



# 国内外健康问答社区用户信息需求对比研究<sup>\*</sup>

## ——基于主题和时间视角的实证分析

石 静<sup>1</sup> 厉臣璐<sup>1</sup> 钱宇星<sup>1</sup> 周利琴<sup>1</sup> 张 斌<sup>1,2</sup>

<sup>1</sup>(武汉大学信息资源研究中心 武汉 430072)

<sup>2</sup>(武汉大学中国传统文化研究中心 武汉 430072)

**摘要:**【目的】识别并对比分析国内外健康问答社区用户信息需求,总结演变规律和探究演变原因。【方法】选取“慢友帮”和“DailyStrength”糖尿病群组用户问答数据,基于主题和时间视角,利用主题编码、社会网络和内容分析方法,对比分析主题分布、演化及共现规律。【结果】国内外社区用户核心需求都是“疾病如何治疗”;且对于慢性病社区而言,“饮食”主题与其共现主题关系很密切。【局限】尚未将问答相关程度纳入分析框架,对主题演变及内容研究不够深入。【结论】国内社区处于发展期,国外社区处于稳定期;国内社区具有“问答属性”,而国外社区兼具“问答属性”和“社交属性”。

**关键词:** 健康问答社区 用户需求 内容分析 共现网络

**分类号:** G203 R587.1

**DOI:** 10.11925/infotech.2096-3467.2018.0813

### 1 引言

目前,越来越多的网络问答社区采用开放的问答模式<sup>[1]</sup>,不再设置“最佳答案”或“被采纳答案”以结束问题,而是鼓励更多用户参与深入探讨,问题之下的观点更加多元,答案更加准确和全面<sup>[2]</sup>。网络问答社区的演化发展,协同用户需求的多样化,推动了专业性问答社区的出现,健康问答社区就是其中之一。用户通过问答形式针对性地在社区中贡献、搜寻和获取健康信息。Pew Research Center 的报告显示:80%的网络用户会在互联网上搜寻与健康主题相关的信息,其中

34%的用户浏览其他人发表在网络论坛或博客上与健康医疗相关的评价和经历,24%的用户在线咨询某些药物和医疗诊断等信息<sup>[3]</sup>。

国内外出现了大量健康问答社区:如“Patients like me”、“DailyStrength”、“慢友帮”、“丁香园”等。以上社区允许用户通过发布问题和提供答案分享健康信息和经验,不同于“百度知道”、“知乎”等综合型问答社区,主要专注于健康知识的交流和共享,通过病友互助圈培养用户之间的情感依赖,得到更多医生、患者和关注健康人士的青睐。然而该类型问答社区中存在的问题同样严峻,如用户参与率参差不齐、问答质量不

通讯作者:张斌, ORCID: 0000-0002-5591-7874, E-mail: zb0205@whu.edu.cn。

\*本文系国家自然科学基金组织间合作研究项目“基于慢病知识管理的智慧养老平台研究”(项目编号: 71661167007)、国家自然科学基金重点国际(地区)合作研究项目“大数据环境下的知识组织与服务创新研究”(项目编号: 71420107026)和国家自然科学基金青年项目“心智空间视角下科学知识生成与演化机理研究”(项目编号: 71704138)的研究成果之一。

\*本文系“首届全国高校数据驱动创新研究大赛”优秀论文选登。

高等。针对这些问题,国内外学者已有一定的研究,主要集中在持续使用意愿、知识共享动机、用户采纳等方面,较少从用户角度出发研究用户信息需求。

为了更有效地识别国内外健康问答社区中的用户信息需求,本文基于“慢友帮”和“DailyStrength”两大健康社区中糖尿病主题下的问答数据,从主题和时间双重视角,结合内容分析、主题编码和社会网络方法识别国内外健康问答社区用户信息需求的特征与异同,为国内健康问答社区发展提供借鉴。

## 2 相关研究

交互问答平台是一种开放的服务平台,基于交互问答平台所建立的问答知识库可纳入到搜索信息源之中,提高搜索引擎的搜索质量<sup>[4]</sup>。健康问答社区作为交互问答平台的一种,通过健康知识共享与获取,更便捷、快速和准确地服务于用户的健康信息搜寻行为,满足用户的健康信息需求。

### 2.1 健康信息需求研究

健康信息泛指与人们身心健康、疾病、营养、养生等相关的信息,可以指导健康及临床行为。用户出于某种特定健康条件的驱使,或因拜访医生的需求,或对营养、锻炼及控制体重等方面存在疑惑,选择通过广播、电视、报纸、杂志、网络等多种方式搜寻健康信息<sup>[5]</sup>。

早期对于健康信息需求的研究主要集中于医学领域,以患有特定疾病、处于特定病程的患者为研究对象,采取问卷调查或访谈方法识别用户健康需求的基本情况、类别以及影响因素等。例如任蔚虹<sup>[6]</sup>、曾倩等<sup>[7]</sup>分别针对处于围手术期和孕期的患者进行健康信息需求研究;赵华等<sup>[8]</sup>、魏永婷等<sup>[9]</sup>分别以白血病、癌症患者为研究对象,识别其信息需求,帮助提供教育和咨询服务。随着互联网的发展,用户越来越倾向于在线搜索健康信息,网络健康信息搜寻行为中的信息需求开始受到关注。克利夫兰等<sup>[10]</sup>较早对达拉斯福和地区华人网络健康信息资源的使用情况进行调研,结论证明用户对营养信息和特定疾病信息最感兴趣,并且十分愿意学习使用互联网查找信息。之后,张馨遥等<sup>[11]</sup>、张洪武等<sup>[12]</sup>也分别对网络用户健康信息需求及其影响因素进行研究,显示用户对特定疾病信息需求较高,并且健康状况差的用户更倾向于搜寻在线健康信息。新媒体广泛应用之后,金碧漪等<sup>[13]</sup>、徐倩等<sup>[14]</sup>、

