INTERNATIONAL JOURNAL OF MEDICAL INFORMATICS 84 (2015) 24–35

journal homepage: www.ijmijournal.com

The divided communities of shared concerns: Mapping the intellectual structure of e-Health research in social science journals

L. Crystal Jiang[a], *Zhen-Zhen Wang*[b], *Tai-Quan Peng*[c], *Jonathan J.H. Zhu*[b,*]

[a] *Department of Media and Communication, City University of Hong Kong, Kowloon, Hong Kong Special Administrative Region*
[b] *Web Mining Lab, Department of Media and Communication, City University of Hong Kong, Kowloon, Hong Kong Special Administrative Region*
[c] *Wee Kim Wee School of Communication and Information, Nanyang Technological University, Singapore*

ARTICLE INFO

Article history:
Received in revised form
9 September 2014
Accepted 9 September 2014

Keywords:
e-Health
Review
Social sciences
Knowledge
Internet
Technology

ABSTRACT

Purpose: Social scientific approach has become an important approach in e-Health studies over the past decade. However, there has been little systematical examination of what aspects of e-Health social scientists have studied and how relevant and informative knowledge has been produced and diffused by this line of inquiry. This study performed a systematic review of the body of e-Health literature in mainstream social science journals over the past decade by testing the applicability of a 5A categorization (i.e., access, availability, appropriateness, acceptability, and applicability), proposed by the U.S. Department of Health and Human Services, as a framework for understanding social scientific research in e-Health.

Methods: This study used a quantitative, bottom-up approach to review the e-Health literature in social sciences published from 2000 to 2009. A total of 3005 e-Health studies identified from two social sciences databases (i.e., Social Sciences Citation Index and Arts & Humanities Citation Index) were analyzed with text topic modeling and structural analysis of co-word network, co-citation network, and scientific food web.

Results: There have been dramatic increases in the scale of e-Health studies in social sciences over the past decade in terms of the numbers of publications, journal outlets and participating disciplines. The results empirically confirm the presence of the 5A clusters in e-Health research, with the cluster of applicability as the dominant research area and the cluster of availability as the major knowledge producer for other clusters. The network analysis also reveals that the five distinctive clusters share much more in common in research concerns than what e-Health scholars appear to recognize.

Conclusions: It is time to explicate and, more importantly, tap into the shared concerns cutting across the seemingly divided scholarly communities. In particular, more synergy exercises are needed to promote adherence of the field.

* *Corresponding author*. Tel.: +852 34427186.
E-mail address: j.zhu@cityu.edu.hk (J.J.H. Zhu).

http://dx.doi.org/10.1016/j.ijmedinf.2014.09.003
1386-5056/

1. Introduction

The continuous development of new information and communication technologies (ICTs) has greatly changed the ways in which people interact with health information, health professionals and healthcare systems [1,2]. As an emerging field of medical informatics, e-Health refers to the organization, delivery, and consumption of health services and information via the Internet and related technologies. In particular, the concept goes beyond technical development to include "a new way of working, an attitude, and a commitment for networked, global thinking, to improve health care locally, regionally, and worldwide by using information and communication technology" [3]. In keeping up with the burgeoning of e-Health practices in global scale, scholars and practitioners have progressed from debating what e-Health is to examining the technical, human, organizational as well as social factors that influence e-Health practices [4–6]. In the past decade, e-Health research has emerged as an active interdisciplinary field where biomedical science, information science and technology, and social sciences work together synergistically to address issues of common interest [7,8].

Unlike biomedical sciences, computer science and information technology which presumably have close connections with e-Health (i.e., the term "e-Health" heuristically links to biomedical sciences, computer science and information technology), the contributions of social sciences have been overlooked. Very few reviews of e-Health literature have included social science databases except for the one conducted by Pagliari et al. [9]. In fact, driven by the increasing recognition for social and behavioral factors in public health sciences [10,11], there has been substantial attention paid to "merging the science of evidence-based medicine with the practice of user-centered research" [12] including information science, psychology, economics, communication and other social sciences [4]. In the past decade, e-Health research from the social scientific perspectives has grown in magnitude and strength. For example, a recent review of 27,000+ Internet studies published in social science journals found that 27% of the studies were devoted to e-Health issues between 2000 and 2009, which makes e-Health the second most popular research domain in Internet studies, ranked only behind Human–Technology Interactions (34%), but ahead of e-Society (21%) and e-Business (18%) [13]. There is no doubt that the social scientific approach provides an important angle to look at e-Health. However, the questions remain unanswered on which topics are pursued by this approach, how these foci are rising or falling in popularity, and what the observable and hidden relationships exist between these research concerns.

Early attempts that seek to map out the scope of the e-Health field have proposed several taxonomies of e-health research. For example, Eng came up with a 5C model and Eysenbach used 10 essential Es to characterize e-Health [3,14]. Richardson discussed the development of e-Health in Europe and claimed clinical applications, healthcare professional continuing education, public health information, and health policy development as four main pillars for e-Health practices [15]. The U.S. Department of Health and Human Services (DHHS) proposed a category of 5A research categories as an analytical framework for the social aspect of e-Health research [16]. However, these taxonomies suggested by field leaders and national agencies are more like normative guidelines of what they think the field should be than the empirical reality of what the field of e-Health actually looks like. There is no clear-cut evidence that the taxonomies sufficiently capture the knowledge inquiry in e-Health. Until recently, several critical reviews and scoping exercises (e.g., special issue of American Journal of Preventive Medicine on e-Health research in 2007) take the initiative to critically examine what the field has already become [4,7,9,17–19]. These reviews have provided qualitative evaluations of the field regarding its research themes, perspectives, methodology and practical implications.

This study expands previous efforts to systematically analyze the research scale, key themes, intellectual structure and knowledge production pattern of e-Health related studies in social sciences for the past decade. It particularly investigates whether this field has converged a set of originally disconnected colonies into a disciplinary structure. Such convergence directly reflects the maturity of a field and has been seen in a number of emerging interdisciplinary fields such as artificial intelligence, nanotechnology, entrepreneurship, communication studies, and innovation studies [20,21].

The study tests the applicability of DHHS's 5A category as a framework for understanding social research in e-Health. The 5As include acceptability, access, applicability, appropriateness, and availability. The 5As, although not explicitly stated in the original framework, are framed as sequential steps of implementation in e-Health practices [16]. The convergence of the e-Health field should be reflected in a recognizable knowledge structure with studies on each A integrally connected to one another (see Fig. 1). Specifically, the *access* theme focuses on the technical access of e-Health systems, such as digital divided and adoption of general information and communication technologies. This issue underlies all e-health practices and is the necessary but insufficient factor for e-Health effectiveness [22]. The *availability* theme addresses the development of meaningful access (i.e., having the tools people want and need), mainly concerning information accessibility and information seeking pattern related to a variety of e-Health tools. Given that different population groups experience disproportionate amounts of diseases and health issues, it is important to design and deploy e-Health access and availability in accordance to population characteristics [5,22,23]. The *appropriateness* theme considers the objective fit between diverse user needs and the e-Health practices, and specific foci include cultural appropriateness, users' perceptions of content credibility, information quality and readability, and the use of tailoring. The theme of *acceptability*, by contrast, is concerned with users' subjective evaluations of tools, including ease of use, satisfaction, usage over time and usability. Last but not least, the *applicability* theme builds on the previous four As and attends to impact and outcomes of e-Health practices, promoting changes in knowledge, beliefs and attitudes, social support and health behaviors in different health contexts.

