Review

Predictors of Health Information–Seeking Behavior: Systematic Literature Review and Network Analysis

Ardalan Mirzaei[1*], BPharm, MPhil, GCertEdStud (Higher Ed); Parisa Aslani[1*], BPharm (Hons), MSc, PhD, GCertEdStud (Higher Ed); Edward Joseph Luca[2], BA, MBA; Carl Richard Schneider[1*], BN, BPharm (Hons), PhD, PGCert (Higher Ed)

[1]The University of Sydney School of Pharmacy, Faculty of Medicine and Health, The University of Sydney, Australia

[2]The University of Sydney Library, The University of Sydney, Australia

*these authors contributed equally

Corresponding Author:
Ardalan Mirzaei, BPharm, MPhil, GCertEdStud (Higher Ed)
The University of Sydney School of Pharmacy
Faculty of Medicine and Health
Pharmacy and Bank Building A15
The University of Sydney, 2006
Australia
Phone: 61 286275225
Email: ardalan.mirzaei@sydney.edu.au

Abstract

Background: People engage in health information–seeking behavior to support health outcomes, and being able to predict such behavior can inform the development of interventions to guide effective health information seeking. Obtaining a comprehensive list of the predictors of health information–seeking behavior through a systematic search of the literature and exploring the interrelationship of these predictors are critical first steps in this process.

Objective: This study aims to identify significant predictors of health information–seeking behavior in the primary literature, develop a common taxonomy for these predictors, and identify the evolution of the concerned research field.

Methods: A systematic search of PsycINFO, Scopus, and PubMed was conducted for all years up to and including December 10, 2019. Quantitative studies identifying significant predictors of health information–seeking behavior were included. Information seeking was broadly defined and not restricted to any source of health information. Data extraction of significant predictors was performed by 2 authors, and network analysis was conducted to observe the relationships between predictors with time.

Results: A total of 9549 articles were retrieved, and after the screening, 344 studies were retained for analysis. A total of 1595 significant predictors were identified. These predictors were categorized into 67 predictor categories, with the most central predictors being age, education, gender, health condition, and financial income. With time, the interrelationship of predictors in the network became denser, with the growth of new predictor grouping reaching saturation (1 new predictor identified) in the past 7 years, despite increasing publication rates.

Conclusions: A common taxonomy was developed to classify 67 significant predictors of health information–seeking behavior. A time-aggregated network method was developed to track the evolution of the research field, showing the maturation of new predictor terms and an increase in primary studies reporting multiple significant predictors of health information–seeking behavior. The literature has evolved with a decreased characterization of novel predictors of health information–seeking behavior. In contrast, we identified a parallel increase in the complexity of predicting health information–seeking behavior, with an increase in the literature describing multiple significant predictors.

(J Med Internet Res 2021;23(7):e21680) doi: 10.2196/21680

KEYWORDS

information seeking; network analysis; health; review; temporal analysis; mobile phone

XSL•FO
RenderX

Introduction

Background

Health information seeking has been defined as "the ways in which individuals obtain information, including information about their health, health promotion activities, risks to one's health, and illness" [1]. A consumer's health information–seeking behavior has the potential to influence the process and outcomes related to coping or adjusting to an illness or condition [1].

The conceptualization of health information–seeking behavior has evolved since Lenz first defined it in 1984, which identified 2 dimensions of health information–seeking behavior: extent and method [2]. Lambert and Loiselle's comprehensive review [1] of the concept of health information–seeking behavior found definitions such as actions or behaviors used to obtain knowledge [3], clarify or confirm knowledge [4], satisfy a query [5], identify information sources [6,7], or demonstrate a coping strategy [8,9]. Other known models from the information science perspective include the Comprehensive Model of Information Seeking, which looks at information carrier characteristics, antecedents, and information-seeking actions [10] and the book by Case in 2002 about the research on information-seeking needs and behaviors [11]. A recent paper by Zimmerman and Shaw describes health information–seeking behavior as an *umbrella term* for many forms of information seeking, such as "direct seeking, information monitoring or browsing, and the passive receipt of information" [12]. Alternatively, people's passive receipt of health information has been defined separately as *information scanning* [13], where people may not be active in their search but are still receptive enough to receive information. Thus, although the concept of health information–seeking behavior has existed for more than 30 years, there remains a lack of consensus on its definition and model theories.

Despite these inconsistencies, previous models and theories commonly describe health information–seeking behavior as involving the action of seeking out information, regardless of where it comes from, how it is sought, or why it is sought [1,14]. Therefore, predictors of health information–seeking behavior can be described as the variables affecting the actions of seeking out information. These predictors can be contextual, such as the environment of an individual or their social networks, and personal, such as sociodemographic characteristics, health status, or internal beliefs. There may also be predictors of persistence in health information–seeking behavior, such as satisficing. Satisficing relates to decision making "through which an individual decides when an alternative approach or solution is sufficient to meet the individuals' desired goals rather than pursue the perfect approach" [15]. In the context of information seeking, it is choosing whether it is worth the cost or effort of continuing to search or whether already acquired information suffices.

Factors that influence why people engage in information seeking include the content of the information, which information sources or channels are frequently used, their credibility, and the barriers they may pose to seeking information [1,16]. A meta-analytic review by Chang and Huang quantified 7 predictors of health information–seeking behavior; however, not all of the predictors, such as behaviors (adherence) and beliefs, were included in their review [17]. Predicting peoples' behavior for health information seeking requires understanding the predictors and their significance and magnitude on information-seeking behavior.

We define *significant predictors* as those shown through empirical research to have a direct effect rather than an association (correlation) to health information seeking. An example of this direct effect is age. Nelissen et al [18] showed that an increase in age led to increased cancer information seeking. An example of an association is the relationship between patient-physician interaction and information-seeking behavior. In this case, it is unclear whether information seeking leads to better patient-physician interaction (an outcome of health information–seeking behavior) or whether better physician-patient interaction leads to increased health information–seeking behavior (a predictor of health information–seeking behavior). Although associations between predictors and health information–seeking behavior may have statistical significance in some empirical studies, knowing the direction of the effect from predictor to variable is more informative. Furthermore, although qualitative research can provide a foundation for identifying predictors of health information–seeking behavior, the ability to quantify the effect size allows for comparison of individual predictors' relative importance.

A comprehensive list of predictors of health information–seeking behavior provides researchers with a focus on identifying new significant predictors or examining the relationship and effect of new interventions. Predictors of health information–seeking behavior can support researchers, clinicians, private institutions, and public health initiatives in optimizing information-related interactions between themselves and consumers, leading to a more positive health care management experience [1,16].

Consistent classification of terms is the first step in formulating a comprehensive list of health information–seeking behavior predictors. For instance, there is concern that the terms *race* and *ethnicity* have been used synonymously in health research despite being separate constructs [19]. Thus, how predictors of health information–seeking behavior are defined in one study may not necessarily be consistent with another.

Anker et al [16] compiled a comprehensive list of predictors for health information–seeking behavior a decade ago through a systematic search of the literature. They extracted and reported on the methods and measures used in health information–seeking behavior research. However, there are 2 critical shortcomings of this review. First, Anker et al [20] restricted their definition of health information–seeking behavior to an active process, in accordance with Niederdeppe's definition of health information–seeking behavior. Consequently, predictors for the nonactive acquisition of health information were not identified [20]. Second, their search strategy was restricted to a single database (PsycINFO), with the justification that health information–seeking behavior is a social psychological construct instead of a medical construct. PsycINFO has focused subject

XSL•FO

RenderX

areas, and it is possible that other health information–seeking behavior researchers may have published in journals not indexed in PsycINFO.

