek comfort and encouragement during treatment.
	Social support	Charity relief	Property and other charitable assistance during treatment.

Man Youbang and 9,179 replies associated with 1,640 questions from DailyStrength. In this study, we explore hot topics and the evolution in different periods by coding content to reflect users' information need.

Manual annotation and word frequency statistics are often used to summarize the characteristics of information needs. Topics are identified manually by two coders considering the coding scheme adapted by Oh, Zhang, and Park [17]. Table 1 shows the coding results. 50 English and 50 Chinese questions and associated replies were selected randomly to test inter-rater reliability. For the level 2 categories, Cohen's kappa scores were 0.892 and 0.913 for the English questions and Chinese questions respectively. For the level 1 categories, Cohen's kappa score were 0.812 and 0.791 for them. It indicated that the two coders achieved good agreement on categories.

## Results

### Hot Topics Comparison of Chinese and American HCQA

Topics in domestic and foreign have similar performance. The most popular topics are treatment, diet, and state of illness. It is surprising that "community health service" is mentioned merely as an important auxiliary treatment for chronic diseases. It reflects that patients still pay more attention to disease-specific information. Most of them expect to obtain practical and reliable treatment and health management information, which can guide medical treatment and drug selection.

We also found some differences. To make numbers comparable, we calculate the proportion of each topic instead of absolute value. In Man Youbang, participants sought information about treatment most frequently, followed by diet, state of illness, complication, and diagnose, which reflects obvious practicability. Especially, "treatment" accounted for more than a quarter of all topics, which was the top priority for patients in Chinese communities. The least were prevention, insurance and, social support. Regarding DailyStrength, information need focuses on treatment, state of illness, and diet. Beyond the disease-specific information, an important non-disease specific term, emotional support, accounts for 15% in DailyStrength. It appears that American users pay great attention to the impact of emotion and gain comfort and encouragement from others in the community. Figure 1 present the top 5 hot topics.

Comparing the two communities, it becomes apparent that the proportion of "emotional support" in DailyStrength is about twice that of Man Youbang. Considering the content of relevant posts, American users intend to use emotional words more frequently out of religious belief, humanistic concern, and so on. Most questioners not only describe conditions but show their moods. At the same time, answers deliver their concern, sympathy, and support. It reflects the fact that DailyStrength not only possesses the attribute of "Q&A" but the attribute of "social contact." Patients express emotions and seek support frequently because they trust and depend on other members. It becomes apparent that western countries pay attention to emotional support in addition to drugs and physical ways.

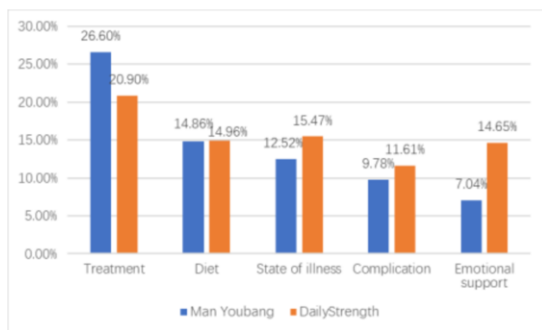


Figure 1— Comparison of the top 5 hot topics in HCQA

### Evolution of Hot Topics

Combined with the time label, Figure 2 and Figure 3 show the topics' composition of diabetes forums' in 2013-2017.