李小虎等<sup>[15]</sup>分别基于社会化问答社区、移动医疗 APP 和微信等新媒体对用户健康信息需求及行为进行分析。研究结论均显示用户具有强烈的健康信息需求,并且对特定疾病信息最为关注。

纵观前人研究,用户健康信息需求与健康信息的获取方式、搜寻行为息息相关。如今专门化的健康问答社区兴起,新的环境下用户信息需求有何特点是值得探究的问题。此外,前人研究以问卷调查、访谈等方法为主,后期有少量的文本挖掘方法。这可以在一定程度上反映用户信息需求,但却很少考虑不同信息需求之间的关系,不便深入分析需求特征、演化及其动力。本文在文本挖掘方法的基础上结合社会网络分析方法,将各类信息需求的关系考虑进来,以期进行全面深入的分析。

### 2.2 健康问答社区主题研究

健康问答社区既包括面向专业人士的医疗健康网站,也有面向普通大众的门户网站。社区将拥有类似需求和经历的用户聚集在一起,发挥集群优势,社区内的提问自发产生聚类,用户可更有针对性地搜集和发布信息。交流形式主要是文本信息的互动,如专家问答、健康博客和病友论坛等;参与者包括普通民众、病人人群、病患家属、专业医师等各类型用户<sup>[16]</sup>。国内外学者对健康问答社区主题的研究大多先采用编码方式得到相关主题,再进行后续分析,如构建回答质量评价体系<sup>[17]</sup>,研究提问行为的激发因素<sup>[18]</sup>,总结消费者健康信息需求类型<sup>[13]</sup>等。

Coulson<sup>[19]</sup>针对酗酒论坛中的 758 条消息进行分析,得到三个主题:分享、支持和醒酒。李重阳等<sup>[20]</sup>以在线问答社区中的癌症信息为例,从时间和主题两个角度综合测度信息需求,采用 LDA 主题挖掘法确定相关编码规则,探求用户信息需求特征、主题关注度及变化趋势。Oh 等<sup>[21]</sup>以 Yahoo! Answers 中有关饮食失调的问答数据为研究样本,通过词频、词性、情感分析等方法总结出相关主题。Arden 等<sup>[22]</sup>选取 400 条涉及英国国家卫生医疗质量标准署发布的关于孕中体质量控制指导文件的论坛帖子为研究对象,探究孕中体质量的影响因素。金碧漪等<sup>[23]</sup>基于 Yahoo! Answers 和 Diabetic Connect 网络健康社区中糖尿病相关的社会化标签和问答记录,通过数据编码和文本处理得到 8 大类主题,比较两大社区内 8 大主题的分布情况,利用统计分析发现糖尿病相关的信息主题分布特征。

## 2 数据分析与知识发现

这些研究多止步于识别热门主题,没有结合问答内容深入分析;并且大多选取单个社区作为研究对象,缺乏对比。本文选择国内外较具代表性的医疗社区,在主题识别的基础上,结合问答内容进行分析对比,为国内外医疗社区的发展提供借鉴。

3 研究设计

选取健康问答社区“慢友帮”和“DailyStrength”中的糖尿病问答数据作为研究对象,利用内容分析、主题编码和社会网络分析方法进行用户健康信息需求研究;探讨国内外健康问答社区用户关注主题热度异同,不同时期主题演化对比等问题。结合社会背景,对国内外健康问答社区整体情况进行基于主题和时间视角的对比,以期有效推动国内健康问答社区发展,更好地满足用户健康信息需求。

3.1 数据采集和处理

由于慢病治疗周期较长、药物种类较多,患者倾向于借助网络查找疾病相关信息,并进行交流。据国际糖尿病联盟(IDF)更新的《2017 年全球糖尿病地图(第 8 版)》显示:2017 年全球约 4.25 亿成人患糖尿病,预计到 2045 年,糖尿病患者可能达到 6.29 亿。2017 年 20-79 岁糖尿病患者数量居前三位国家分别为中国、印

度和美国,分别高达 1.144 亿、7 290 万和 3 020 万<sup>[24]</sup>。全球来看,糖尿病已经成为威胁健康的一大“杀手”;而 2017 年我国糖尿病患病率高达 10.9%,约有 84 993 人死于糖尿病。而相对于国外,国内健康问答社区起步较晚,尚未形成机制完备的大型社区,多为实力一般的小型社区,并且出现衰退趋势。

“慢友帮”和“DailyStrength”是国内外健康问答社区中的典型代表,前者建立于 2013 年,目前已形成 84 个慢病互助圈,糖尿病互助圈就是其中一个,注册用户 7 995 人;后者建立于 2006 年,目前存在 421 个疾病讨论组,糖尿病组注册用户 1 096 人。选取两个社区中的糖尿病群组作为数据来源,爬取群组中 2013 年-2017 年的问答数据。前者共计 777 个帖子、3 553 条评论,后者共计 2 115 个帖子、11 754 条评论。某些问答虽然发布在糖尿病群组,但与糖尿病没有关系或者没有实际意义,对这部分数据进行清洗后得到“慢友帮”帖子 653 个、评论 2 556 条;“DailyStrength”帖子 1 640 个、评论 9 179 条。

3.2 主题编码

研究中常常采用人工标注、词频统计等方法总结用户信息需求特征<sup>[20]</sup>,本文选择三级编码形式<sup>[25]</sup>针对数据集进行人工主题编码,用二级编码主题词代表用户信息需求。编码表如表 1 所示。