Importantly, there is a need for an updated review to account for the evolution of information seeking as a result of the rapid emergence and dominance of mobile digital information technology. The use of the internet has been increasing in the past three decades [21]. Advances in technologies such as smartphones have led to increased availability and access to the internet. Since 2011, smartphone ownership by the American population has increased from 35% to 81% in 2019, with 96% of the population owning a cellular device [22]. Similarly, the use of smartphones has led to greater access to the internet, with the exclusive use of smartphones for internet access doubling from a reported 17.5% in 2013 to 37% of the American population in 2019 [21,23]. A smartphone user is estimated to spend a daily average of 2.6 hours on their device [24]. Although the internet has become a common source of health information [25], how the influence of the internet has modified predictors of health information–seeking behavior throughout time is yet to be well characterized. Previous studies have compared sources of information used by people as part of health information–seeking behavior; however, most studies have only compared the findings from the early 2000s with those from the early 2000s [26-28]. Such comparison studies report that internet use was not a predictor of information-seeking behavior, yet Huerta et al [26] reported an increase in internet use, especially in older age groups. In contrast, Li et al [28] performed a hierarchical regression analysis comparing 2002 and 2012 cohorts from the Pew Database to examine changes in health information–seeking behavior on the web. They identified internet access as a predictor for health information decline with time. The authors hypothesized that this could be partly because of the increase in misinformation and rise in smartphone use, resulting in increased accessibility of the internet as a source of health information in the United States. The extent to which these changes have affected predictors of health information–seeking behavior as a whole and across information sources and settings has yet to be reviewed.

Objectives

This review aims to identify predictors of health information–seeking behavior, as reported in the primary literature, and explore the relationships between predictors with time. The specific objectives are as follows:

- identify significant predictors of health information–seeking behavior in the primary literature;
- develop a common taxonomy for predictors of health information–seeking behavior;
- identify the evolution of the health information–seeking behavior research field using quantitative studies.

Methods

Selection Criteria

The following section outlines the inclusion and exclusion criteria for this review.

Types of Participants

The papers included defined participants as health consumers or caregivers. The intent of the search for information was important: a health consumer searches for information for their own self or treatment, as opposed to a health professional who may search for information to provide therapy. Caregivers were also included, as they were in a nontherapeutic relationship with the health consumer.

Papers in which the participants were health students (university or college) or simulation studies in which the participants sought information prospectively in hypothetical scenarios were excluded. Students studying a health-related degree were removed if in their health-related disciplines, they were searching for information for their future role as health professionals.

Types of Studies

Quantitative studies were eligible for inclusion in this study. Relevant study designs included experimental and epidemiological studies, including randomized controlled trials, nonrandomized controlled trials, quasi-experimental studies, before-and-after studies, prospective and retrospective cohort studies, case-control studies, and analytical cross-sectional studies.

An article was included if it reported the significance level of a predictor. That is, the study showed that a certain *predictor* was significant in causing *information seeking* rather than associations. Significance was determined by either *P* values <.05 or in the case of odds ratio if the CI did not cross 1. Articles were also included if further work demonstrated a causation or effect size, which was often shown through logistic or linear regression, confirmatory factor analysis, or structural equation modeling.

Studies were excluded if they did not have a quantitative focus. These included qualitative studies such as focus groups, semistructured interviews, mixed methods studies (with no quantitative component), and content analysis of websites or interviews.

Search Procedures

Search Limits

Papers published in English up to December 10, 2019, since database inception were considered for inclusion. No data range was applied. Participants' information seeking was not restricted to any source, and all sources (eg, web, health care practitioner) were included.

Databases

The databases searched were PsycINFO, Scopus, and PubMed. Scopus is considered the largest abstract and citation database of peer-reviewed literature and incorporates the results from Embase and MEDLINE [29].

Search Terms or Phrases

The keywords used were the following:

Health OR *Drug* OR *Medicine*; AND *Information Seeking* OR *Information Behavior* OR *Information Search* or *Satisficing* (Tables S1-S3 in Multimedia Appendix 1 for full syntax).

Screening

Two authors (AM and EJL) independently screened a random split of articles for inclusion by title and abstract using the selection criteria. Pilot tests were conducted to calibrate the screening process before the records were split.

Data Extraction

Each included paper was counted as a data source for the extraction; 2 authors (AM and EJL) independently extracted the following variables: year of publication, country of the study, participant recruitment, disease states, theories used, and significant predictors. Significant predictors were those variables for which direct effects were reported (not correlations) on health information–seeking behavior and provided significance with either *P* values <.05 or CIs (in the case of odds ratio) that did not cross 1. Significant relative predictors of health information–seeking behavior were also extracted from studies reporting comparative health information–seeking behavior between 2 or more groups. Any uncertainty associated with extraction was mediated by a third author (CRS), who performed an audit of approximately 10% (34/344) of the screened articles and extracted variables.

Analysis

Content Analysis

The first author (AM) analyzed the variables and categorized the significant predictors into emerging categories. Individual predictors identified in the individual papers were extracted, and an iterative process of clustering was undertaken by 2 authors (AM and CRS); the 2 authors reached a consensus for terms and categories. The categories consisted of identifying similarity between predictor terms: where predictors were the same, they were categorized together; where predictors were similar with a common definition, they were categorized together; and finally, where there was no common definition, but predictors were described similarly in text, they were categorized together [30].

Predictor Frequency

As part of the content analysis, a *word frequency* analysis was performed. In this case, the *words* chosen were identified-predictor terms. Examining the predictor term frequency assists in analyzing the strength and importance of a predictor with regard to other terms [31-34]. Each predictor extracted into a category was counted as 1 for that article. Multiple predictors, if categorized, were categorized as 1. For example, if the article reported "Age 25-30" and "Age 65-70" as significant factors, then they would be categorized as *age*; however, they would contribute only 1 to the *age* category for that article instead of 2. The total predictors were then reported, and the predictor frequency was used to develop the network structure for network analysis.

Network Analysis

Network analysis has been used in previous systematic reviews to identify relationships among authors of the included papers [35] or in a health context to compare drug treatments [36]. Traditionally, quantitative data from a systematic review are pooled via meta-analysis, which requires homogeneous data. Network analysis allows the examination of relationships among heterogeneous entities [37]. A network analysis was conducted to observe the relationships between the predictor terms. This method can help identify nodes (or predictor terms) connected to other nodes and show the relationships between terms in the literature [38,39].

The weights of the nodes were based on the frequency of the predictor, whereas their size was based on the number of articles mentioning the predictor term. An analysis of changes throughout time was undertaken to compare the networks of articles before and after 2008 (articles dated up to December 31, 2008)—while Pew Research started reporting smartphone ownership in 2011 [22], the iPhone was the beginning of a new phone era [40]. The year 2008 was chosen to distinguish between the availability of smartphones following the introduction of the iPhone in 2007, allowing for the uptake of the device to have begun [41]. Accordingly, in 2008, global mobile broadband subscriptions overtook fixed broadband subscriptions [42]. Time-based comparisons of temporal and atemporal network features were observed using time-varying networks. Such an approach has been used in ecology, transport, and social media [43-45].

The co-occurrence of individual predictors within an article was calculated based on the predictor frequency. Each individual predictor term was connected bidirectionally to another predictor from the same article. Each connection adds a weight of 1 to the edge. Edges were formed where a pair of predictor terms was mentioned together in an article. The visualization was created using R software (R Foundation for Statistical Computing), with the code available on GitHub [46]. The co-occurrence of predictors and visualization of the network was created using the igraph package, a software package used for network visualizations between different objects on a network map [47].

The number of nodes and edges, along with modularity, were captured to compare the different networks. Modularity measures the clustered communities of nodes, which is how the nodes cluster together, forming a community group of nodes distant to another community group of nodes. The full setup and parameters are available in GitHub [46]. See Multimedia Appendix 1 for further methods [38,39,48,49]

Results

Search Results

The literature search process is illustrated in Figure 1. From the 2 databases, a total of 9549 articles were retrieved, of which 2866 were duplicates. Following deduplication, title and abstract screening was performed, followed by full-text screening. A total of 344 papers were included in the final analysis. The results of the categorized predictors are reported in Table S4 in

Multimedia Appendix 1. The included articles contained papers published between 1993 and 2019, with a peak publication year in 2019 (Figure 2).