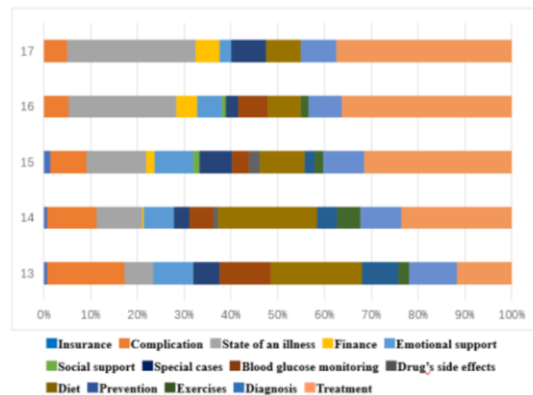


Figure 2— Theme Tags of Man Youbang

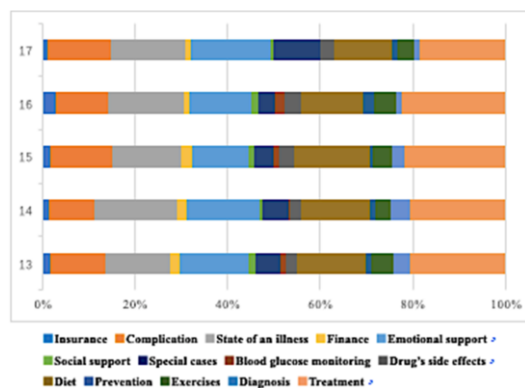


Figure 3— Theme Tags of DailyStrength

First, the overall trend of topics evolution is different. Annual topics' composition changes obviously in Man Youbang while not in DailyStrength. It is related to the development of HCQA in Chinese and American. Founded in 2006, DailyStrength is at a stable stage now. Users of various conditions and courses are balanced. By contrast, the heats have a larger fluctuation with members in different stages altering in Man Youbang.

Second, specific topics have different trends. In Man Youbang, the proportion of "treatment," "state of illness," and "finance" increase while "complication" and "diet" decrease. "Emotional

support" and "diagnosis" are stable. By contrast, topics proportion is more balanced in DailyStrength. "Complication" usually posted by severe patients while "treatment" and "state of illness" may be issued by newcomers. In China, changing topics' proportion reflects Diabetes incidence and treatment burden are growing.

### Comparison of Co-occurrence Network of Topics

Figure 4 shows the co-occurrence networks of Man Youbang and DailyStrength, which demonstrates the overall patterns and holistic properties. The larger the node, the higher the betweenness centrality. It shows that these topics are more likely to co-occurrence with others. The thicker the line, the higher the co-occurrence frequency between two topics.

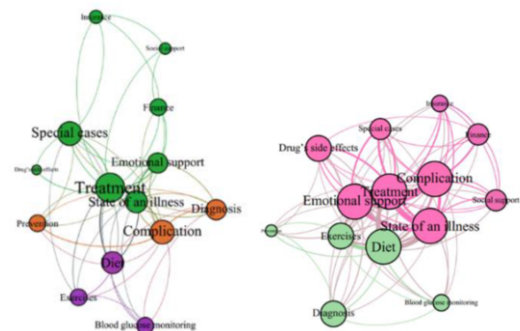


Figure 4— Co-occurrence Network of Topics in Man Youbang(left) and DailyStrength(right)

Figure 4 indicates the heats distribution is similar. We also focus on detailed information on different topics.

First, the top degree centrality topics are "treatment," "special cases," "state of illness," and "diet" in both Man Youbang and DailyStrength. All of them appear alongside more than 10 themes. "Treatment" is the most critical theme. Even if the main purpose of the question is not "treatment," the related narrative emerge. Second, we found that although "diet" does not co-occur with many topics, its average co-occurrence rate is high. It indicates "diet" has a closer correlation with its co-occurrence.

Diabetes is closely related to dietary habits for improper diet leading to the fluctuation of blood sugar. Therefore, patients discuss whether certain food can be eaten and how to control blood sugar. Food therapy is also used in the treatment of diabetes, which explains why "diet" is frequently associated with "treatment" and "state of illness." Also, obesity is one of the causes, and losing weight is a hot topic for diabetic patients. As two major ways to lose weight, "exercise" and "diet" are connected closely.

On the other hand, there are also significant differences. First, the overall network attributes are different. Topics in DailyStrength are connected more closely than Man Youbang. The co-occurrence network of Man Youbang includes 14 nodes and 50 lines. Average degree is 7.143, that is each topic co-occurs with other 7 topics on average. The density is 0.549, and average weighting degree is 48.571. Overall, the topic themes are not closely connected. The co-occurrence network of DailyStrength includes 14 nodes and 77 lines. Average degree is 11. The density is 0.846, and average weighting degree is 208.571. Overall, these themes are closely connected.

Second, topics play different roles. The high-degree nodes in DailyStrength are more than that in Man Youbang. In DailyStrength, nodes with a degree centrality above 10 accounts for 79% where "complication," "state of illness,"

"emotional support," "diet," and "treatment" co-occur with all other topics. We found that DailyStrength has a distinct role in "sharing." Patients ask and answer questions, also share their conditions, moods, and even matters in life. This increases the number of words and the probability of topics occurrence. While obvious attribute of "practice" in Man Youbang, patients ask and answer with short words involving fewer topics.