表 1 用户问答数据涉及主题的编码表

三级编码	二级编码	一级编码	内涵
疾病 直接相关	预防	如何预防糖尿病	糖尿病高危人群,如何从遗传、环境、饮食、运动等因素预防糖尿病
	诊断	症状	确诊之前机体内的一系列异常变化所引起的病人主观上的异常感觉或某些客观病态改变
		病因	导致患糖尿病的原因
		分型	糖尿病分型,一般分为 1 型和 2 型
	病情	病情描述	患者对于当前病情的描述
	治疗	胰岛素治疗	采用注射胰岛素的方式治疗
		其他方式治疗	使用仪器、其他药物等方案进行治疗
	并发症	糖尿病足	患者因周围神经病变与外周血管病合并过高的机械压力,引起的足部问题
		酮症酸中毒	患者因糖尿病引起的多尿、烦渴、乏力等糖尿病症状加重
		视网膜病变	糖尿病性微血管病变中最重要的表现,具有特异性改变的眼底病变
		其他并发症	如糖尿病神经病变、糖尿病肾病等
	血糖监测	患者自我血糖监测	患者借助于仪器、试纸等对自己的血糖变化进行观察记录
	特殊情况	妊娠糖尿病	孕妇在妊娠前或妊娠时患糖尿病
		青少年	患者在 13-25 岁确诊糖尿病
		老年糖尿病	患者在 60 岁之后确诊糖尿病
		其他情况	患者出糖尿病之外还患有其他疾病
非疾病 直接相关	药物副作用	药物不良反应	患者因服用糖尿病药物产生的恶心、呕吐、脱发、疲倦等症状
	饮食	饮食注意事项	患病期间帮助治疗、恶化治疗的食物
	运动	锻炼减肥	患病期间通过锻炼减肥以辅助治疗
	保险	医疗保障	患病治病时可以获得报销补偿的保障
		其他保险	除医疗保障外,患者办理的其他保险
	财务	经济问题	患者治疗期间的经济压力
	情感支持	精神支持	治病期间向其他人寻求安慰、鼓励等
	社会支持	慈善救助	患病治病期间获得的财物等慈善救助

### 3.3 主题共现网络

提取编码表中的二级编码作为主题词进一步分析。由于大多数问答数据并非只涉及一个主题,例如,2014年,“DailyStrength”一个帖子“……我23岁了。我19岁时被诊断出患有2型糖尿病。当我第一次被诊断时,我改变了我的饮食习惯……我锻炼了……我尽了一切努力减肥,已经尝试了无数次,但总是失败。但在这一刻,我觉得如此低落,我的自尊心很低。我不知道该怎么办,如果你们谁能帮助我,我将不胜感激。”就同时涉及到“病情”、“饮食”、“运动”和“情感支持”4个

主题。这种不同主题之间的共现关系在一定程度上反映患者问答过程中信息需求的关联,提取主题共现关系构造矩阵,利用可视化软件 Gephi 构造共现网络。

## 4 结果分析

### 4.1 国内外主题热度对比

本文对糖尿病问答帖全文进行采集与清洗,并依照三级编码规则对样本进行主题编码,对编码结果进行统计。由于国内外社区整体数据量存在差异,因此计算各主题所占比例,结果如图1所示。

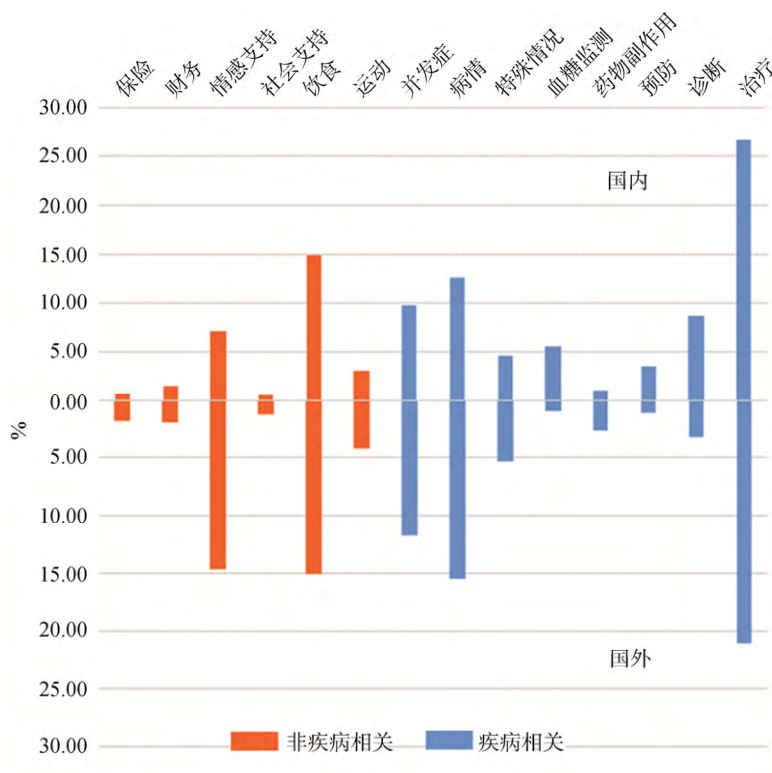


图1 国内外问答主题分布比例

通过数据对比发现,国内外健康问答社区糖尿病专题下用户关注热点大体相似,关注度最高的主题均为:治疗、饮食、病情。此外,值得关注的一点是“社区卫生服务”这一主题在国内外健康问答中均较少提及,可以推测,社区卫生服务作为慢性病的重要辅助治疗手段之一,在实践中并没有得到恰当的落实或足够的重视,致使慢性病治疗方面存在部分空白,对以糖尿病为代表的慢性疾病治疗基础设施进一步改进与完善提出了方向。

但从主题分布上来讲,又存在差异。国内健康社区问答主题热点集中于“治疗、病情、并发症、诊断”

等“疾病直接相关”主题,具有实用性特征,“治疗”主题占比超过全部主题的四分之一,是国内糖尿病用户关注的重中之重。而国外健康问答社区的问答主题热点在“治疗,病情”等直接相关主题之外,“情感支持”这一“非疾病直接相关”主题占比突出,达到15%。表明国外用户非常注重情感变化对疾病治疗的影响,也更加倾向于在社区中寻求他人的安慰和鼓励。