Unlike previous qualitative, top–down reviews, this study adopts a quantitative, bottom–up approach to examine the body of e-Health literature in social sciences, which minimizes the influence of a priori assumptions and arbitrarily-defined

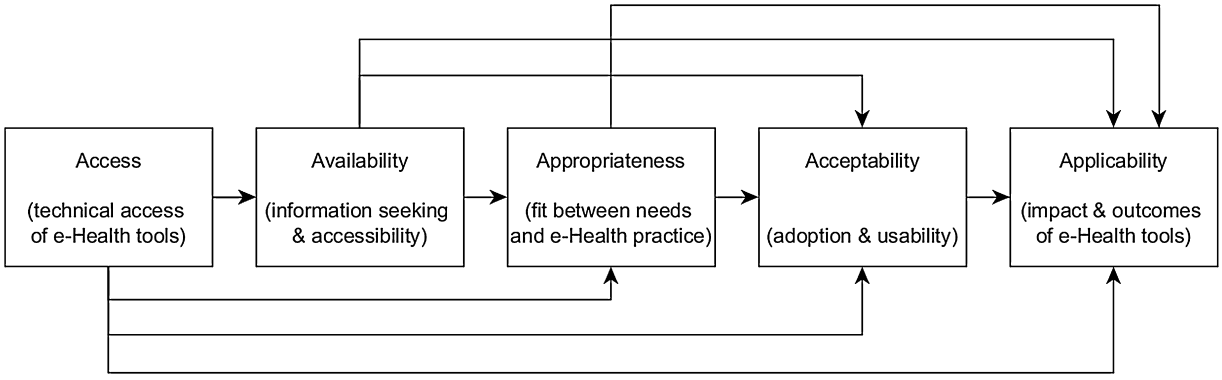

Fig. 1 – Expected interrelationships among 5A clusters.

categorizations on findings. Specifically, this study investigates two databases of social science journals (i.e., Social Sciences Citation Index, or SSCI, and Arts & Humanities Citation Index, or A&HCI) with a mix of several novel methods, including topic modeling, and network analysis of research vocabularies, co-citations as well as scientific food web. These quantitative methods have recently become popular in studies of scientific literature, such as bibliographic analysis of academic journal articles [24] and systematic reviews of scientific disciplines [25].

2. Method

2.1. Data

The study is an in-depth analysis of a subset data, used in a previous study that reviewed 27,340 Internet studies from SSCI and A&HCI of ISI Web of Science published from 2000 to 2009 [13]. In that study, through analyzing article titles, abstracts, and keywords, four primary domains in Internet research were extracted: e-Health, e-Business, e-Society, and Human–Technology Interactions. However, as an overview of the entire field of Internet studies, the authors did not, understandably, zoom in on the inner structure of any domain, including e-Health. The current study departed from that study by analyzing the e-health studies ($n = 6926$) and describing the research scale, foci and intellectual structure of e-Health research in social sciences.

In Peng et al.'s [13] review, Internet studies were defined as those studies that consider the Internet either as a means (i.e., the Internet is a new platform for data collection) or as an end (Internet-relevant variables are substantive variables in a study). The current study, however, is only interested in the studies that substantially focused on online health issues. Hence, the studies that used Internet for recruitment or other methodological purposes were dropped. To do so, a longer list of Internet-related keywords were searched for in titles and keywords, including Internet, web, cyberspace/cyber-space, online/on-line, technology, computer, network, forum, electronic, tele (telephone excluded), mobile, digital, email/e-mail, messaging/messenger, virtual, mediated, channel, e-, e-Health, informatics, information system, and etc. The studies whose title and keywords did not include any Internet-related words were dropped, resulting in a final sample of 3005 articles qualified as e-Health studies. In the dataset, each article was listed with its title, authors, published year, author-supplied keywords, abstract, publication name, publication's subject category, references cited and other parameters provided by SSCI and A&HCI.

2.2. Analytical procedure

The analysis followed a three-step procedure: topic extraction, topic coding, and article labeling. Specifically, Latent Dirichlet allocation (LDA) [26] was employed to extract topics from the 3005 studies. This is a bottom–up approach that treats each article abstract as a mixture of latent topics and identifies clusters of topics based on word occurrences across abstracts. R package "topicmodels" [27] was used for performing the topic modeling procedure (details on this methods are provided as supplementary material). A total of 60 topics were extracted (see Table 1 for the 10 most frequently occurring words for each topic). Based on the most frequently occurring words associated with each topic, three coders independently assigned the 60 topics to one of the 5A clusters (i.e., acceptability, access, applicability, appropriateness, or availability). The initial overall inter-coder reliability (Krippendorff's alpha) was .75 and the pairwise reliability (Cohen's kappa) ranged from .68 to .91. Disagreements were discussed among the coders until 100% agreement was reached. Of the 60 initially extracted topics, 12 (e.g., about research methods, healthcare IT in medical institutions; involving 18.4% of the articles) were coded as irrelevant to any of the 5A clusters and thus dropped from the subsequent analysis. Because topic modeling identified the most prominent topic for each article abstract, each article was thus exclusively assigned to the cluster of acceptability, access, applicability, appropriateness, or availability. The words used in abstracts were aggregated by cluster and the top 50 most popular words in each cluster were used to generate word cloud (see Fig. 2).

Cluster-level analysis was further performed to examine the intellectual structure among the five clusters. The analysis compared the cluster similarity by looking at the extent to which different clusters used similar words to describe studies (called "co-word") and the extent to which they are cited by same articles (co-citation). Specifically, co-word was calculated based on vocabularies (i.e., number of unique

Table 1 – Top 10 frequently occurring words for each topic.