In addition, members have different emotional needs. As shown in Figure 4, "emotional support" not only co-occurs with all topics but appear a higher average co-occurrence rate In DailyStrength. Diabetes has a long treatment period accompanied by serious complications, which may cause fatigue and negative emotions. Diabetes is closely related to living habits, especially diet and exercise. Most patients have to reduce living quality and take on a greater economic burden.

At last, Topic clustering of Man Youbang is relatively scattered but more concentrated in DailyStrength. Table 2 presents the clusters. Cluster 1 refers to the main direct-related topics. Other topics are divided into 2 categories in Man Youbang. Cluster 2 refers to few disease-specific topics about diabetes or complications, while cluster 3 refers to non-disease specific topics about living habits. In DailyStrength, the remaining topics do not be divided. Therefore, we can infer American HCQA users pay more attention to the impact of diet and exercise on prevention. Improper living habits is an important cause of diabetes, such as high-calorie eating habits.

Table 2–Topics clusters results

Man Youbang		DailyStrength	
Clusters	Topics	Clusters	Topics
Treatment and protection of disease	Treatment, Special cases, State of an illness, Emotional support, Finance, Drug's side effects, Insurance	Treatment and protection of disease	Emotional support, Treatment, State of an illness, Complication, Drug's side effects, Special cases, Social support, Insurance, Finance
Prevention and complications	Prevention, Complication, Diagnosis	Diagnosis and management of disease	Diet, Exercises, Diagnosis, Blood glucose monitoring, Prevention
Daily management of disease	Diet, Exercises, Blood glucose monitoring		

## Discussion

Using topic analysis and social network analysis, we examine hot topics evolution, related terms, and co-occurrence of HCQA in China and the USA. Then we will summarize the main similarities and differences of user information needs.

### Common characteristics

First, the information needs of users in HCQA are similar. The main purpose is to provide patients with sharing disease and treatment experience. The core need is "how to treat diseases," which involves such topics as "state of illness," "treatment," "complication," and so on. It mainly depends on the function of HCQA. HCQA are working to provide information exchange channel for participants. Even if users do not intend to ask

questions about treatment, most of the narratives relate to the core. Second, "diet" and "finance" are the most popular non-disease related topics. This is related to characteristics of chronic diseases, such as long treatment period, the healing process, multiple complications, high cost, etc. which brings greater financial burden to patients. Most patients expect to get practical and reliable information about treatment and health management in the community.

### Different Characteristics

First, communities stand at different stages. American communities started early and develop stable now. The composition of members and demands are relatively stable, obvious, and connected more closely. Chinese communities started late and possess development period. Users' composition and demands fluctuate widely and are not closely related to each other. Second, there are different attributes. The attribute of "Q&A" is remarkable in Chinese communities. Patients prefer direct and short narrative for paying more attention to practicality. But they seldom seek mood expression and emotional support. However, American communities also show the attribute of "social connect" in addition to "Q&A." The former is even more significant. According to the content of related posts, it is related to differences in the social environment and living habits between Chinese and American communities. Due to religious beliefs and humanistic care, American HCQA users are better at expressing emotions and communicating with people. They intend to use more emotional words in expressions and eager to be comforted by others. Chinese users implicitly express their emotions and do not express their emotions to strangers or seek encouragement. It also reflects low trust and dependence in community members. Although interesting findings were made, this study has its limitations. First, correlation degree and the quality of Q&A have not been considered. It remains to be studied to take correlation degree into the multi-dimensional analysis framework. Second, researches on topic and content evolution are not enough. We should consider time slice analysis by season, quarter, etc.

## Conclusions

In this study, we analyze hot topics evolution and co-occurrence on the topic of diabetes. On the one hand, the information needs of HCQA in Chinese and American are similar focusing on "how to treat diseases." Other living-habits-specific needs, such as "diet," are clearly expressed in chronic disease groups. On the other hand, Chinese communities appear the attribute of "Q&A," and members pursue practicality. By contrast, American communities have both the attribute of "Q&A" and "social connect." Furthermore, topics fluctuate frequently in Chinese communities while stable in American communities.