Most of the studies were conducted in the United States (n=202); 26 articles reported studies from China, 12 from Australia and South Korea, and 9 from the United Kingdom and Germany.

In 203 articles, participants were recruited specifically for the study. However, another source of participants was using existing databases of respondents such as the Health Information Trends Survey (HINTS; n=65 papers), The Pew Research Center's Internet & American Life Project (n=9), and the Pennsylvania Cancer Registry (n=8).

Participants were predominantly seeking information for chronic diseases, with cancer being the most studied condition (n=76).

Figure 1. Flowchart for article inclusion. Reasons for exclusion do not sum to the number of excluded articles because some articles overlapped in their reasons for exclusion.
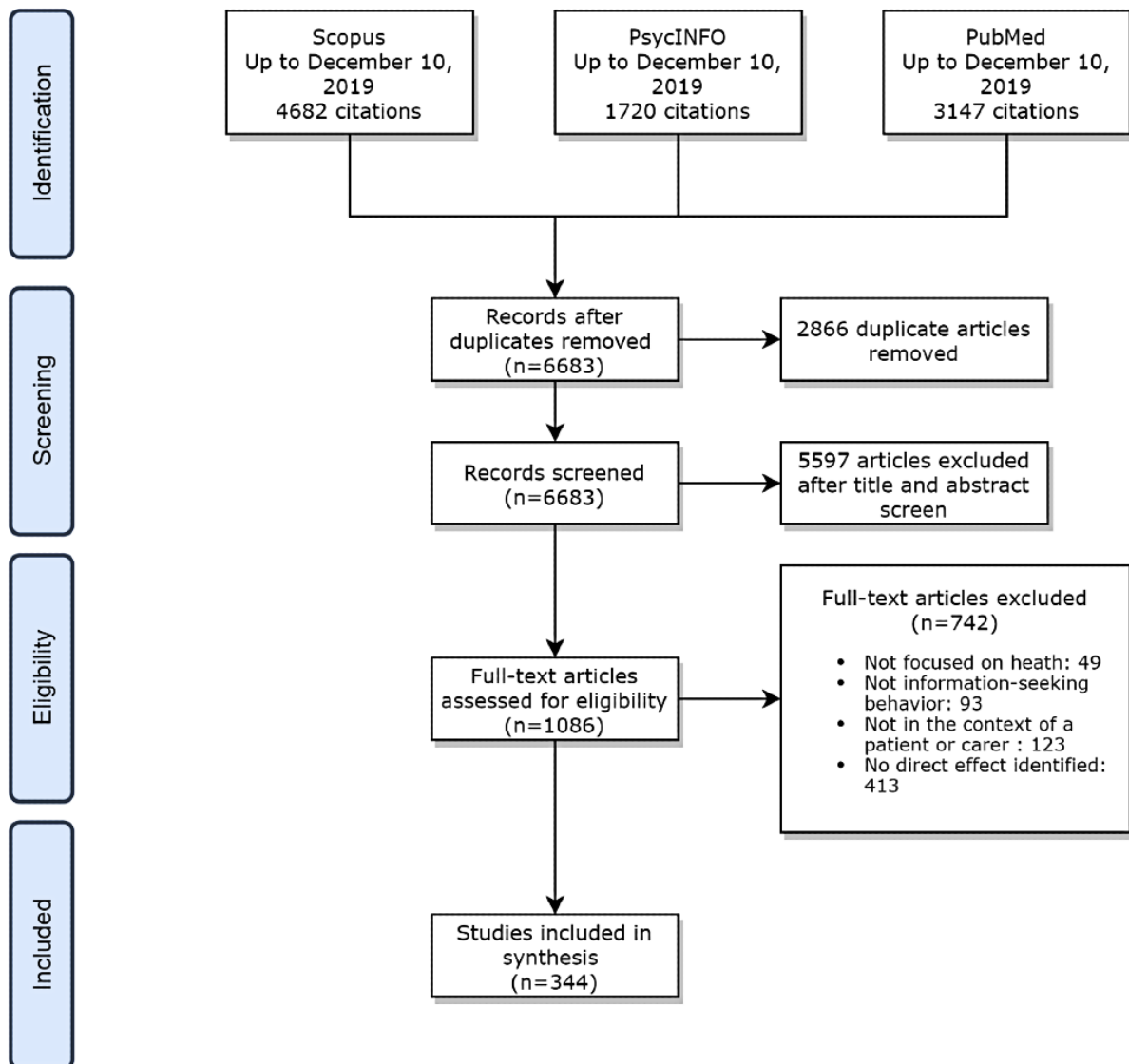
Figure 2. Total articles published each year and the number of articles that used a theoretically grounded approach.
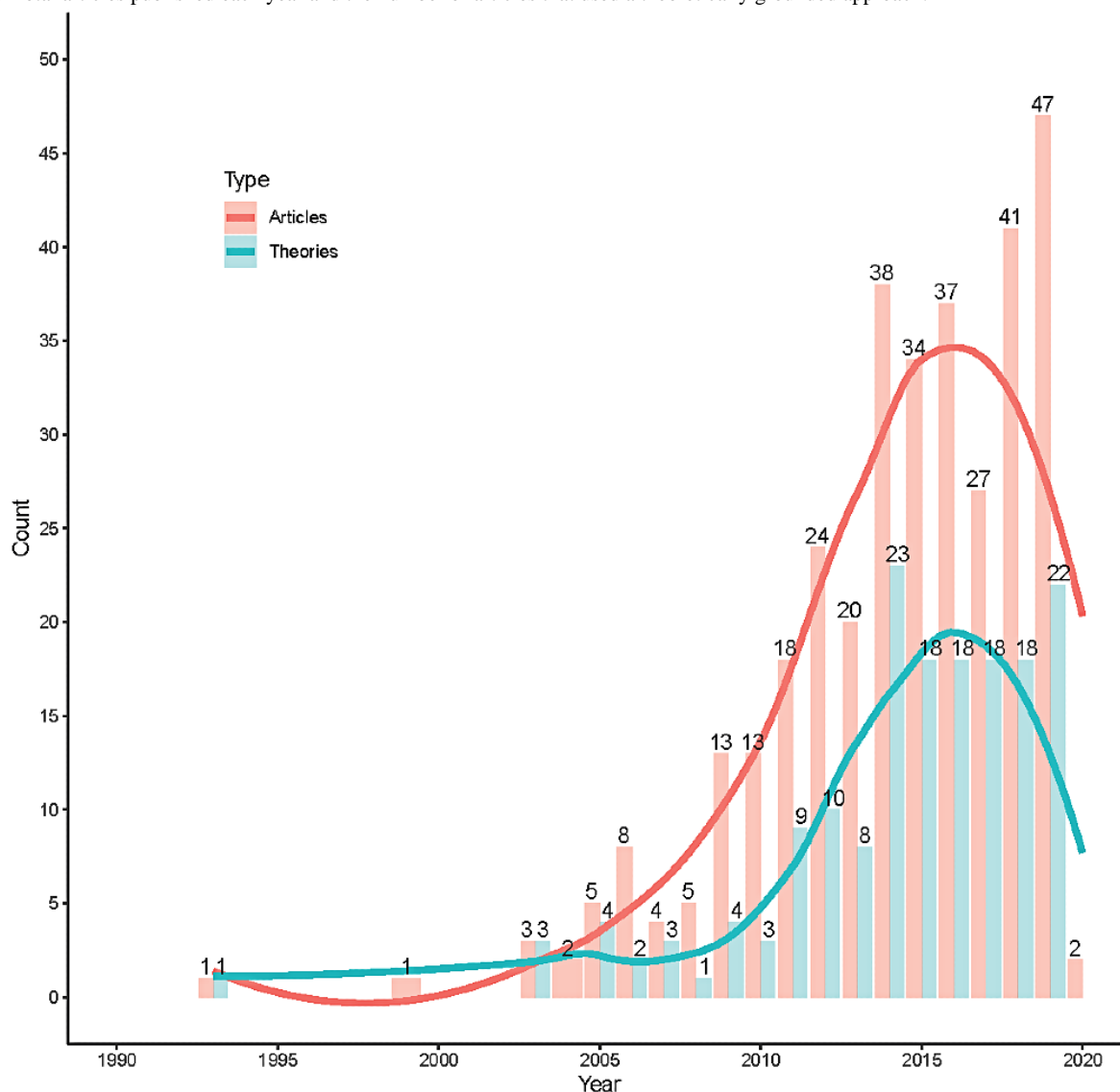
Content Analysis

Fewer than half of the papers (n=167) were underpinned by a theory or model. As shown in Figure 2, there was an increase in the number of papers with a theoretical underpinning until 2014, at which point a plateau developed from 2015 to 2018. In 2019, there was a second peak of articles with theoretical underpinnings. However, less than half of the papers published in 2019 were supported by a theory or model (Figure 2).