在以上数据基础上,选取国内外健康问答社区5大热点主题“治疗,饮食,病情,并发症,情感支持”进一步对比,结果如图2所示。

## 4 数据分析与知识发现



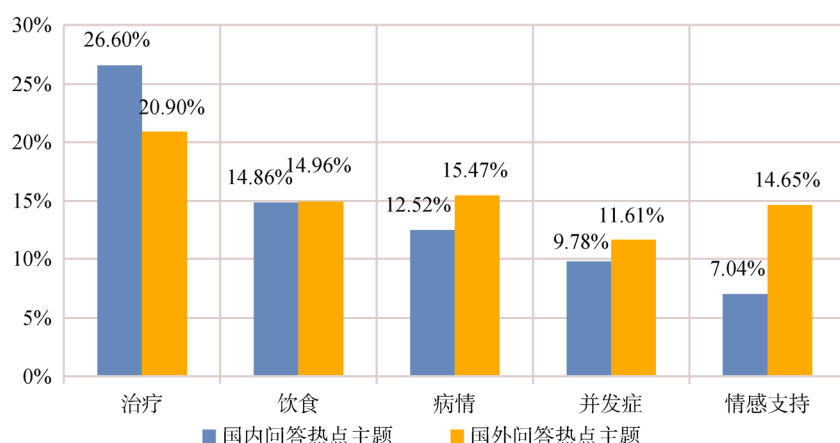


图 2 国内外健康问答社区问答五大热点主题对比

国外健康问答社区问答内容中“情感支持”主题占比约为国内数据的两倍。大多提问者不仅会对自身病情进行描述,还会呈现治疗过程中的情感变化,甚至单纯地寻求情感支持;回答者也会对提问者提供关心、慰问、支持等。

这些差异表明国外健康问答社区不仅具有“问答属性”,还具有“社交属性”,患者对社区信任、依赖程度较高,会在问答过程中表达情感,寻求支持。此外,也表明相对于国内,国外患者更容易表达情感,这与国内外的社会环境、文化差异等不无关系。最后,国外患者在糖尿病等慢性疾病的治疗过程中,除药物和物理手段(如运动)之外,也注重情绪对患者的支撑和辅助治疗作用。

#### 4.2 主题热度演变

结合时间标签,以一年为单位,分别统计国内外

糖尿病论坛 2013 年—2017 年各年主题的构成情况,结果如图 3 和图 4 所示。

除“情感支持”外,国内外问答主题的比例分布较为相似,关注较多的主题在各年主题构成中都有所体现。但 2013 年—2017 年国内外问答主题演化趋势存在明显差别。

(1) 主题演化总体趋势不同。国内每年的主题热度变化明显,而国外各个主题的热度变化不大。这与国内外问答社区的发展情况有关,以“DailyStrength”和“慢友帮”为例,前者创立于 2006 年,已经度过了初始期、发展期,进入稳定期;后者创立于 2013 年,新确诊的患者不断加入到论坛中分享病情、咨询治疗方法,导致各类病情、病程的用户波动性较大。