In theory, we combine Social Network Analysis and content analysis to capture and contrast the pattern of Q&A in HCQA. Some clear conclusions are obtained referring to the similarities and differences of users' information needs of HCQA in China and the USA, which has certain theoretical significance for following researches. In practice, the founding provides enlightenment for the development of HCQA. Chinese communities should focus on improving the content quality and strengthen emotional connection among users to retain them. In addition, the rising incidence of diabetes must be taken seriously by government departments, which can take certain measures to further improve disease treatment infrastructure. For example, they can promote the role of community service and strengthen medical security to reduce the financial burden of patients.

## Acknowledgements

This work was supported by the National Natural Science Foundation of China (Nos. 71661167007).

## References

- [1] E. Manafo and S. Wong, Exploring older adults' health information seeking behaviors, *Journal of Nutrition Education & Behavior* **44** (2012), 85-89.
- [2] V.M. Conley, Beyond knowledge deficit to a proposal for information-seeking behaviors, *International Journal of Nursing Terminologies & Classifications* **9** (2010), 129-135.
- [3] J. Wu, X. Huang, K. Dong, A review of online medical research based on knowledge graph, *Journal of Information Resources Management* **2** (2016), 4-21. [4] N.J. Gray, J.D. Klein, P.R. Noyce, T.S. Sesselberg, and J.A. Cantrill, Health information-seeking behaviour in adolescence: the place of the internet, *Social Science & Medicine* **60** (2005), 1467-1478.
- [5] Y. Li and W. Cai. A review of the studies on health information seeking behavior overseas, *Library & Information Service* **19** (2012), 128-132.
- [6] Tustin N. The role of patient satisfaction in online health information seeking. *Journal of Health Communication* **1** (2010), 3-17.
- [7] Y. Zhang, Contextualizing consumer health information searching: an analysis of questions in a social Q&A community, in: *ACM International Health Informatics Symposium*, 2010, pp. 210-219.
- [8] C. Betsch, N.T. Brewer, P. Brocard, P. Davies, W. Gaissmaier, N. Haase, J. Leask, F. Renkewitz, B. Renner, and V.F. Reyna, Opportunities and challenges of Web 2.0 for vaccination decisions, *Vaccine* **30** (2012), 3727-3733..
- [9] N. Priya, Information seeking and social support in online health communities: impact on patients' perceived empathy, *Journal of the American Medical Informatics Association* **18** (2011), 298-304.
- [10] S. Kim, S. Oh, and J.S. Oh, Evaluating health answers in a social Q&A site, *Proceedings of the Association for Information Science & Technology* **45** (2010), 1-6.
- [11] J.S. Oh, D. He, J. Wei, E. Mattern, and L. Bowler, Linguistic characteristics of eating disorder questions on Yahoo! Answers – content, style, and emotion, in: *Asis&t Meeting: Beyond the Cloud: Rethinking Information Boundaries*, 2013, p. 87.
- [12] S.C. Kim, D.V. Shah, K. Namkoong, F.M. Mctavish, and D.H. Gustafson, Predictors of online health information seeking among women with breast cancer: the role of social support perception and emotional well-being, *Journal of Computer-Mediated Communication* **18** (2013), 98-118.
- [13] A.S. Ahadzadeh, S.P. Sharif, and F.S. Ong, Online health information seeking among women: the moderating role of health consciousness, *Online Information Review* **42** (2017), 58-72.
- [14] N.S. Coulson. Sharing, supporting and sobriety: a qualitative analysis of messages posted to alcohol-related online discussion forums in the United Kingdom, *Journal of Substance Use* **19** (2014), 176-180.
- [15] M.A. Arden, A.M.S. Duxbury, H. Soltani. Responses to gestational weight management guidance: a thematic analysis of comments made by women in online parenting forums. *BMC Pregnancy and Childbirth* **14** (2014), 216.
- [16] M.S. Park, H. Park. Topical network of breast cancer in formation in a Korean American online community: a

semantic network analysis. *Information Research: An International Electronic Journal* **21** (2016), n4.

- [17] J.S. Oh, D. He, W. Jeng, et al. Linguistic characteristics of eating disorder questions on Yahoo! Answers – content, style, and emotion. *Proceedings of the American Society for Information Science & Technology* **50** (2014), 1-10.

## Address for correspondence

Corresponding author: Liqin Zhou, PhD student. Center for Studies of Information Resources, Wuhan University, Wuhan, Hubei, China  
E-mail: zhoulq92@163.com