A total of 1595 nonunique, significant predictors were identified. Table S4 in Multimedia Appendix 1 lists the predictor categories. Predictors were classified into 68 categories (1 category was labeled *unclassifiable/other*; unclassifiable predictors were not carried forward for network analysis.). The categories were further grouped into sociodemographic, health, information source, information content, and affective predictor groups. Sociodemographic variables of education (n=160), age (n=156), and gender (n=120) were the most commonly reported significant predictor categories, followed by the health-related predictor categories of health condition (n=87). A noted increase in the number of predictors reported in the literature began in 2005 and peaked in 2019.

Network Analysis

The complete network with all terms and years combined resulted in 67 nodes (*other* node not included) and 4128 edge connections. The modularity of the groups revealed 3 clusters. The largest group of variables was predominately composed of psychosocial predictors (n=41). The second-largest group was sociodemographic predictors (n=22), followed by a third group that did not have any strong focus on any particular grouping (n=4). However, most sociodemographic variables (age, education, gender) had the greatest eigenvector centrality before other variables (health condition and financial income), thus appearing in the center of the network (Table S5 in Multimedia Appendix 1).

The network statistics reported in Table S5 of Multimedia Appendix 1 as well as in Figures 3 and 4 show a difference in structural characteristics before and after 2008. After 2008, only 15 new nodes were identified, with no new nodes identified after 2014 (Table S4 in Multimedia Appendix 1). There was a 7.8 times greater number of edges in the post-2008 network than in the pre-2008 network. The combination of an increased number of edges and a limited increase in nodes resulted in a

more connected network after 2008, with the average number of adjacent edges to each node (mean degree of the nodes) increased by 6.1 times compared with before 2008. Age, education, gender, and health condition were the nodes with the greatest degree of centrality, indicating the greatest influence on adjacent nodes (Figure 5). Modularity was greatest before 2008; it decreased to 2 in the post-2008 period, indicating tighter communities of nodes clustering together, with all years

compared being 3. This tighter clustering is because of the greater co-occurrence of the predictor terms being researched. That is, individual articles reported more significant predictors than previous articles. A sensitivity analysis was conducted to compare networks before and after 2014 (the most recent new node), which confirmed the dynamic shifts in network statistics after 2008 (Table S5 in Multimedia Appendix 1).