Domain	Topic	Topic label	Top ten words
Access	1	Medical service	Service, doctor, consultation, suicide, delivery, tele-Health, consult, nhs, accessible, anonymous
	2	Healthcare	Nurse, facility, home, professional, nurses, care, resident, work, available, discharge
	3	Access	Access, barrier, accessed, accessing, free, equipment, ability, economic, disable, chess
	4	Learning	Learn, course, professional, skill, knowledge, teach, teacher, educational, recognize, posttest
	5	Adoption	Evidence, clinician, practitioner, adoption, clinical, evidence-based, ehr, confidence, available, gps
	6	Digital divide	Rural, feasibility, pilot, delivery, telemedicine, acceptability, usability, urban, stroke, videoconferencing
	7	Education	Education, knowledge, educational, awareness, cohort, curriculum, retention, resident, educate, breastfeeding
Availability	8	Database	Search, database, engine, genetic, source, query, evidence-based, google, poor, link
	9	Self help	Seek, source, disease, available, heart, hivaids, role, seeker, information-seeking, channel
	10	Chat	Chat, united, room, states, read, grade, literacy, readability, education, society
	11	Website	Web, site, page, presence, america, sites, proeating, commercial, available, die
	12	Website	Website, promote, log, link, campaign, http, accessible, presence, epilepsy, bandwidth
Appropriateness	13	Elderly	Adult, age, older, young, younger, old, educate, employ, hip, middle-aged
	14	Parents	Child, family, parent, sleep, mother, parental, pediatric, parenting, father, parents
	15	Doctor–patient	Patient, physician, clinic, patients, portal, clinical, doctor-patient, physicians, medical, demand
	16	Race/ethnicity	American, digital, white, income, african, minority, race, ethnic, divide, americans
	17	Regional (adoption)	Public, policy, country, privacy, canada, department, domain, surveillance, canadian, funded
	18	Gender groups	Gender, male, female, age, cross-sectional, stereotype, cafe, males, orientation, occupation
	19	Women	Woman, email, men, women, pregnancy, age, pregnant, longitudinal, menopause, postpartum
	20	Service evaluation	Care, provider, primary, cost, economic, patient-provider, efficient, register, secure, delivery
	21	Training	Training, skill, disability, injury, rehabilitation, ability, brain, train, tbi, traumatic
Acceptability	22	Monitoring	Monitor, safety, encourage, acceptance, vaccine, device, compliance, class, hepatitis, vaccination
	23	Utility	Preference, utility, language, profile, visual, protocol, pc, divide, book, express
Applicability	24	Pornography	Youth, sexual, exposure, pornography, survey, harassment, solicitation, age, victimization, symptomatology
	25	STD	Men, sex, sexual, partner, HIV, msm, risk, gay, anal, sexually
	26	Internet addition	Addiction, disorder, personality, diagnostic, addict, psychiatric, adhd, excessive, interpersonal, compulsive
	27	Cancer	Cancer, breast, screening, survivor, prostate, oncology, mailing, diagnose, aid, icsgs
	28	Substance use	Drug, medication, substance, abuse, prescription, availability, pharmacy, treat, ecstasy, illicit
	29	Risk prevention	Risk, prevention, counseling, pressure, braden, preventive, disease, demographic, fit, ulcer
	30	Bullying	Adolescent, school, bully, girl, peer, boy, cyberbullying, victim, medicine, teen
	31	Drinking	Alcohol, drink, feedback, consumption, drinker, heavy, screening, intervention, audit, alcohol-related

Table 1 (*Continued*)

Domain	Topic	Topic label	Top ten words
	32	Social network	Social, network, loneliness, relationship, friend, self-esteem, networking, peer, interpersonal, friendship
	33	Interaction	Message, emotional, board, post, emotion, word, posting, expression, bulletin, text
	34	Gambling	Forum, gamble, screen, gambler, hearing, pathological, harmful, bet, gum, sample
	35	Diabetes	Diabetes, maintain, utilization, adherence, utilize, observe, coach, blood, elderly, cellular
	36	Chronic illness	Illness, chronic, pain, story, narrative, news, fatigue, syndrome, hta, adapt
	37	Physical exercises	Physical, tailor, exercise, self-reported, lifestyle, goal, arm, leisure, feedback, sedentary
	38	Relationship health	Relationship, phone, mobile, messaging, couple, text, friend, instant, call, infidelity
	39	Depression	Depression, symptom, anxiety, depressive, mood, depress, stigma, moodgym, primary, help-seeking
	40	Weight loss	Weight, loss, food, body, nutrition, intake, diet, eat, obesity, dietary
	41	Gaming	Game, play, role, video, player, bed, gamers, sleep, mmorpgs, clip
	42	Mental health	Work, stress, employee, job, distress, worker, workplace, recovery, workforce, productivity
	43	Wellbeing	Wellbeing, task, identity, sense, disclosure, felt, experiment, self, engagement, anonymity
	44	Cognitive disorder	Cognitive, caregiver, family, schizophrenia, ability, memory, impairment, control, speed, cognition
	45	Smoking	Smoke, cessation, smoker, tobacco, abstinence, cigarette, nicotine, dependence, efficacy, engagement
	46	Mental health	Treatment, therapy, disorder, self-help, cbt, therapist, phobia, cognitive, therapeutic, control
	47	Posttrauma	Mental, client, professional, ethical, belief, offender, legal, psychologist, code, psychiatrist
	48	Empower	Support, empower, supportive, cmsg, accessible, poster, listserv, fibromyalgia, ireland, efficient
Dropped	49	Questionnaire	Questionnaire, response, structure, fear, test-retest, sensitivity, sample, self-report, dutch, dental
	50	Survey	Survey, counsel, software, us, counselor, career, available, infertility, goal, appropriate
	51	Survey	Survey, email, response, mail, telephone, postal, randomly, reminder, responder, respondent
	52	Gender groups	Gender, male, female, age, cross-sectional, stereotype, cafe, males, orientation, occupation
	53	Qualitative	Interview, qualitative, diagnosis, quantitative, headache, indepth, diary, facilitator, telephone, net
	54	Quantitative/experiment	Group, control, experimental, randomly, volunteer, empowerment, pd, relaxation, exercise, urban
	55	Mixed topics	Attitude, self-efficacy, sample, usefulness, posttest, pretest, relationship, taiwan, phns, cross-sectional
	56	Quantitative/experiment	Intervention, control, trial, randomized, efficacy, randomly, confidence, randomized, registration, primary
	57	College students	Student, college, university, undergraduate, chinese, graduate, campus, freshman, mark, express
	58	Hospitalization	Medical, hospital, professional, care, library, medicine, trust, available, acute, librarian
	59	Clinical system	System, clinical, call, security, telerehabilitation, speech, informatics, infrastructure, protection, cpoe
	60	Marketing	Expert, promotion, market, advertise, television, tv, asthma, aor, ad, multimedia

words) shared by clusters, and co-citation patterns were analyzed by aggregating the co-citation links at article level to those at cluster level. The inter-citation patterns were extracted and interpreted in light of the scientific food web perspective [28] to develop an ecosystem structure for e-Health. The five clusters were connected by the extent to which they provided knowledge (being cited) to and consumed knowledge (citing others) from each other. For each pair of clusters, there was a knowledge source (being cited more frequently than citing the other cluster) and a knowledge receiver (citing the other cluster more