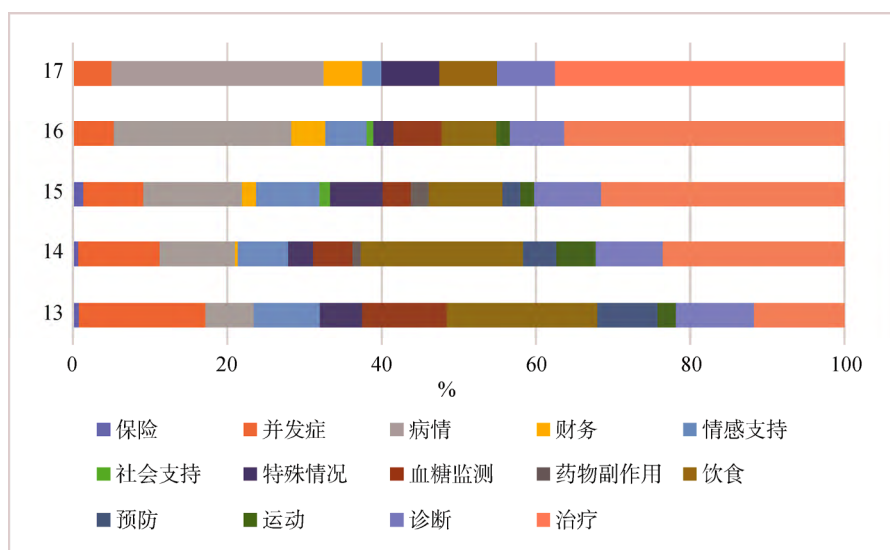


图 3 “慢友帮”糖尿病论坛主题标签构成

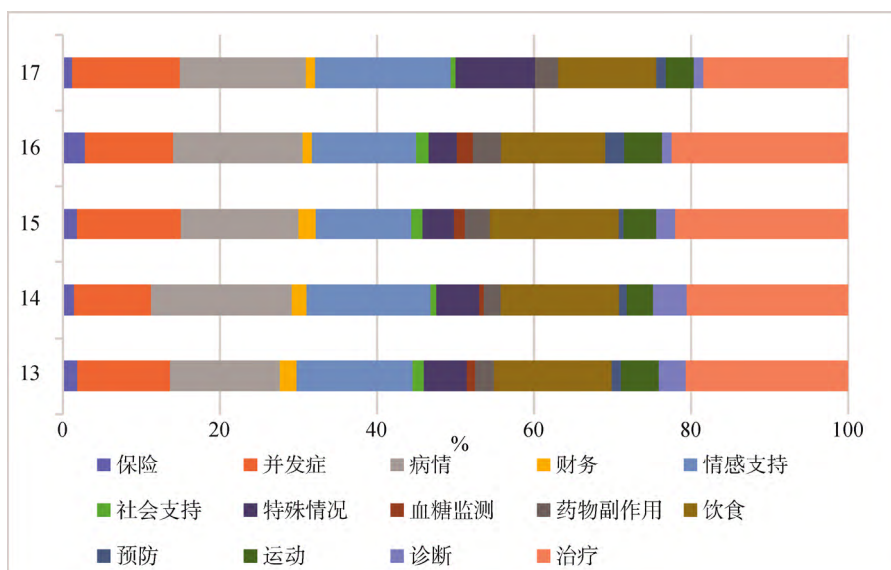


图 4 “DailyStrength”糖尿病论坛主题标签构成

(2) 具体主题变化趋势不同。国内糖尿病论坛中,“治疗”、“病情”和“财务”主题占比有明显提高的趋势,而“并发症”、“饮食”主题占比逐年降低,“情感支持”和“诊断”等主题占比较为稳定。国外糖尿病群组每年各标签占比均较为均衡,没有明显增减。“并发症”主题一般由病情较重、病程较长的患者发布,“治疗”和“病情”等多为新确诊患者发布的治疗咨询、病情介绍。国内“并发症”主题占比逐年明显降低,而“治疗”、“病情”和“财务”主题占比明显提高,反映出国内的新确诊患者数量逐年增加、发病率上升、治疗经济负担增大的情况。

#### 4.3 主题共现网络对比

去除只涉及单个主题的数据,利用主题之间的共现关系构造共现网络,国内外主题共现网络图如图 5 所示。节点代表不同的主题,节点越大,中介中心性越高,表明该主题更容易与其他主题共现;节点颜色越深,度中心性越高,表明该主题越重要。连边代表主题之间的共现关系,连边越粗,主题之间的共现频率越高,关系越密切。

(1) 国内外问答主题热度分布大体相似,表 2 给出了不同主题的统计信息。

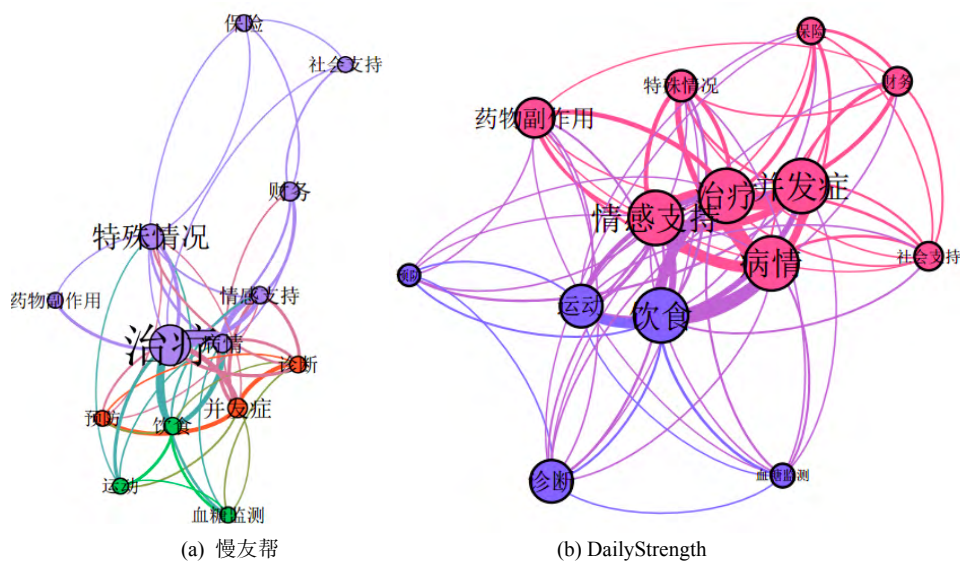


图 5 主题共现网络图

表 2 主题属性统计表

主题	国内			国外		
	度中心性	加权重	节点平均加权重	度中心性	加权重	节点平均加权重
治疗	13	168	12.9	13	561	43.2
并发症	10	31	3.1	13	340	26.2
特殊情况	10	73	7.3	10	132	13.2
病情	9	64	7.1	13	508	39.1
饮食	9	125	13.9	13	414	31.8
情感支持	8	47	5.9	13	380	29.2
诊断	8	67	8.4	11	40	3.6
运动	6	23	3.8	11	180	16.4
财务	6	21	3.5	10	84	8.4
血糖监测	6	18	3.0	9	39	4.3
预防	6	24	4.0	8	29	3.6
保险	4	5	1.3	9	72	8.0
社会支持	3	5	1.7	10	65	6.5
药物副作用	2	9	4.5	11	76	6.9

①国内外度中心性较高的主题均为“治疗”、“特殊情况”、“病情”和“饮食”，在其他共计 13 个主题中都 与 10 个以上主题具有共现关系，尤其是“治疗”与所有主题都有共现关系，共现总次数最高，是最为关键的主题。这不难理解，对于一个健康问答社区而言，大多数用户倾向于了解应该服用什么药物，不同治疗方法的效果如何等与治 疗直接相关的问题；即使问答目的并非“治疗”，也会出现与其相关的叙述。

②国内外问答中都存在一个有趣的现象，“饮食”主题虽未与较多主题共现，但平均共现次数却很高，即“饮食”与其共现主题的关联性较高。以国内“慢友帮”社区为例，与“饮食”联系最为密切的主题依次是“治疗”、“病情”和“运动”。在糖尿病的治疗方法中也有食疗，这解释了为何“饮食”与“治疗”、“病情”共现频繁。同时，过度肥胖也是糖尿病病因之一，对于糖尿病患者而言，减肥是另一个热点主题。“运动”与“饮食”作为减肥的两大途径，关系必然密切。