Figure 3. Fruchterman-Reingold layout algorithm of the network analysis comparing the pre- and including 2008 network models with color coding according to the modularity group membership of the complete model. The repository is available on GitHub [46].
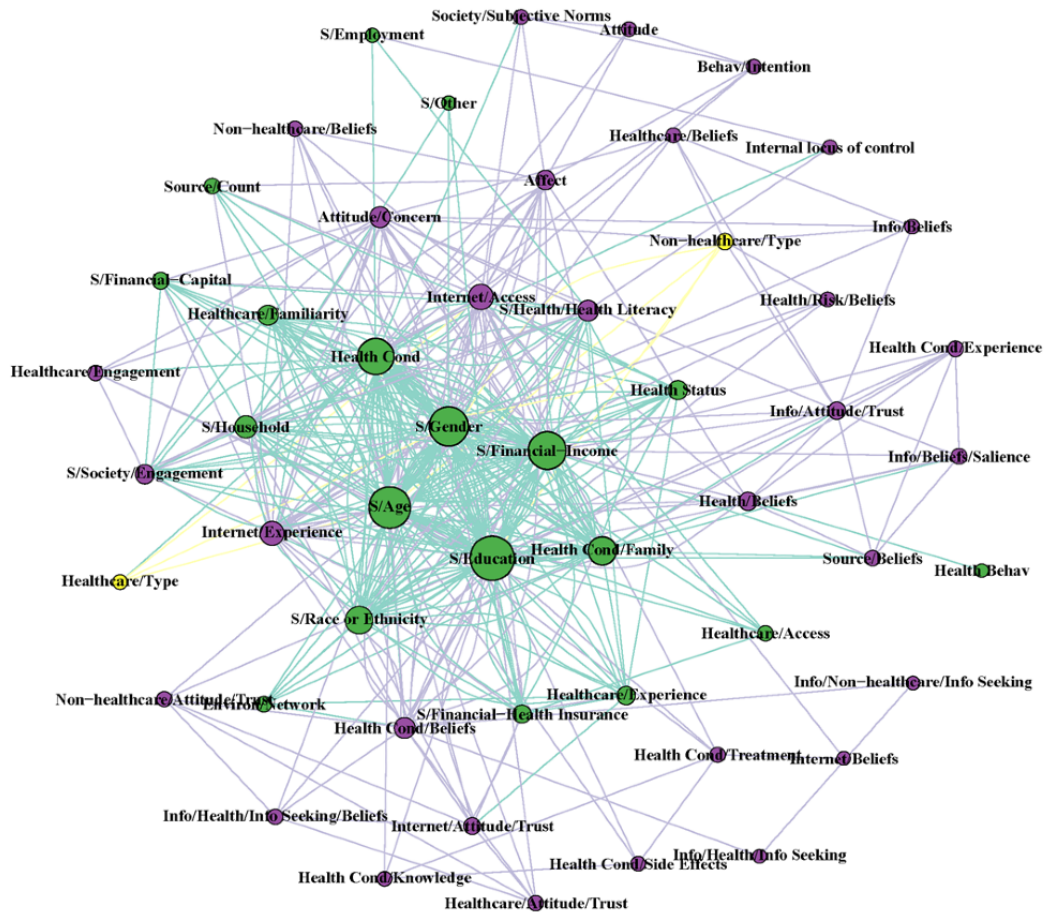
Figure 4. Fruchterman-Reingold layout algorithm of the network analysis comparing the post-2008 network model with color coding according to the modularity group membership of the complete model. The repository is available on GitHub [46].
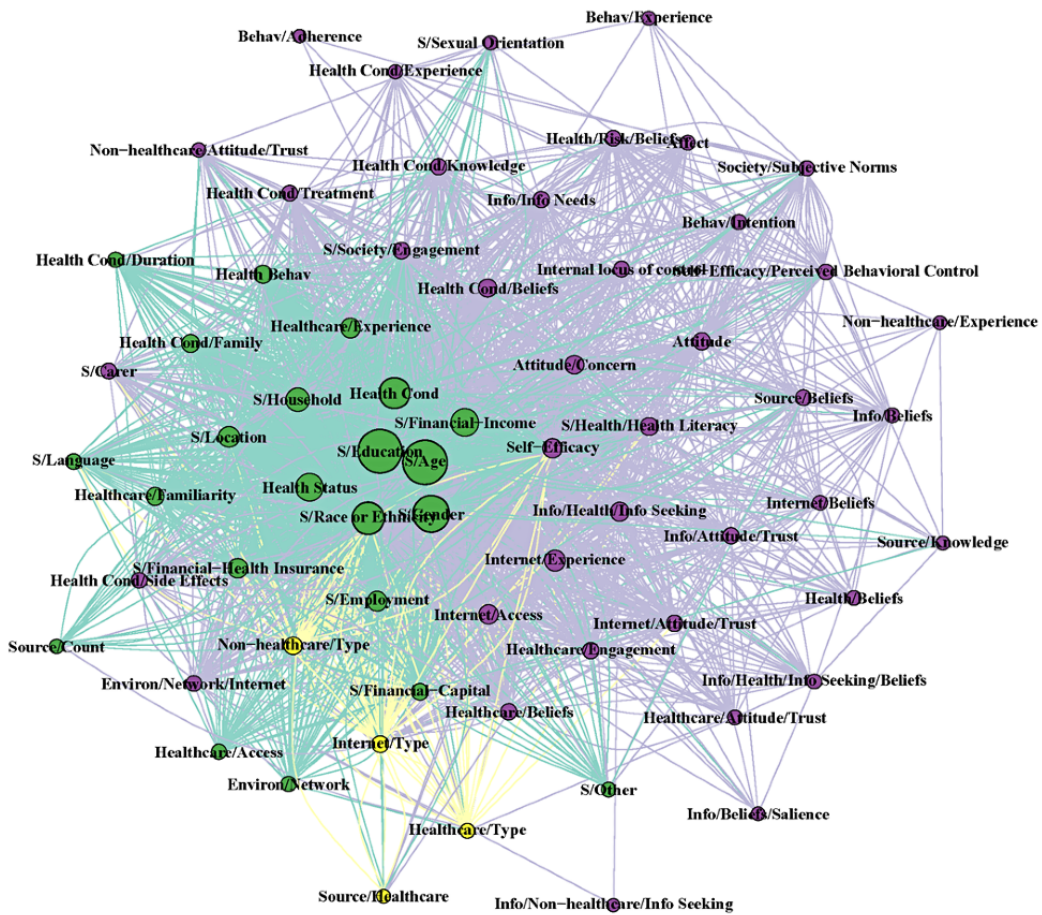
Figure 5. Fruchterman-Reingold layout algorithm of the network analysis comparing the complete model with all years until 2019 with color coding according to the modularity group membership of the complete model. The repository is available on GitHub [46].
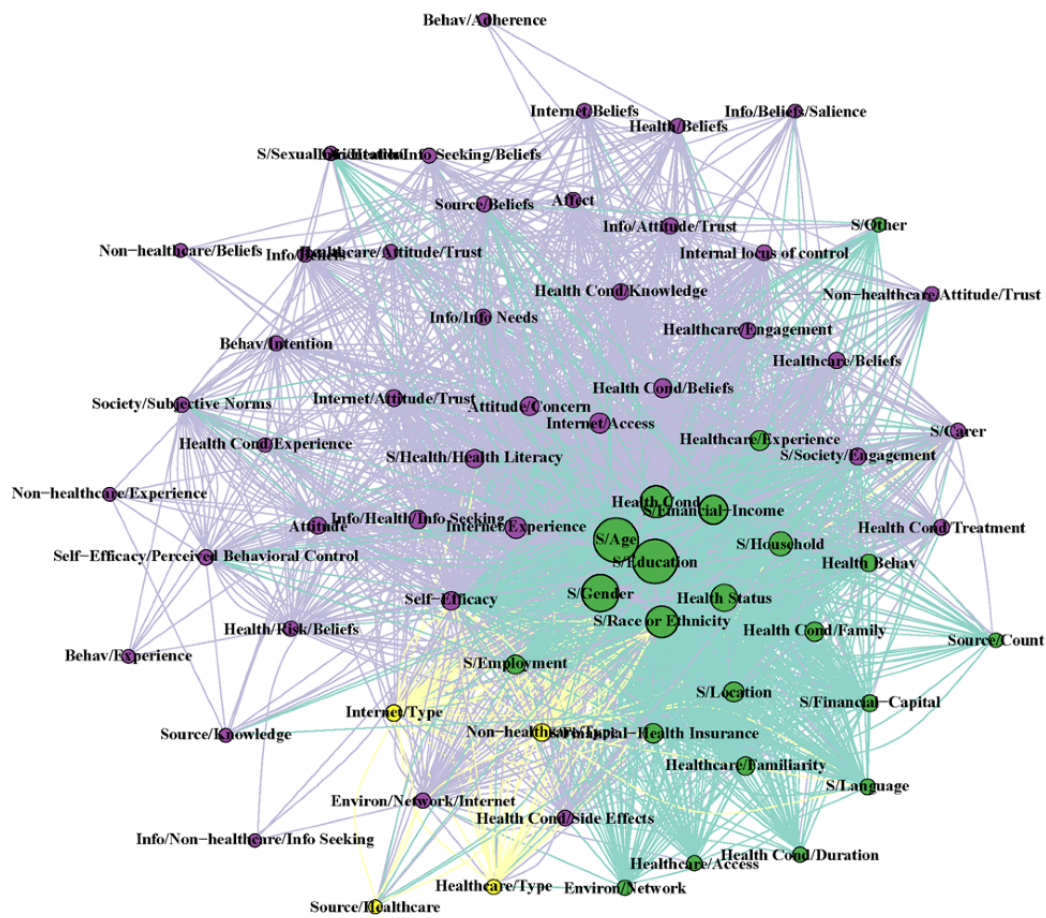
Discussion

Principal Findings

This paper reports a systematic literature search to identify and characterize predictors of health information–seeking behavior. The 344 included papers report 1595 significant predictor terms of health information–seeking behavior that can be classified into 68 categories. A comprehensive list of health information–seeking behavior predictor terms was developed. A novel temporal network analysis through the comparison of 2 sequential time-aggregated networks was conducted to characterize the relationships between health information–seeking behavior predictors and identify changes throughout time. This approach has never been previously used to characterize such relationships. Key findings were an increase in papers reporting on multiple significant predictors of health information–seeking behavior within a paper and a reduced rate in the identification of new predictors. The use of network analysis to map the relationships within a research field throughout time demonstrates the evolving nature of research and provides insight into how the understanding of predictors of health information–seeking behavior has developed. Van de Wijngaert et al [50] conducted a network approach to examine the current state in a research field pertaining to the adoption of eGovernment services. However, they used structural equation modeling and cross-sectional analysis, as opposed to our network comparison. The advantage of network comparison throughout time is that it allows for characterizing the evolution of research fields. Future work would be to explore the use of temporal dynamics of networks, ideally through the analysis of longitudinal data sets [43-45].

A recent meta-analysis conducted by Chang and Huang [17] on the antecedents predicting health information seeking aggregated the antecedents into 7 categories; through their review, they were able to quantify the effect sizes of the 7 categories. However, in their methodological design, not all the predictors of health information–seeking behavior were captured. Although a valuable review, we note the differences in design and the papers that could be retrieved. Since the initial review by Anker et al [16] in 2010, an increased number of papers exploring predictors of health information–seeking behavior has allowed for greater granularity in the identification of health information–seeking behavior predictors. Specifically, we have been able to develop a sociodemographic group of predictors from 5 predictors (age, gender, education, race, and health literacy) to 16 predictors (caregiver, employment, household, language, sexual orientation, finances, societal engagement, and location of residence). A potential benefit of such granularity is the improved targeting of interventions to optimize health information–seeking behavior.

Anker et al [16] reported that medication adherence is an outcome of engaging in health information seeking. However, according to our review, adherence was identified as a predictor.

This suggests that health information–seeking behavior is affected by a feedback loop, where outcomes from health information–seeking behavior can be a predictor for further health information–seeking behavior. This relationship should be examined in longitudinal studies. Few longitudinal studies have examined health information–seeking behavior, but our findings suggest that this could lead to unidentified health information–seeking behavior predictors. A longitudinal study described the reciprocal relationship between health anxiety and health information–seeking behavior on the web and how *cyberchondriac* people have health anxiety exacerbated [51]. Another study looked at clinician information engagement and information seeking [52], whereas others have shown that information needs and preferences change with time [53,54]. These initial findings demonstrate the utility of further longitudinal studies to measure additional predictors and outcomes of health information–seeking behavior.