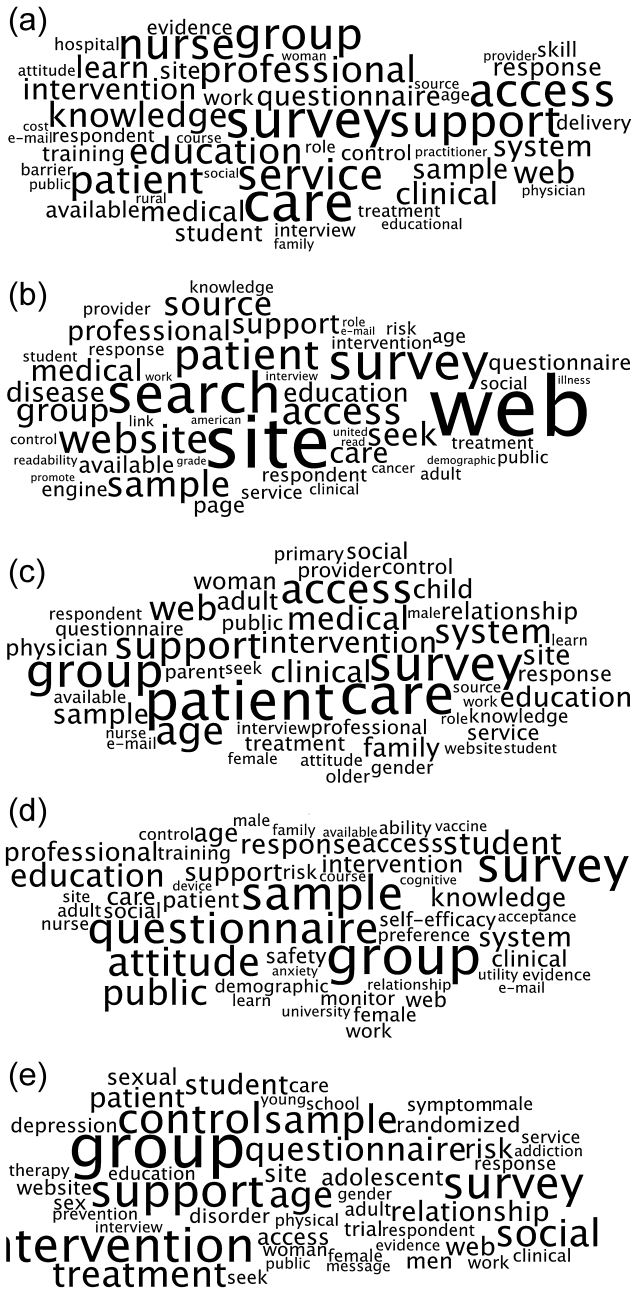

Fig. 2 – Word clouds for the 5A clusters.

frequently than being cited) based on the net flow between the two.

3. Results

3.1. *Article-level analysis*

Fig. 3 shows the year-to-year distribution of the e-Health publications from 2000 to 2009. The number of e-Health studies has experienced an exponential growth during the period, increasing from 102 in 2000 to 664 in 2009 (average growth rate = 23.1%). Clearly, e-Health studies have gained increasing attention from social sciences in the past decade.

The 3005 articles identified on e-Health appear in a total of 756 social science journals from 2000 through 2009. These journals span 123 subject categories in the 2011 Journal Citation Reports (JCR). The subject categories were further collapsed by general discipline. For the journals that cover more than one subject categories, they were listed under multiple categories with a weighted contribution. It turns out that 54% of the journals are related to Health and Medicine Sciences, 22% Psychology, 10% Information Science and Communication, and 14% other disciplines (e.g., Biology, Engineering, Business & Management, etc.), which indicates e-Health as an interdisciplinary field with Health and Medicine playing a dominant role. Among all the journal outlets, *Cyberpsychology & Behaviors* has published the most articles ($N = 246$), followed by *Journal of Medical Internet Research* ($N = 110$), *Computers in Human Behavior* ($N = 63$), *Patient Education & Counseling* ($N = 41$), *Journal of the American Medical Informatics Association* ($N = 40$), *Journal of Health Communication* ($N = 37$), *CIN-Computers Informatics Nursing* ($N = 36$), *Journal of Adolescent Health* ($N = 36$), *Journal of Advanced Nursing* ($N = 29$) and *Health Education & Research* ($N = 27$). About 40% of the identified e-Health articles were published in these top 10 journals. Each journal on average published 4 e-Health articles (SD = 11, skewedness = 16.20). Such highly right-skewed distribution follows a common pattern in many other fields that there are a few regular homes of e-Health and many occasional shelters.

Citation analysis was performed to identify the most influential publications that have laid down the groundwork of the e-Health field. Table 2 presents the top 20 publications cited by the 3005 e-Health studies. The most frequently cited is Kraut's [29] paper *Internet paradox: A social technology that reduces social involvement and psychological wellbeing*, followed by Tate's [30] *Using Internet technology to deliver a behavioral weight loss program*, and Baker's [31] *Use of the Internet and e-mail for healthcare information*. More than half of the top 20 citations were published in Health and Medicine journals while about one third published in social sciences journals, particularly psychology journals. Only one of the 20 was a theoretical reference [Bandura's social cognitive theory, 32] whereas other 16 of them were empirical papers concerned with specific health practices (e.g., health information seeking, Internet addiction, e-interventions). These findings seem to suggest that e-Health studies are highly oriented to evidence accumulation.

As mentioned earlier, the analysis identified a primary research cluster for each article based on its topic loadings on the 5A clusters. More than half of the e-Health studies have focused on applicability (56%), followed by appropriateness (20%), access (13%), availability (9%) and acceptability (3%). While one might expect to see that the applicability cluster received a lot of attention in social sciences, it was surprising that it dominated the body of literature to such a great extent. Fig. 4 further illustrates the annual research outputs in normalized percentages across research clusters. The applicability cluster has received increasing attention over years, but the research outputs produced in the other four research clusters remained relatively constant or slightly decreased by year. The analysis also considered the contributions of different disciplines to the 5A clusters. However, no significant difference was detected across cluster: Health and Medicine Sciences contributed the most, followed by Psychology, Information