(2) 国内外问答主题共现情况存在明显差异。

①整体网络属性不同，“DailyStrength”社区中各主题联系更为紧密。图 5(a)“慢友帮”主题共现网络图，包含 14 个节点，50 条连边。从网络整体来看，平均度为 7.143，即每个主题约与其他 7 个主题存在共现关系；图密度为 0.549，网络平均加权重为 48.571，各个主题之间连接不甚紧密。图 5(b)“DailyStrength”主题共现网络图，包含 14 个节点，77 条连边。从网络整体来看，平均度为 11，即每个主题约与其他 11 个主题存在共现关系；图密度为 0.846，网络平均加权重为 208.571，各个主题之间联系很

紧密。

②网络角色构成不同，“慢友帮”中大节点相对较少，而“DailyStrength”中较多，度中心性在 10 以上的节点占比 79%，其中“并发症”、“病情”、“情感支持”、“饮食”与“治疗”均与其他所有主题共现。在网站调研过程中，发现国外问答社区具有明显的“分享性质”，患者在群组中不是单纯地以问问题为目的，而是分享自己最近的病情与心情，甚至是身边发生的小事，这就使得问答字数较多，可能涉及到的主题也就较多。而“慢友帮”的“实用性质”明显，国内患者在提问时大多较为简短、目的性强，开门见山提出疑问，因此涉及到的主题较少。

③用户情感需求不同，国外用户情感表达更多，希望通过社区获得情感支持。由图 5 可以看到，“情感支持”不仅与全部主题都具有共现关系，并且平均共现次数也较高。糖尿病治疗周期长、并发症多，长期的治疗过程极易使患者产生疲倦消极情绪；糖尿病治疗与生活习惯关系密切，尤其是饮食与运动两方面，因此很多患者生活习惯被迫改变，生活质量下降；并且长期患病治疗，患者丧失劳动能力，产生较大的经济负担。这些都是患者在问答社区中寻求情感支持的原因。

④主题聚类情况不同，国内较为分散，而国外较为集中。国内外主题聚类情况如表 3 所示。两者类别 1 构成相似，涉及主题较多，基本是与糖尿病直接相关的主题。在健康问答社区中，患者首要关注的问题是糖尿病如何治疗，在求助过程中，必然涉及对自身病情、使用过何种治疗方法、效果如何的阐述。

表3 国内外主题聚类结果

国内		国外	
类别	主题	类别	主题
类别1: 疾病治疗与保障	治疗、特殊情况、病情、情感支持、财务、药物副作用、保险	类别1: 疾病治疗与保障	情感支持、治疗、病情、并发症、药物副作用、特殊情况、社会支持、保险、财务
类别2: 预防与并发症	预防、并发症、诊断	类别2: 疾病诊断与管理	饮食、运动、诊断、血糖监测、预防
类别3: 疾病日常管理	饮食、运动、血糖监测		

其他主题国内呈现两类,类别2中涉及主题较少,也是与疾病紧密相关的类别,围绕糖尿病或并发症确诊之前的问题展开;类别3与疾病本身并非直接相关,而是围绕生活习惯展开,包括饮食、运动两大因素对血糖的影响,说明国内饮食、运动等主要是为了控制血糖值。而国外类别2类似于“慢友帮”中的类别2和类别3,涉及饮食、运动等生活习惯和疾病预防、诊断等。表明国外较为关注饮食、运动等生活习惯对糖尿病预防的影响,也从侧面反映出不当的生活习惯是国外糖尿病的一个重要诱因,这很可能与国外高热量的饮食习惯有关。

## 5 国内外健康问答社区整体对比

通过主题编码、热点分布、演化与共现分析,总结得到国内外健康问答社区的异同。

### 5.1 国内外健康问答社区共同点

(1) 国内外健康问答社区用户信息需求相似。无论是国内还是国外,用户核心需求都是“疾病如何治疗”,具体涉及“病情”、“治疗”、“并发症”等主题。这主要取决于健康问答社区本身的功能定位,作为一种专业性社区,其参与者以患者及其家属为主,主要目的是为这些用户提供病情及治疗交流。即使用户本意并非提问关于治疗的问题,在其叙述过程中,大多也会围绕核心需求展开。

(2) 在非疾病直接相关主题中,关注度较高的是“饮食”与“财务”。这与慢性病的自身特征有关。糖尿病作为一种与饮食习惯密切相关的慢性病,不恰当的饮食会导致患者血糖波动较大,因此很多患者会在互助圈中讨论三餐食谱、某种食品是否可以食用、是否有助于控制血糖等问题。慢性病治疗周期长、治愈慢,病情控制困难且并发症多,患者极易丧失劳动能力,治疗费用高,带来了较大的经济负担。大部分患者期望通过社区获得切实、可靠的疾病治疗与健康管理信息,解决患者本身及其家属的疑惑和困难。

### 5.2 国内外健康问答社区不同点

(1) 国内外健康问答社区发展阶段不同。国外社区起步较早,目前处于稳定期,经过用户的积累,各

类病情、病程的用户结构相对平衡,用户参与率高,需求主题较为稳定显著,不同需求之间联系更紧密。国内社区起步较晚,目前处于发展期,用户“进出”频繁、参与率低,尚未达到均衡状态,需求主题波动较大,彼此之间联系不甚紧密。

(2) 国内外健康问答社区总体属性存在差异。国内社区“问答属性”显著,患者注重实用性,倾向于开门见山、较为简短的叙述方式,较少表达及寻求情感支持。而国外社区除“问答属性”之外,还具有“社交属性”,患者不仅想要获得答案,还想借助社区表达情感、获得情感支持,后者甚至更为重要。因此,患者提出问题之前,还会分享自己的病情、治疗经验及情感变化,倾向于情绪化、较为冗长的叙述方式。结合相关帖子内容分析可知,这与国内外社会环境、文化差异及生活习惯的差异有关。国外用户出于宗教信仰、人文关怀等原因,更善于表达情感,更乐于与人交流,因此面对疾病带来的疲倦消极情绪与财务压力,患者会直接表达心情,在问答中使用更多与情感相关的词汇;并且渴望得到他人的安慰。而国内用户在情感表达上较为含蓄,并不倾向于向陌生人表达情感、寻求鼓励,也侧面反映出用户对社区及其他用户的信任、依赖程度较低。