The rate of article publication on health information–seeking behavior increased after 2005, with a doubling of articles published in the past 10 years compared with the prior 30 years. This finding mirrors the overall increase in the academic publishing rate [55]. Another explanation might be the establishment of data gathering institutions, such as the HINTS from the National Cancer Institute, which was established in 2003. Such cohort studies provide researchers with important opportunities to examine health information–seeking behaviors across large nationally representative sampling frames. HINTS is the most used data set across papers; therefore, it is the main source of identified health information–seeking behavior predictors. HINTS comprises 12 cross-sectional surveys that have been conducted in the past 15 years [56-65]. The data set has the advantage of being a representative sample frame of the United States.

The number of articles using theory to underpin their research has also increased with time. The use of theory has become a consistent theme in describing significant predictors. Interestingly, in the past 7 years, there has been a plateau in the frequency of publications reporting health information–seeking behavior predictors. A possible reason for this is maturation in the literature, with apparent saturation of identified predictors. Li et al [66] also identified an increase in publications until 2014. Our findings have extended this trend to demonstrate a plateau in publication rates since 2015.

Participants' interactions with social media, including social networking sites and health blogs, which were categorized as environment/network/internet, were identified as new predictors since 2008. Hamid et al [67] reviewed the role of social media in information-seeking behavior among international students, highlighting that specific information needs were satisfied by using social media. Although social media can be a medium for public health intervention [68], it can also present a challenge as a source of misinformation [69,70]. Competing misinformation has implications for providers of information using web media to target their audience. Providers or information creators could address the rise of misinformation by ensuring that the content delivered through social media is verified for quality and that continued monitoring is implemented.

The terminology used to describe predictors varied significantly between papers and, at times, lacked precision. Articles may have mentioned race as a predictor, but on closer inspection of the survey, items used for race, ethnicity, and culture overlap. A potential contributor to the lack of clarity in terminology is the low number of studies that used a theoretically grounded approach. Ambiguous terminology poses a challenge when comparing findings between papers on health information–seeking behavior. In response to this issue, this review developed a common classification structure for predictor terms. This structure has the potential to be developed into a future consensus taxonomy for predictors of health information–seeking behavior using domain ontologies [71].

Conducting a network analysis for predictors of health information–seeking behavior is a novel approach for analyzing the health information–seeking behavior research field. The overall network analysis shows the interrelationship of the predictor variables; however, the interaction between these variables in predicting health information–seeking behavior is still unclear. A concern is the issue of terms being correlated with each other, resulting in collinearity. The collinearity of predictor terms may affect how an individual predictor term affects the health information–seeking behavior. The temporal network analysis finding of a 6.1 greater mean degree demonstrates an increase in publications reporting multiple significant predictors. Increased reporting was also supported by an increase in co-reporting, represented by reducing network modularity with time. The high centrality of age, education, gender, and health condition indicates that these were the most commonly reported predictors when multiple predictors were reported in the included studies. Such studies have a greater ability to identify the predictors of collinearity. The network analysis approach allowed us to examine how our understanding of predictors of health information–seeking behavior and their interrelationships changed with time. Such changes have occurred in the presence of a shift toward mobile technology becoming commonplace.

Strengths and Limitations

This review has several strengths. Multiple authors followed a rigorous methodology to extract data and reach an agreement on definitions. However, because of the sheer number of articles returned in the initial database searches, there is a potential risk that articles meeting the inclusion criteria could have been potentially omitted. Nevertheless, the likelihood of omitted articles affecting the findings of this review was low because of the number of included articles. The developed taxonomy of predictors was directly informed by the included articles via a theoretical agnostic approach and consensus between 2 authors (AM and CRS). The content validity of the developed taxonomy would benefit from validation via conceptual synthesis and a consensus approach from experts in the field of health information–seeking behavior. The reliability of data extraction could be considered a limitation of our review, which we tried to mitigate through a 10% audit by a third author (CRS). Using a quantitative measure of intercoder reliability would increase confidence in reliability.

Limitations to the search strategy are, first, the inclusion of only articles published in English. This is a potential issue in this field, as geographic and cultural differences have been identified. The United States is the most represented country. However, a bibliometric analysis of the internet health information–seeking behavior literature has been previously performed by Li et al [66]. The authors similarly identified a majority of articles from the United States. A skew toward a single country may introduce geographic bias in the literature and subsequent identification of significant health information–seeking behavior predictors. There is evidence that context can directly influence individuals' health information–seeking behavior, such as being in a low-resource setting [72]. Therefore, it is important to be mindful of the number of studies from high-resource settings when considering the implications of our findings in low-resource settings. For instance, the presence of the predictor variable *health care source accessibility* may be more pertinent for countries without universal health care coverage, such as the United States, where access to physicians is variable [73,74] (Multimedia Appendix 1). A second limitation is the restriction of the definition of health information–seeking behavior as an active behavior. This limited the ability of the review findings to represent the predictors of passive health information–seeking behavior or *scanning*. Third, the review findings do not represent predictors of health information–seeking behavior for university students because of the exclusion of this subpopulation.

Finally, systematic reviews typically include an assessment of the risk of bias. The heterogeneity of the studies and the observational nature of most study designs meant that an assessment of bias was not appropriate. This led to this systematic review not adhering to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-analyses) guidelines and protocols in their entirety. However, the systematic approach adds to the strengths of this study.

Conclusions

A systematic literature search identified 344 papers reporting the predictors of health information–seeking behavior. A common taxonomy was developed to classify the predictors of health information–seeking behavior into 67 categories. Only 24% (16/67) of the predictor groupings have emerged since the invention of smartphones. Novel network analysis identified that the growth of new predictor groupings had approached saturation with only a single new predictor identified in the past 7 years, despite increasing publication rates. Publication network analysis is a promising methodology for measuring trends across scientific fields.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search terms, list of predictors and their definitions, network statistics, and network analysis methods.
[DOCX File , 31 KB-Multimedia Appendix 1]

References

1. Lambert SD, Loiselle CG. Health information seeking behavior. Qual Health Res 2007 Oct 01;17(8):1006-1019. [doi: 10.1177/1049732307305199] [Medline: 17928475]
2. Lenz ER. Information seeking: a component of client decisions and health behavior. ANS Adv Nurs Sci 1984 Apr;6(3):59-72. [doi: 10.1097/00012272-198404000-00010] [Medline: 6426379]
3. Barsevick AM, Johnson JE. Preference for information and involvement, information seeking and emotional responses of women undergoing colposcopy. Res Nurs Health 1990 Feb;13(1):1-7. [doi: 10.1002/nur.4770130103] [Medline: 2305105]
4. Conley VM. Beyond knowledge deficit to a proposal for information-seeking behaviors. Nurs Diagn 1998 Apr;9(4):129-135. [doi: 10.1111/j.1744-618x.1998.tb00182.x] [Medline: 9987316]
5. Baker LM, Connor JJ. Physician-patient communication from the perspective of library and information science. Bull Med Libr Assoc 1994 Jan;82(1):36-42 [FREE Full text] [Medline: 8136759]
6. Johnson J. On contexts of information seeking. Inf Process Manag 2003 Sep;39(5):735-760. [doi: 10.1016/s0306-4573(02)00030-4]
7. Czaja R, Manfredi C, Price J. The determinants and consequences of information seeking among cancer patients. J Health Commun 2003 Nov;8(6):529-562. [doi: 10.1080/716100418] [Medline: 14690888]
8. van der Molen B. Relating information needs to the cancer experience: 1. Information as a key coping strategy. Eur J Cancer Care (Engl) 1999 Dec;8(4):238-244. [doi: 10.1046/j.1365-2354.1999.00176.x] [Medline: 10889622]
9. Rees CE, Bath PA. The psychometric properties of the Miller Behavioural Style Scale with adult daughters of women with early breast cancer: a literature review and empirical study. J Adv Nurs 2000 Aug;32(2):366-374. [doi: 10.1046/j.1365-2648.2000.01485.x] [Medline: 10964184]
10. Johnson JD, Donohue WA, Atkin CK, Johnson S. A comprehensive model of information seeking. Sci Commun 2016 Aug 18;16(3):274-303. [doi: 10.1177/1075547095016003003]
11. Case DO, Given LM. Looking for information: a survey of research on information seeking, needs, and behavior (4th edition)). Bingley, UK: Emerald Group Publishing; 2016.