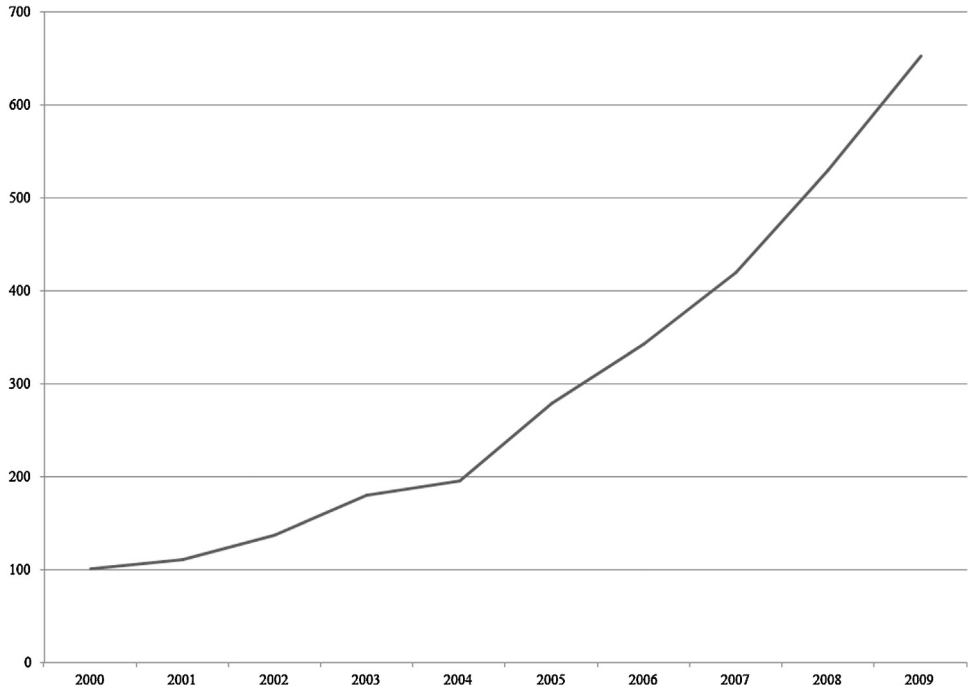

Fig. 3 – Year-to-year distribution of e-Health studies in social sciences from 2000 to 2009.

Science and Communication, and other disciplines, all in proportion to their respective shares of the articles.

3.2. *Cluster-level analyses*

Intellectual structure of e-Health was further mapped out by cluster-level analyses. The similarity among five clusters was reflected in the structures of co-word network and co-citation network (see Fig. 5). The thickness of edges indicated the number of co-word (Fig. 5a) or co-citation (Fig. 5b) shared between clusters. The resulting co-word network was tightly connected with thick edges. The thickest edge appeared between appropriateness and access and the thinnest edge appeared between availability and applicability. Such cohesion suggests that different e-Health clusters share a common set of vocabulary and have similar research concerns. The co-citation

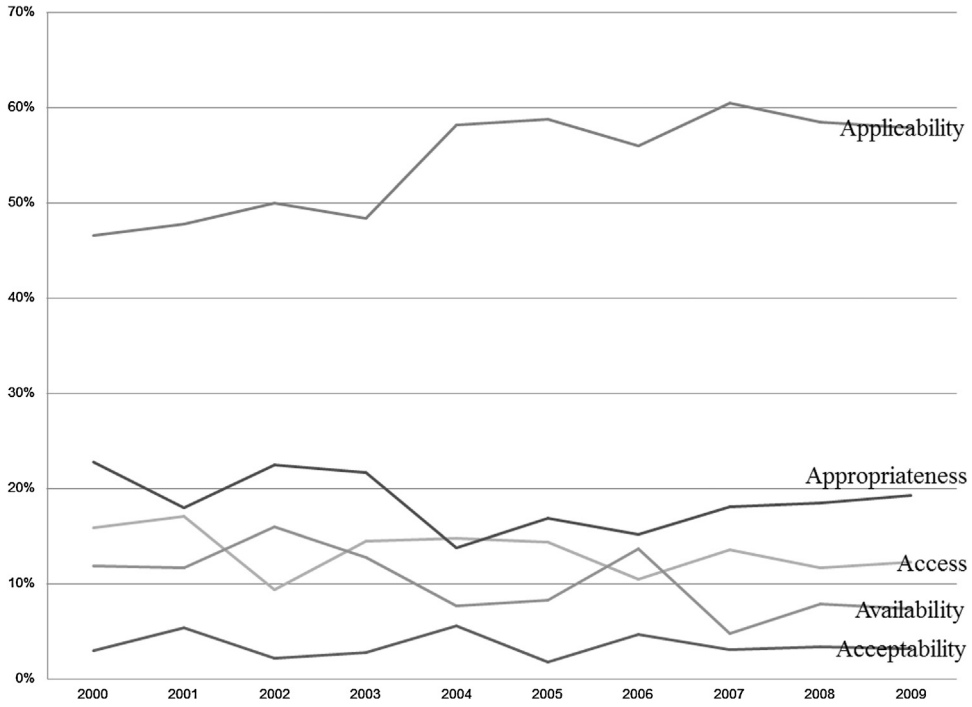

Fig. 4 – Year-to-year distribution of e-Health studies across five research clusters in normalized percentages.

Table 2 – The most frequently cited publications in e-Health: 2001–2009.

Citations	Full citation index	Discipline/field	Type
181	Kraut R, 1998, Internet paradox. A social technology that reduces social involvement and psychological well-being? Am Psychol, V53, P1017	Social sciences	Practice
94	Tate DF, 2001, Using Internet technology to deliver a behavioral weight loss program. JAMA-J Am Med Assoc, V285, P1172	Medicine	Practice
88	Baker L, 2003, Use of the Internet and e-mail for health care information: Results from a national survey. JAMA-J Am Med Assoc, V289, P2400	Medicine	Practice
86	Cohen J, 1988, Statistical power analysis for the behavioral sciences	Methodology	n/a
81	*Am Psych Ass, 1994, Diagnostic and Statistical Manual of Mental Disorder	Methodology	n/a
80	Eysenbach G, 2002, What is e-health? JAMA-J Am Med Assoc, V287, P2691	Medicine	Review
76	Morahanmartin J, 2000, Incidence and correlates of pathological Internet use among college students. Comput Hum Behav, V16, P13	Social sciences	Practice
74	Berland GK, 2001, Health information on the Internet: accessibility, quality and readability in English and Spanish. JAMA-J Am Med Assoc, V285, P2612	Medicine	Practice
72	Bandura A, 1986, Social Foundations of Thought and Action.	Social sciences	Theory
71	Eysenbach G, 2002, Using the Internet for surveys and health research. Brit Med J, V324, P573	Medicine	Practice
66	Christensen H, 2004, Delivering interventions for depression by using the Internet: randomised controlled trial. Brit Med J, V328, P265	Medicine	Practice
65	Young K, 1998, The relationship between depression and Internet addiction. Cyberpsychol Behav, V1, P237	Social sciences	Practice
64	Kraut R, 2002, Internet paradox revisited. J Soc Issues, V58, P49	Social sciences	Practice
63	Mcfarlane M, 2000, The Internet as a newly emerging risk environment for sexually transmitted diseases. JAMA-J Am Med Assoc, V284, P443	Medicine	Practice
62	Tate DF, 2003, Effects of Internet behavioral counseling on weight loss in adults at risk for type 2 diabetes: a randomized trial. JAMA-J Am Med Assoc, V289, P1833	Medicine	Practice
61	Sharf BF, 1997, Communicating breast cancer on-line: support and empowerment on the Internet. Women Health, V26, P65	Medicine	Practice
58	Fox S, 2003, Internet health resources. Pew Internet & American Life Project.	Social sciences	Practice
56	Cline RJW, 2001, Consumer health information seeking on the Internet: the state of the art. Health Educ Res, V16, P671	Medicine	Practice
56	Silberg WM, 1997, Assessing, controlling, and assuring the quality of medical information on the Internet: Caveant lector et viewor–Let the reader and viewer beware. JAMA-J Am Med Assoc, V277, P1244	Medicine	Practice
56	Winzelberg AJ, 2000, Effectiveness of an Internet-based program for reducing risk factors for eating disorders. J Consult Clin Psych, V68, P346	Social sciences	Practice

network, by contrast, was much less balanced. The clusters of applicability and appropriateness most frequently co-cited similar literature, followed by the pair of applicability and availability, and the pair of applicability and access. The co-citations among the other four clusters were much weaker. In particular, acceptability has very limited citation traffic with other clusters. As an indication of scholars' unconscious behavior, the tightly connected co-word network implies that the five e-Health clusters were not distinct from each enough in terms of research concerns. Meanwhile, as an indication of scholars' conscious beliefs about cluster similarity, such co-citation pattern reveals that the five e-Health clusters were not aware of how similar they are.