## 6 结 语

本文选取“DailyStrength”和“慢友帮”两个典型社区,借助于问答主题对国内外健康问答社区用户信息需求进行实证分析,识别热点主题、主题演化及共现规律,总结其异同。一方面,国内外健康问答社区用户信息需求基本相同,围绕“疾病如何治疗”展开;其他与慢性病日常管理相关的需求,如“饮食”、“运动”等生活习惯表现明显。在国内,生活习惯与疾病日常管理关系密切;在国外,生活习惯与疾病预防诊断等关系密切,这与国内外生活方式的差异不无关系。另一方面,国内社区具有“问答属性”,用户追求实用性,而国外社区兼具“问答属性”和“社交属性”,用户情感需求较为强



烈;国内社区起步晚,主题波动频繁,处于发展期,而国外社区起步早,主题数量与结构稳定,处于稳定期。这为国内外健康问答社区的管理与发展提供了启示,尤其是起步较晚的国内社区,目前正处于关键时期,充分考虑患者需求的同时,应当采取措施提高用户对社区及其他用户的信任程度与情感依赖,增加用户粘性,把重点放在提高问答质量及“留住用户”上。

本研究尚有不足,未考虑问答之间的相关程度,下一步研究中考虑将该指标纳入多维分析框架,分析问答相关度与用户满意度、信息需求的联系;此外,对主题演变及帖子内容研究不够深入,应考虑进行时间切片或利用更细致的文本挖掘方法加强对帖子内容的分析。

### 参考文献:

- [1] 刘佩,林如鹏.网络问答社区“知乎”的知识分享与传播行为研究[J].图书情报知识,2015(6):109-119.(Liu Pei, Lin Rupeng. Research on Knowledge Sharing and Dissemination Behavior of Online Q&A Services: Taking Zhihu as an Example[J]. Documentation, Information & Knowledge, 2015(6): 109-119.)
- [2] 龚凯乐,成颖.基于“问题-用户”的网络问答社区专家发现方法研究[J].图书情报工作,2016,60(24):115-121.(Gong Kaile, Cheng Ying. Finding Experts in Community Question Answering Based on “Question-User” [J]. Library and Information Service, 2016, 60(24): 115-121.)
- [3] Pew Research Center. Internet Health Resources[EB/OL]. [2013-06-16]. <https://www.pewinternet.org/2003/07/16/internet-health-resources/>.
- [4] 邓胜利.国内外交互问答平台的比较及其对策研究[J].情报理论与实践,2009,32(3):50-55.(Deng Shengli. Comparison and Countermeasures of the Interactive Q&A Platforms at Home and Abroad[J]. Information Studies: Theory & Application, 2009, 32(3): 50-55.)
- [5] Manafo E, Wong S. Exploring Older Adults' Health Information Seeking Behaviors[J]. Journal of Nutrition Education & Behavior, 2012, 44(1): 85.
- [6] 任蔚虹.围手术期病人健康信息需求的调查分析[J].中华护理杂志,1999,34(8):458-460.(Ren Weihong. Patient's Need of Health Education During Perioperative Period[J]. Chinese Journal of Nursing, 1999, 34(8): 458-460.)
- [7] 曾倩,林雪霞,霍魏魏,等.孕期孕妇健康信息需求调查分析[J].中国妇幼保健,2004,19(2):26-28.(Zeng Qian, Lin Xuexia, Huo Weiwei, et al. Investigation and Analysis of the Health Information Needs of Pregnant Women[J]. Maternal and Child Health Care of China, 2004, 19(2): 26-28.)
- [8] 赵华,刘晓英,单玉香.住院初诊白血病人健康信息需求及其影响因素[J].护理研究,2008,22(25):2279-2283.(Zhao Hua, Liu Xiaoying, Shan Yuxiang. Health Information Requirement of Preliminary Diagnosed Leukemia Inpatients and Its Influencing Factors[J]. Chinese Nursing Research, 2008, 22(25): 2279-2283.)
- [9] 魏永婷,陈英,许亚红.癌症患者住院化疗期间健康信息需求状况调查分析[J].护理实践与研究,2013,10(11):152-153.(Wei Yongting, Chen Ying, Xu Yahong. Investigation and Analysis of the Health Information Needs Among Patients with Cancer During Chemotherapy in Hospital[J]. Nursing Practice and Research, 2013, 10(11): 152-153.)
- [10] 安娜·德·克利夫兰,潘雪群,陈江萍,等.健康信息需求分析及相关网络资源的使用——对达拉斯福和地区华人的调查[J].图书情报工作,2008,52(3):112-116.(Cleveland A D, Pan D, Chen Jiangping, et al. Analysis of the Health Information Needs and Health Related Internet Usage of a Chinese Population in the United States[J]. Library and Information Service, 2008, 52(3): 112-116.)
- [11] 张馨遥,曹锦丹.网络环境下用户健康信息需求的影响因素分析[J].医学与社会,2010,23(9):25-27.(Zhang Xinyao, Cao Jindan. The Analysis of Influence Factors of Health Information Network Users' Requirement[J]. Medicine and Society, 2010, 23(9): 25-27.)
- [12] 张洪武,冯思佳,赵文龙,等.基于网络用户搜索行为的健康信息需求分析[J].医学信息学杂志,2011,32(5):13-18.(Zhang Hongwu, Feng Sijia, Zhao Wenlong, et al. Analysis of Health Information Needs Based on Network Users Retrieval Behavior[J]. Journal of Medical Informatics, 2011, 32(5): 13-18.)
- [13] 金碧漪,许鑫.社会化问答社区中糖尿病健康信息的需求分析[J].中华医学图书情报杂志,2014,23(12):37-42.(Jin Biyi, Xu Xin. Health Information Needs of Diabetics in Social Q & A Community[J]. Chinese Journal of Medical Library and Information Science, 2014, 23(12): 37-42.)
- [14] 徐倩,赵文龙.基于移动医疗 App 的用户健康信息需求分析[J].现代情报,2015,35(11):79-82.(Xu Qian, Zhao Wenlong. The Research and Analysis on Demands of Health Information for Mobile Health Application[J]. Journal of Modern Information, 2015, 35(11): 79-82.)
- [15] 李小虎,刘熹,田红梅,等.微信时代用户群体健康信息需求调查分析[J].预防医学情报杂志,2017,33(7):689-692.(Li Xiaohu, Liu Xi, Tian Hongmei, et al. Users Concerning About Health Information Needs in the Era of WeChat[J]. Journal of Preventive Medicine Information, 2017, 33(7): 689-692.)
- [16] 吕英杰.网络健康社区中的文本挖掘方法研究[D].上海:上海交通大学,2013.(Lv Yingjie. Research on Text Mining in Online Health Community[D]. Shanghai: Shanghai Jiao Tong University, 2013.)
- [17] Kim S, Oh J S. Evaluating Health Answers in a Social Q&A Site[J]. Proceedings of the American Society for Information Science & Technology, 2008, 45(1): 1-6.