12. Zimmerman MS, Shaw G. Health information seeking behaviour: a concept analysis. Health Info Libr J 2020 Sep 12;37(3):173-191. [doi: 10.1111/hir.12287] [Medline: 32052549]

13. Kelly S, Eldredge SA, Dalton ED, Miller LE. Health-information behavior: an initial validity portfolio for active and passive measures. Commun Res Rep 2014 Apr 25;31(2):171-182. [doi: 10.1080/08824096.2014.907145]

14. Zare-Farashbandi F, Lalazaryan A. A review of models and theories of health information seeking behavior. Int J Health Syst Disaster Manage 2014;2(4):193. [doi: 10.4103/2347-9019.144371]

15. Simon HA. Designing organizations for an information-rich world. International Library of Critical Writings in Economics. 1969. URL: https://zeus.zeit.de/2007/39/simon.pdf [accessed 2021-06-14]

16. Anker AE, Reinhart AM, Feeley TH. Health information seeking: a review of measures and methods. Patient Educ Couns 2011 Mar;82(3):346-354. [doi: 10.1016/j.pec.2010.12.008] [Medline: 21239134]

17. Chang C, Huang M. Antecedents predicting health information seeking: a systematic review and meta-analysis. Int J Inf Manag 2020 Oct;54:102115. [doi: 10.1016/j.ijinfomgt.2020.102115]

18. Nelissen S, Beullens K, Lemal M, Van den Bulck J. Fear of cancer is associated with cancer information seeking, scanning and avoiding: a cross-sectional study among cancer diagnosed and non-diagnosed individuals. Health Info Libr J 2015 Jun 23;32(2):107-119 [FREE Full text] [doi: 10.1111/hir.12100] [Medline: 25809822]

19. Sheldon T, Parker H. Race and ethnicity in health research. J Public Health (Oxf) 1992;14(2):104-110. [doi: 10.1093/oxfordjournals.pubmed.a042706]

20. Niederdeppe J, Hornik RC, Kelly BJ, Frosch DL, Romantan A, Stevens RS, et al. Examining the dimensions of cancer-related information seeking and scanning behavior. Health Commun 2007 Aug 08;22(2):153-167. [doi: 10.1080/10410230701454189] [Medline: 17668995]

21. Demographics of internet and home broadband usage in the United States. Pew Research Center. URL: https://www.pewresearch.org/internet/fact-sheet/internet-broadband/ [accessed 2021-05-28]

22. Demographics of mobile device ownership and adoption in the United States. Pew Research Center. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2021-05-28]

23. Anderson M. Mobile technology and home broadband 2019. Pew Research Center. 2019. URL: https://www.benton.org/headlines/mobile-technology-and-home-broadband-2019 [accessed 2021-05-28]

24. Montag C, Błaszkiewicz K, Sariyska R, Lachmann B, Andone I, Trendafilov B, et al. Smartphone usage in the 21st century: who is active on WhatsApp? BMC Res Notes 2015 Aug 04;8(1):331 [FREE Full text] [doi: 10.1186/s13104-015-1280-z] [Medline: 26238512]

25. Percheski C, Hargittai E. Health information-seeking in the digital age. J Am Coll Health 2011 Apr 13;59(5):379-386. [doi: 10.1080/07448481.2010.513406] [Medline: 21500056]

26. Huerta TR, Walker DM, Johnson T, Ford EW. A time series analysis of cancer-related information seeking: hints from the Health Information National Trends Survey (HINTS) 2003-2014. J Health Commun 2016 Sep;21(9):1031-1038. [doi: 10.1080/10810730.2016.1204381] [Medline: 27565190]

27. Rutten LJ, Agunwamba AA, Wilson P, Chawla N, Vieux S, Blanch-Hartigan D, et al. Cancer-related information seeking among cancer survivors: trends over a decade (2003-2013). J Cancer Educ 2016 Jun 26;31(2):348-357. [doi: 10.1007/s13187-015-0802-7] [Medline: 25712202]

28. Li J, Theng Y, Foo S. Predictors of online health information seeking behavior: changes between 2002 and 2012. Health Inform J 2016 Dec 26;22(4):804-814 [FREE Full text] [doi: 10.1177/1460458215595851] [Medline: 26261218]

29. Falagas ME, Pitsouni EI, Malietzis GA, Pappas G. Comparison of PubMed, Scopus, Web of Science, and Google Scholar: strengths and weaknesses. FASEB J 2007 Sep 20;22(2):338-342. [doi: 10.1096/fj.07-9492lsf]

30. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov 01;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]

31. Du H, Li N, Brown MA, Peng Y, Shuai Y. A bibliographic analysis of recent solar energy literatures: the expansion and evolution of a research field. Renew Energy 2014 Jun;66:696-706. [doi: 10.1016/j.renene.2014.01.018]

32. Li J, Zhang Y, Wang X, Ho YS. Bibliometric analysis of atmospheric simulation trends in meteorology and atmospheric science journals. Croatica Chemica Acta 2009;82(3):695-705 [FREE Full text]

33. Mao N, Wang M, Ho Y. A bibliometric study of the trend in articles related to risk assessment published in science citation index. Hum Ecol Risk Assess 2010 Aug 23;16(4):801-824. [doi: 10.1080/10807039.2010.501248]

34. Wang L, Zhao L, Mao G, Zuo J, Du H. Way to accomplish low carbon development transformation: a bibliometric analysis during 1995–2014. Renew Sustain Energy Rev 2017 Feb;68:57-69. [doi: 10.1016/j.rser.2016.08.021]

35. Leppink J, Pérez-Fuster P. Social networks as an approach to systematic review. Health Prof Edu 2019 Sep;5(3):218-224 [FREE Full text] [doi: 10.1016/j.hpe.2018.09.002]

36. Caldwell DM. An overview of conducting systematic reviews with network meta-analysis. Syst Rev 2014 Sep 29;3(1):109 [FREE Full text] [doi: 10.1186/2046-4053-3-109] [Medline: 25267336]

37. Marsden PV. Measures of network centrality. In: Wright JD, editor. International Encyclopedia of the Social & Behavioral Sciences (Second Edition). Amsterdam: Elsevier; 2015:1-23185.

38. Cherven K. Mastering Gephi Network Visualization. Birmingham, United Kingdom: Packt Publishing; 2015:1-378.

39. Khokhar D. Gephi Cookbook. Birmingham, United Kingdom: Packt Publishing Limited; 2015.

40. Sarwar M, Soomro T. Impact of smartphone's on society. Eur J Sci Res 2013;98(2):216-226 [FREE Full text]

41. Islam N, Want R. Smartphones: past, present, and future. IEEE Pervasive Comput 2014 Oct;13(4):89-92. [doi: 10.1109/mprv.2014.74]

42. The World in 2009: ICT facts and figures. ITU-D. 2009. URL: https://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2009.pdf [accessed 2021-05-28]

43. Blonder B, Wey TW, Dornhaus A, James R, Sih A. Temporal dynamics and network analysis. Methods Ecol Evol 2012 Aug 01;3(6):958-972. [doi: 10.1111/j.2041-210x.2012.00236.x]

44. Santoro N, Quattrociocchi W, Flocchini P, Casteigts A, Amblard F. Time-varying graphs and social network analysis: temporal indicators and metrics. arXiv : Computer Science - Social and Information Networks. 2011. URL: https://arxiv.org/abs/1102.0629 [accessed 2021-05-28]

45. Tang J, Musolesi M, Mascolo C, Latora V. Temporal distance metrics for social network analysis. In: Proceedings of the 2nd ACM Workshop on Online Social Networks. 2009 Presented at: SIGCOMM '09: ACM SIGCOMM 2009 Conference; August, 2009; Barcelona Spain p. 31-36. [doi: 10.1145/1592665.1592674]