The resulting knowledge chain of e-Health, based on scientific food web analysis, is presented in Fig. 6. Note that the bars in Fig. 6 indicate the net knowledge flow between every two clusters with higher bars showing greater influence. The cluster of availability, although only taking up a small portion (9%) of the e-Health studies, was the leading knowledge source in the field. It produced knowledge for applicability (i.e., cited by

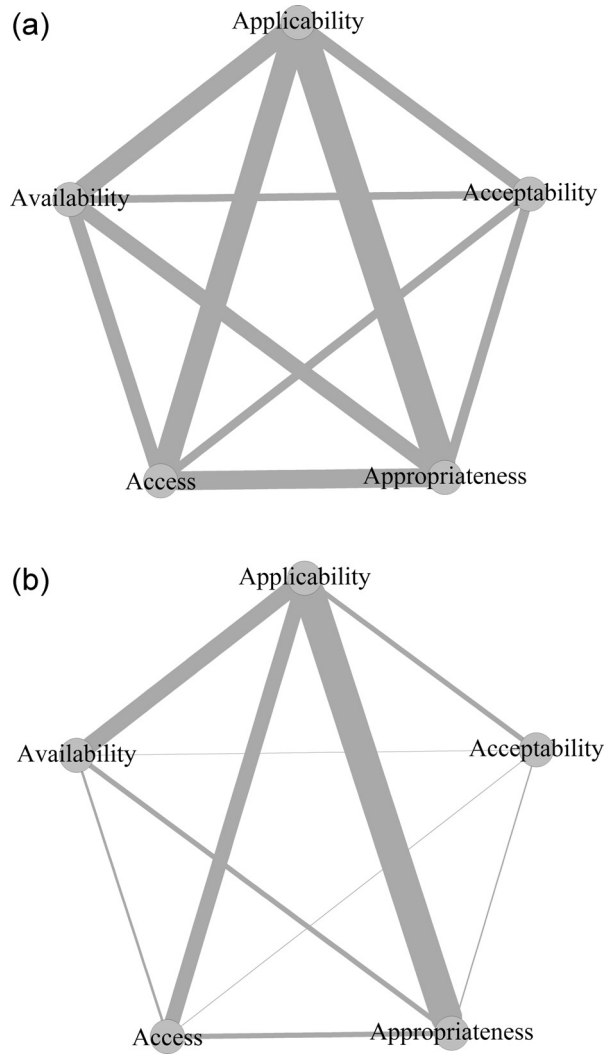

Fig. 5 – Cluster similarity.

applicability research), which feeds in turn appropriateness, acceptability and access studies, in that order.

4. Discussion and conclusion

This study adopts novel text mining techniques to map out the landscape and intellectual structure of e-Health scholarship by analyzing 3005 articles on e-Health published in SSCI and A&HCI journals during 2000–2009. To our best knowledge, this is the first review of the field that seeks for understanding a large body of e-Health literature with bottom-up exercises. The findings enrich the understanding of knowledge production and diffusion in e-Health field and raise important issues for future inquiries.

The past decade has been a period of substantial growth for e-Health research in social sciences, as illustrated by a dramatic increase in the number of publications and a diverse range of publication venues in different disciplines. E-Health research carried out by social sciences could roughly be mapped onto the five primary research clusters as identified by previous research [16], including access, availability, appropriateness, acceptability and applicability. Of them, applicability appears to be the dominant research area, followed by appropriateness, access, availability and acceptability. The 5A clusters indeed share many common interests, as indicated by a coherent co-word network. However, such shared common ground is not recognized in scholarly communication. The co-citation network is unbalanced, with applicability having much stronger citation links with each of other four clusters as compared to the links among the four clusters. The cluster of availability, which concerns acquisition, retrieval, and exchange of health and lifestyle information, is identified as the major knowledge producer. This confirms Pagliari et al.'s conclusion that e-Health was not discernible from the wider field of health informatics at its infant stage and was mainly conceptualized as a broad range of medical informatics applications that facilitate the management and delivery of healthcare [9].

Disciplinary integration is considered as key to interdisciplinarity [33], and transdisciplinarity is the fully synthesized stage where research communities share a common language and stand together as mutually informative networks [34]. The wide scope and complexities of e-Health scholarship definitely signal a growing impact. However, as pinpointed by Ahern, the diversity could present challenges for synthesis and integration [19]. Indeed, the current study demonstrates a significant lack of communication, as reflected in citation traffic, among different research clusters. These five research clusters are more like emerging colonies that have not been adequately networked into a new specialty. The unawareness of shared research concerns is likely to generate silo or parallel working, which leads to the ignorance of network potentials and harms the pursuit of research efficiency and quality in the longer term. Pagliari has similarly noted that the increasing heterogeneity in e-Health would pose significant barriers to interdisciplinary research collaboration and to the translation of research to policy and practice [4]. Hence, there remains a strong need to promote adherence of the field, not only in terms of shared terminologies, theoretical bases and methodologies but also in terms of established standards, values and principles in knowledge production. From a managerial point of view, more institutionalized efforts may help increase field synergy by establishing academic journals, specialized conferences, professional associations and joint programs that invites scholars from different research clusters and perspectives. For example, research on acceptability can be potentially enhanced and bridged by connecting e-Health scholars in information science to those in social sciences [35].