- [18] Zhang Y. Contextualizing Consumer Health Information Searching: An Analysis of Questions in a Social Q&A Community[C]//Proceedings of the 2010 ACM International Health Informatics Symposium. 2010: 210-219.
- [19] Coulson N S. Sharing, Supporting and Sobriety: A Qualitative Analysis of Messages Posted to Alcohol-related Online Discussion Forums in the United Kingdom[J]. Journal of Substance Use, 2014, 19(1): 176-180.
- [20] 李重阳, 翟姗姗, 郑路. 网络健康社区信息需求特征测度——基于时间和主题视角的实证分析[J]. 数字图书馆论坛, 2016, 10(9): 34-42. (Li Chongyang, Zhai Shanshan, Zheng Lu. Measurement of Information Demand Characteristics in Online Health Community: An Empirical Analysis Based on Time and Theme Perspective[J]. Digital Library Forum, 2016, 10(9): 34-42.)
- [21] Oh J S, He D, Jeng W, et al. Linguistic Characteristics of Eating Disorder Questions on Yahoo! Answers—Content, Style and Emotion[J]. Proceedings of the American Society for Information Science & Technology, 2014, 50(1): 1-10.
- [22] Arden M A, Duxbury A M, Soltani H. Responses to Gestational Weight Management Guidance: A Thematic Analysis of Comments Made by Women in Online Parenting Forums[J]. BMC Pregnancy and Childbirth, 2014, 14(1): 216.
- [23] 金碧漪, 许鑫. 网络健康社区中的主题特征研究[J]. 图书馆情报工作, 2015, 59(12): 100-105. (Jin Biyi, Xu Xin. Research on Theme Features in Online Health Community[J]. Library and Information Service, 2015, 59(12): 100-105.)
- [24] International Diabetes Federation. IDF Diabetes Atlas [EB/OL]. [2018-08-13]. <https://www.idf.org/our-activities/advocacy-awareness/resources-and-tools/134:idf-diabetes-atlas-8th-edition.html>.
- [25] Park M S, Park H. Topical Network of Breast Cancer Information in a Korean American Online Community: A Semantic Network Analysis[J]. Information Research: An International Electronic Journal, 2016, 21(4): 729.

### 作者贡献声明:

石静: 论文框架构建, 数据处理及内容撰写;  
厉臣璐, 钱宇星: 数据处理, 内容撰写;  
周利琴: 论文修改;  
张斌: 提出研究思路, 确定研究框架, 论文修改和修订。

### 利益冲突声明:

所有作者声明不存在利益冲突关系。

### 支撑数据:

支撑数据由作者自存储, E-mail: 18844503817@163.com。

[1] 张斌, 石静, 厉臣璐, 钱宇星, 周利琴. 国内问答社区“慢友帮”问答数据。

[2] 张斌, 石静, 厉臣璐, 钱宇星, 周利琴. 国外问答社区“DailyStrength”问答数据。

收稿日期: 2018-07-23

收修改稿日期: 2018-09-12

## Information Needs of Domestic and International HCQA Users ——An Empirical Analysis

Shi Jing<sup>1</sup> Li Chenlu<sup>1</sup> Qian Yuxing<sup>1</sup> Zhou Liqin<sup>1</sup> Zhang Bin<sup>1,2</sup>

<sup>1</sup>(Center for Studies of Information Resources, Wuhan University, Wuhan 430072, China)

<sup>2</sup>(Center of Traditional Chinese Cultural Studies, Wuhan University, Wuhan 430072, China)

**Abstract:** [Objective] This paper identifies and analyzes the information needs of domestic and international question and answer health communities, aiming to find the patterns of regulation evolutions and explore the reasons. [Methods] First, we selected diabetes related data from ManYouBang and DailyStrength. Then, we compared topic evolution and co-occurrence, from the perspectives of theme and time, and using theme coding, social network and content analysis. [Results] The essential needs of HCQA users were “how to treat the disease”. For the chronic disease community, the “diet” theme was closely related to its co-occurrence themes. [Limitations] Our research did not examine the relevance of answers to questions, thus more in-depth study is needed on the topic evolution and content. [Conclusions] Domestic HCQA community is still developing while their foreign counterparts are stable. The former has only “question and answer attribute” while latter has both “Q&A” and “social” attributes.

**Keywords:** HCQA Information Needs Content Analysis Co-occurrence Network