46. Mirzaei A. Algorithm and code used for calculating and displaying a network analysis for the predictors of health information seeking behaviour. Github: HISB Predictors Network Analysis. 2020. URL: https://github.com/ardimirzaei/HISB_Predictors_Network_Analysis [accessed 2021-05-28]

47. Csardi G, Nepusz T. The igraph software package for complex network research. InterJ Compl Sys 2006;1695(5):1-9 [FREE Full text]

48. Fruchterman TM, Reingold EM. Graph drawing by force-directed placement. Softw Pract Exp 1991 Nov;21(11) [FREE Full text]

49. Mullarkey MC, Marchetti I, Beevers CG. Using network analysis to identify central symptoms of adolescent depression. J Clin Child Adolesc Psychol 2019;48(4):656-668 [FREE Full text] [doi: 10.1080/15374416.2018.1437735] [Medline: 29533089]

50. van de Wijngaert L, Bouwman H, Contractor N. A network approach toward literature review. Qual Quant 2012 Nov 10;48(2):623-643. [doi: 10.1007/s11135-012-9791-3]

51. Poel FT, Baumgartner SE, Hartmann T, Tanis M. The curious case of cyberchondria: a longitudinal study on the reciprocal relationship between health anxiety and online health information seeking. J Anxiety Disord 2016 Oct;43:32-40. [doi: 10.1016/j.janxdis.2016.07.009] [Medline: 27497667]

52. Moldovan-Johnson M, Tan AS, Hornik RC. Navigating the cancer information environment: the reciprocal relationship between patient-clinician information engagement and information seeking from nonmedical sources. Health Commun 2014 Dec 20;29(10):974-983 [FREE Full text] [doi: 10.1080/10410236.2013.822770] [Medline: 24359259]

53. Harrison DE, Galloway S, Graydon JE, Palmer-Wickham S, Rich-van der Bij L. Information needs and preference for information of women with breast cancer over a first course of radiation therapy. Patient Educ Couns 1999 Nov;38(3):217-225. [doi: 10.1016/s0738-3991(99)00009-9]

54. Matsuyama RK, Kuhn LA, Molisani A, Wilson-Genderson MC. Cancer patients' information needs the first nine months after diagnosis. Patient Educ Couns 2013 Jan;90(1):96-102. [doi: 10.1016/j.pec.2012.09.009] [Medline: 23058682]

55. Bjork B, Roos A, Lauri M. Scientific journal publishing: yearly volume and open access availability. Information Research. 2009. URL: http://www.informationr.net/ir/14-1/paper391.html [accessed 2021-05-28]

56. National Cancer Institute. Health Information National Trends Survey. NCI Thesaurus, Code C20093 2003:A [FREE Full text] [doi: 10.32388/6P9GLV]

57. Westat. Health Information National Trends Survey 2005 (HINTS 2005) - final report. National Cancer Institute, Bethesda, MD. 2005. URL: https://hints.cancer.gov/docs/HINTS_2005_Final_Report.pdf [accessed 2021-05-28]

58. Westat. Health Information National Trends Survey 4 (HINTS 4) Cycle 1 Methodology Report. National Cancer Institute, Bethesda, MD. 2012. URL: https://hints.cancer.gov/docs/HINTS4_Cycle1_Methods_Report_revised_Jun2012.pdf [accessed 2021-05-28]

59. Westat. Health Information National Trends Survey 4 (HINTS 4) Cycle 2 Methodology Report. National Cancer Institute, Bethesda, MD. 2013. URL: https://hints.cancer.gov/docs/HINTS_4_Cycle2_Methods_Report.pdf [accessed 2021-05-28]

60. Westat. Health Information National Trends Survey 4 (HINTS 4) Cycle 3 Methodology Report. National Cancer Institute, Bethesda, MD. 2014. URL: https://hints.cancer.gov/docs/methodologyreports/HINTS_4_Cycle_3_Methodology_Report.pdf [accessed 2021-05-28]

61. Westat. Health Information National Trends Survey 4 (HINTS 4) HINTS-FDA Methodology Report. National Cancer Institute Bethesda, MD. 2015. URL: https://hints.cancer.gov/docs/Instruments/HINTS-FDA_Methodology_Report.pdf [accessed 2021-05-28]

62. Westat. Health Information National Trends Survey 4 (HINTS 4) HINTS-FDA2 Methodology Report. National Cancer Institute Bethesda, MD. 2017. URL: https://hints.cancer.gov/docs/methodologyreports/HINTS_FDA_Cycle2_Methodology_Rpt.pdf [accessed 2021-05-28]

63. Westat. Health Information National Trends Survey 5 (HINTS 5) Cycle 1 Methodology Report. National Cancer Institute Bethesda, MD. 2017. URL: https://hints.cancer.gov/docs/methodologyreports/HINTS5_Cycle_1_Methodology_Rpt.pdf [accessed 2021-05-28]

XSL•FO

RenderX

64.   Westat. Health Information National Trends Survey 5 (HINTS 5) Cycle 2 Methodology Report. National Cancer Institute Bethesda, MD. 2018. URL: https://hints.cancer.gov/docs/methodologyreports/HINTS5_Cycle_2_Methodology_Report.pdf [accessed 2021-05-28]

65.   Westat. Health Information National Trends Survey 4 (HINTS 4) Cycle 4 Methodology Report. National Cancer Institute Bethesda, MD. 2015. URL: https://hints.cancer.gov/docs/Instruments/HINTS_4_Cycle_4_Methodology_Report_Final.pdf [accessed 2021-05-28]

66.   Li F, Li M, Guan P, Ma S, Cui L. Mapping publication trends and identifying hot spots of research on internet health information seeking behavior: a quantitative and co-word biclustering analysis. J Med Internet Res 2015 Mar 25;17(3):e81 [FREE Full text] [doi: 10.2196/jmir.3326] [Medline: 25830358]

67.   Hamid S, Bukhari S, Ravana SD, Norman AA, Ijab MT. Role of social media in information-seeking behaviour of international students. Aslib J Inf Manag 2016 Sep 19;68(5):643-666. [doi: 10.1108/ajim-03-2016-0031]

68.   Dunn AG, Mandl KD, Coiera E. Social media interventions for precision public health: promises and risks. NPJ Digit Med 2018 Sep 19;1(1):1-47 [FREE Full text] [doi: 10.1038/s41746-018-0054-0] [Medline: 30854472]

69.   Rolls K, Massey D. Social media is a source of health-related misinformation. Evid Based Nurs 2021 Apr 11;24(2):46. [doi: 10.1136/ebnurs-2019-103222] [Medline: 32046968]

70.   Wang Y, McKee M, Torbica A, Stuckler D. Systematic literature review on the spread of health-related misinformation on social media. Soc Sci Med 2019 Nov;240:112552 [FREE Full text] [doi: 10.1016/j.socscimed.2019.112552] [Medline: 31561111]

71.   McDaniel M, Storey VC. Evaluating domain ontologies. ACM Comput Surv 2019 Sep 18;52(4):1-44. [doi: 10.1145/3329124]

72.   Dutta R. Information needs and information-seeking behavior in developing countries: a review of the research. Int Inf Library Rev 2013 Dec 02;41(1):44-51. [doi: 10.1080/10572317.2009.10762796]

73.   Naylor KB, Tootoo J, Yakusheva O, Shipman SA, Bynum JP, Davis MA. Geographic variation in spatial accessibility of U.S. healthcare providers. PLoS One 2019 Apr 9;14(4):e0215016 [FREE Full text] [doi: 10.1371/journal.pone.0215016] [Medline: 30964933]

74.   Aldridge ML, Daniels JL, Jukic AM. Mammograms and healthcare access among US Hispanic and non-Hispanic women 40 years and older. Fam Community Health 2006;29(2):80-88. [doi: 10.1097/00003727-200604000-00004] [Medline: 16552286]

Abbreviations

HINTS: Health Information Trends Survey
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-analyses

---