Theory should presumably play a substantial role in guiding scientific inquiry and defining the identity of a discipline. The convergence of artificial intelligence, nanotechnology, and communication studies, for example, involves using theory and core literature to distinguish themselves from parental or neighboring disciplines [21]. In the case of e-Health field, despite its increasing breadth and complexity, it is fair to say that the field so far is rarely theory-driven and has yielded little theoretical novelty. The lack of theoretical orientation has been documented in Peng et al.'s general review of Internet studies, where they found the Internet studies in e-Health were the least theoretically driven in comparison to the studies in e-Business, e-Society, and Human–Technology

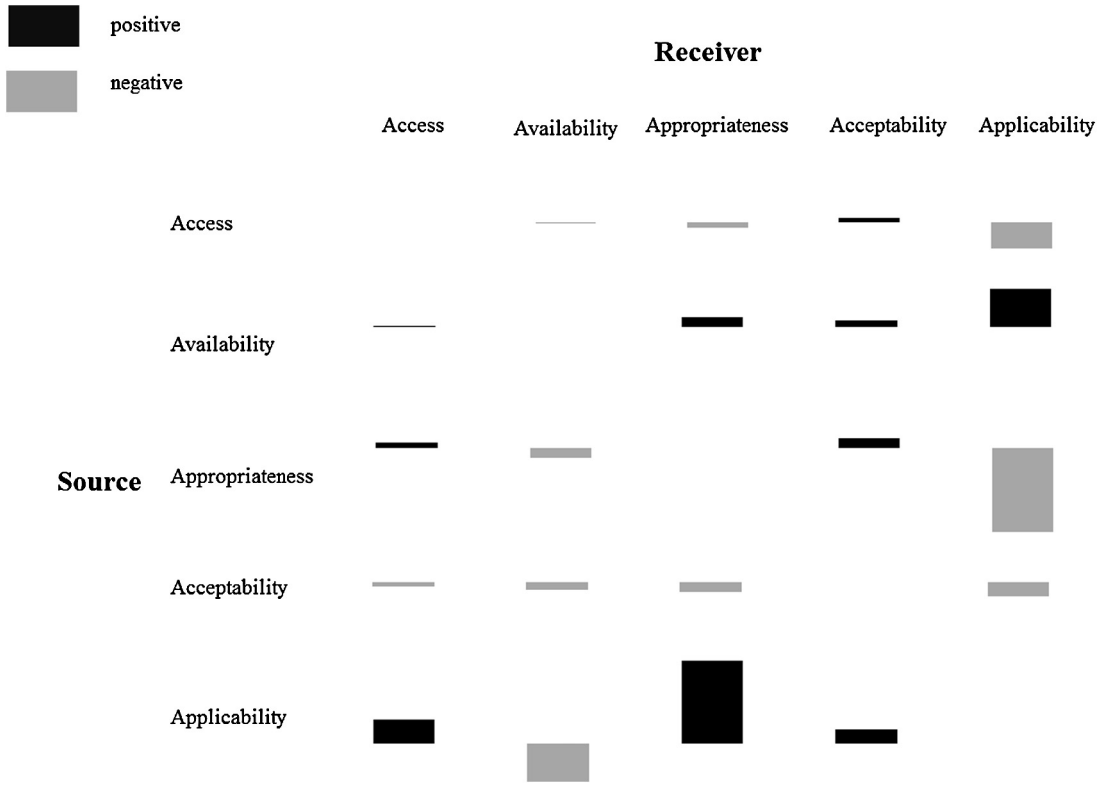

Fig. 6 – Relative knowledge flows among e-Health clusters.

Interactions. In the present study, theory-related content did not merge as an independent topic of the 60 topics identified by topic modeling. Only one of the top 20 citations appeared to be a theoretical piece, which is a classical social science theory. The authors speculate that this is closely allied to the evidence-based preference in Health and Medicine Sciences, which seeks for practical decisions that work for specific groups in specific contexts. The lack of theory has contributed to the lack of synergies mentioned above, and also presents serious challenges to e-Health practices, as theory provides guiding frameworks for practice design, delivery, and evaluation [19,36,37]. In future research, it is particularly important for researchers to ascertain the conceptual or theoretical frameworks for high standards of research design and practice implementation.

The findings also reveal the true nature of e-Health scholarship – health issues in online contexts. Technology-related terms did not emerge as principal research topics; instead, general health practices (e.g., monitoring, prevention, and education), targeted diseases and risk factors (e.g., cancer, substance use, and gaming) and targeted population groups (e.g., women, elderly, and ethnic groups) are identified as research concerns. Although the data in this study only include e-Health literature in social sciences, this result is aligned with Oh et al.'s argument that e-Health field considers technology "as a means to expand, to assist, or to enhance human activities, rather than as a substitute for them" [38]. The literature reviewed here tends to view e-Health as embodiments of social processes that affect the health behaviors and health outcomes of different populations.

4.1. *Limitations*

As with all studies the current investigation has limitations. First, the findings are potentially limited by the varying quality of the abstracts and keywords of the e-Health articles under study. While most abstracts are well-written, in terms of structured formats, some e-Health articles provided little substantive information in the abstracts or did not provide keywords. A more informative (and of course significantly more costly) approach would be to mine the full text of the articles. Second, this investigation approached knowledge production via the product of scholarly practices, not the practices per se, although the product is a good reflection of the phenomena. Personal experience and thoughts of e-Health scholars are also important for understanding the "divided communities of shared concern". Bibliometric methods should be complemented with ethnographic methods or interviews to develop a deeper understanding of the e-Health field [39]. Finally, this investigation is limited by its study period (2000–2009). Future review exercises should target to capture more recent dynamics in e-Health scholarship in social sciences.

4.2. *Conclusions*

The study examines the knowledge structures of social e-Health research over the past decade with a quantitative bottom–up approach. There is strong evidence that the burgeoning social e-Health research is primarily driven by five research clusters that have not integrally connected to one

Summary points

What was already known on the topic

- e-Health has been broadly defined as a new way of integrating various resources to serve health care via the Internet and related technologies.
- Human and social factors are highlighted in relation to many aspects of e-Health.
- Social sciences has become an important approach to study the field of e-Health.

What this study added to our knowledge

- Social e-Health research has experienced an exponential growth in terms of the numbers of publications, journal outlets and participating disciplines during 2000–2009.
- The quantitative bottom–up analyses empirically confirm the presence of the 5A clusters in social e-Health research, with the cluster of applicability as the dominant research area and the cluster of availability as the major knowledge producer for other clusters.
- The five distinctive research clusters share a common set of vocabulary and have similar research concerns but such cohesion has not been recognized by the knowledge flow among clusters.

another, calling for more synergy exercises to promote adherence of the field.

Author contributions

The four authors are justifiably credited with authorship, according to the authorship criteria. In detail:

L. Crystal Jiang – conception, analysis and interpretation of data, drafting of the manuscript, final approval given; Zhenzhen Wang – acquisition of data, analysis and interpretation of data, drafting of the methods, final approval given; Tai-Quan Peng – acquisition of data, critical revision of manuscript, final approval given; Jonathan J. H. Zhu – conception, analysis and interpretation of data, critical revision of manuscript, final approval given.

Competing interests

The authors declare that they have no conflict of interest for this study.

Acknowledgements

The study was supported in part by two Faculty Research Grants (Grant no: 9610249 and 9610292) from Department of Media and Communication, City University of Hong Kong.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ijmedinf.2014.09.003.

REFERENCES

[1] S. Fox, M. Duggan, Health Online 2013, Pew Center, 2013.
[2] C.A. Meier, M.C. Fitzgerald, J.M. Smith, Ehealth: extending, enhancing, and evolving health care, Annu. Rev. Biomed. Eng. 15 (2013) 359–382.
[3] G. Eysenbach, What is e-health? J. Med. Internet Res. 3 (2001) e20.
[4] C. Pagliari, Design and evaluation in eHealth: challenges and implications for an interdisciplinary field, J. Med. Internet Res. 9 (2007) e15.
[5] S.T.M. Peek, E.J.M. Wouters, J. van Hoof, K.G. Luijkx, H.R. Boeije, H.J.M. Vrijhoef, Factors influencing acceptance of technology for aging in place: a systematic review, Int. J. Med. Inform. 83 (2014) 235–248.
[6] S. De Rouck, A. Jacobs, M. Leys, A methodology for shifting the focus of e-health support design onto user needs: a case in the homecare field, Int. J. Med. Inform. 77 (2008) 589–601.
[7] A.A. Atienza, B.W. Hesse, T.B. Baker, D.B. Abrams, B.K. Rimer, R.T. Croyle, L.N. Volckmann, Critical issues in ehealth research, Am. J. Prev. Med. 32 (2007) S71–S74.
[8] M. Murero, R. Rice, E-health research, in: M. Murero, R. Rice (Eds.), The Internet and Health Care: Theory, Research and Practice, Lawrence Erlbaum Associates, New Jersey, 2006, pp. 3–26.
[9] C. Pagliari, D. Sloan, P. Gregor, F. Sullivan, D. Detmer, P.J. Kahan, W. Oortwijn, S. MacGillivray, What is eHealth (4): a scoping exercise to map the field, J. Med. Internet Res. 7 (2005) e9.
[10] C.A. Bachrach, R.P. Abeles, Social science and health research: growth at the National Institutes of Health, Am. J. Public Health 94 (2004) 22–28.
[11] D. Mechanic, Emerging trends in the application of the social sciences to health and medicine, Soc. Sci. Med. 40 (1995) 1491–1496.
[12] B.W. Hesse, B. Shneiderman, eHealth research from the user's perspective, Am. J. Prev. Med. 32 (2007) S97–S103.
[13] T.-Q. Peng, L. Zhang, Z.-J. Zhong, J.J. Zhu, Mapping the landscape of Internet studies: text mining of social science journal articles 2000–2009, New Media Soc. 15 (2012) 644–664.
[14] T. Eng, The e-Health Landscape – A Terrain Map of Emerging Information and Communication Technologies in Health and Health Care, The Robert Wood Johnson Foundation, Princeton, NJ, 2001.
[15] R. Richardson, eHealth for Europe, Stud. Health Technol. Inform. 96 (2003) 151–156.
[16] ODPHP, Expanding the Reach and Impact of Consumer e-Health Tools, Office of Disease Prevention and Health Promotion, 2006.
[17] J. Li, A. Talaei-Khoei, H. Seale, P. Ray, R.C. MacIntyre, Health care provider adoption of ehealth: systematic literature review, Interact. J. Med. Res. 2 (2013) e7.
[18] K.D. Ahern, M.J. Kreslake, M.J. Phalen, What is eHealth (6): perspectives on the evolution of ehealth research, J. Med. Internet Res. 8 (2006) e4.
[19] K.D. Ahern, Challenges and opportunities of ehealth research, Am. J. Prev. Med. 32 (2007) S75–S82.

[20] B.R. Martin, P. Nightingale, A. Yegros-Yegros, Science and technology studies: exploring the knowledge base, Res. Policy 41 (2012) 1182–1204.
[21] L. Leydesdorff, R.L. Goldstone, Interdisciplinarity at the journal and specialty level: the changing knowledge bases of the journal cognitive science, J. Assoc. Inf. Sci. Technol. 65 (2014) 164–177.
[22] N.R. Hardiker, M.J. Grant, Factors that influence public engagement with eHealth: a literature review, Int. J. Med. Inform. 80 (2011) 1–12.
[23] T. Heart, E. Kalderon, Older adults: are they ready to adopt health-related ICT? Int. J. Med. Inform. 82 (2013) e209–e231.
[24] T.L. Griffiths, M. Steyvers, Finding scientific topics, Proc. Natl. Acad. Sci. 101 (2004) 5228–5235.
[25] A. Ahmed, E.P. Xing, W.W. Cohen, R.F. Murphy, Structured correspondence topic models for mining captioned figures in biological literature, in: The 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Paris, France, 2009, pp. 39–47.
[26] D.M. Blei, A.Y. Ng, M.I. Jordan, Latent dirichlet allocation, J. Mach. Learn. Res. 3 (2003) 993–1022.
[27] B. Grun, K. Hornik, Topicmodels: an R package for fitting topic models, J. Stat. Softw. 40 (2011) 1–30.
[28] A. Mazloumian, D. Helbing, S. Lozano, R.P. Light, K. Börner, Global multi-level analysis of the 'Scientific Food Web', Sci. Rep. 3 (2013).
[29] R. Kraut, M. Patterson, V. Lundmark, S. Kiesler, T. Mukopadhyay, W. Scherlis, Internet paradox. A social technology that reduces social involvement and psychological well-being? Am. Psychol. 53 (1998) 1017–1031.
[30] D.F. Tate, R.R. Wing, R.A. Winett, Using Internet technology to deliver a behavioral weight loss program, JAMA 285 (2001) 1172–1177.
[31] L. Baker, T.H. Wagner, S. Singer, M. Bundorf, Use of the Internet and e-mail for health care information: results from a national survey, JAMA 289 (2003) 2400–2406.
[32] A. Bandura, Social Foundations of Thought and Action, Englewood Cliffs, NJ, 1986.
[33] J.T. Klein, W.H. Newell, Advancing interdisciplinary studies, in: J. Gaff, J. Ratcliff (Eds.), Handbook of the Undergraduate Curriculum: A Comprehensive Guide to Purposes, Structures, Practices, and Change, Jossey-Bass, San Francisco, CA, 1997, pp. 393–415.
[34] S.W. Aboelela, E. Larson, S. Bakken, O. Carrasquillo, A. Formicola, S.A. Glied, J. Haas, K.M. Gebbie, Defining interdisciplinary research: conclusions from a critical review of the literature, Health Serv. Res. 42 (2007) 329–346.
[35] K.M. Åkesson, B.-I. Saveman, G. Nilsson, Health care consumers' experiences of information communication technology—a summary of literature, Int. J. Med. Inform. 76 (2007) 633–645.
[36] D.F. Lobach, D.E. Detmer, Research challenges for electronic health records, Am. J. Prev. Med. 32 (2007) S104–S111.
[37] S. Pingree, R. Hawkins, T. Baker, L. DuBenske, L.J. Roberts, D.H. Gustafson, The value of theory for enhancing and understanding e-Health interventions, Am. J. Prev. Med. 38 (2010) 103–109.
[38] H. Oh, C. Rizo, M. Enkin, A. Jadad, What is ehealth (3): a systematic review of published definitions, J. Med. Internet Res. 7 (2005) e1.
[39] L.A. Lievrouw, The invisible college reconsidered: bibliometrics and the development of scientific communication theory, Commun. Res. 16 (1989) 615–628.