

Concept of digital health literacy revisited: Using text network and topic model analysis

Jiyoung Park^{1,2}, Seohyun Won¹, Mingee Choi^{3,4} , Chul Hee Kang⁵
and Han Shi Jocelyn Chew²

Abstract

Background: The rapid integration of ICT into healthcare has elevated the critical role of digital health literacy (DHL). However, the conceptual relationship between DHL and electronic health (eHealth), along with the impact of transformative events such as the Fourth Industrial Revolution and the COVID-19 pandemic, remains inadequately investigated.

Objective: This study seeks to analyze research trends in DHL across four distinct historical periods to uncover key themes and their temporal evolution.

Methods: A comparative analysis of 2645 abstracts from Scopus publications (1977–2022) was conducted, segmented into four periods: (I) emerging era (1977–2006), (II) establishment era (2007–2016), (III) diffusion era (2017–2020), and (IV) post-pandemic era (2021–present). Text network analysis identified core keywords, and Latent Dirichlet Allocation (LDA) extracted dominant topics and their temporal evolution.

Results: Since 2006, DHL research has exhibited consistent growth, underpinned by transformative advancements during the Fourth Industrial Revolution and further amplified by a significant surge in scholarly engagement in the post-pandemic era. Importantly, during the diffusion era (Period III), a divergence in the trajectories of “digital health literacy” and “electronic health literacy” emerged.

Text network analysis revealed a progression toward greater uniformity in node sizes over time, coupled with an increase in the complexity and intricacy of connections between nodes. These findings indicate a growing diversity and nuanced understanding of concepts associated with DHL. Moreover, research in the post-pandemic era (Period IV) emphasized the critical role of DHL in addressing health disparities and advancing equitable access to healthcare.

Conclusion: The study reveals the dynamic progression of DHL research, catalyzed by technological advancements and global health crises. Strengthening DHL, particularly among vulnerable populations, is crucial for mitigating health disparities in a rapidly digitalizing world. Future research should prioritize the development of targeted interventions and examine DHL’s impact across diverse sociocultural contexts.

Keywords

COVID-19, digital health literacy, electronic health literacy, fourth industrial revolution, text network analysis

Received: 6 October 2024; accepted: 24 March 2025

¹College of Nursing, Institute for Health Science Research, Inje University, Busan, Republic of Korea

²Alice Centre for Nursing Studies, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore

³Department of Biostatistics, Korea University College of Medicine, Seoul, Republic of Korea

⁴Department of Preventive Medicine, Yonsei University College of Medicine, Seoul, Republic of Korea

⁵Graduate School of Social Welfare, Yonsei University, Seoul, South Korea

Corresponding author:

Mingee Choi, Department of Biostatistics, Korea University College of Medicine, Seoul, Republic of Korea; Department of Preventive Medicine, Yonsei University College of Medicine, Seoul, Republic of Korea.
Email: min903@korea.ac.kr



Introduction

The rapid advancement of digital technology and the Fourth Industrial Revolution have fundamentally reshaped how health information is accessed and utilized.¹ These transformations have introduced unprecedented opportunities and challenges in the health sector, particularly in how individuals engage with digital health resources.^{2,3} The COVID-19 pandemic further amplified these dynamics,⁴ underscoring the critical role of digital health literacy (DHL) in equipping individuals to adapt to these changes.⁵

The World Health Organization (WHO) defines digital health as the comprehensive integration of digital technologies into health systems, encompassing all stages from inception to operation.⁶ This paradigm extends beyond the conventional scope of eHealth by incorporating advanced technologies, including the Internet of Things (IoT), big data analytics, and artificial intelligence (AI).

Although DHL and eHealth are closely related, they are conceptually distinct.⁷ eHealth primarily refers to the application of information and communication technologies (ICTs) to healthcare services, encompassing electronic health records (EHRs), telemedicine platforms, and mobile health (mHealth) applications.⁸ Its primary focus lies in technology-mediated healthcare delivery and communication, emphasizing the access, storage, and exchange of health information within digital ecosystems⁶.

Conversely, DHL is a critical competency within this domain that transcends the functional use of digital health technologies. It encompasses advanced cognitive and social skills necessary for individuals to effectively seek, understand, evaluate, and apply health information from digital sources. Beyond merely distinguishing accurate from inaccurate information, DHL requires competencies such as navigating complex digital tools, promoting evidence-based decision-making, and addressing health inequities.^{9–11} Recently recognized as a “super social determinant of health,”^{12–14} DHL significantly influences health behaviors and outcomes by empowering individuals to make informed health decisions and engage in self-management of health and well-being.^{15,16} However, insufficient DHL not only hinders effective health promotion and disease management but also exacerbates the digital divide, thereby perpetuating health disparities and systemic inequities.⁹

Since its initial conceptualization by Norman and Skinner in 2006 as part of eHealth literacy (eHL),¹⁷ the concept of DHL has evolved significantly in response to global shifts and digital development.^{7,18} However, the relationship between eHL and DHL, and the impact of societal events such as the Fourth Industrial Revolution and the COVID-19 pandemic on DHL research, remain underexplored.^{19,20} Additionally, previous literature reviews^{2,20} have reported that DHL-related studies began emerging in 2006 and saw a dramatic increase after 2014. However, these studies do not sufficiently account for the impact of

the COVID-19 pandemic. While prior studies have largely relied on bibliometric reviews and literature analyses to provide an overarching view of DHL research,^{2,19,21,22} they often fail to delve into the specific trends and shifts associated with distinct historical events.

This study addresses these gaps by employing text network analysis, a robust methodology capable of identifying patterns and relationships among key terms within academic texts.^{23,24} Unlike traditional methods, text network analysis allows for a nuanced exploration of how research priorities emerge and evolve over time.²⁴ In addition, topic modeling is employed to consolidate and reduce complex textual data into coherent and interpretable topics.²⁵ Analyzing research trends is essential for understanding the progression of scholarly inquiry and informing strategic future directions. By focusing on temporal patterns rather than specific research questions, trend analysis reveals areas of contemporary academic interest.^{23,24} Together, these analytical techniques enable the identification of prevailing research themes and the projection of future trajectories.²³ By categorizing research trends into four distinct periods—defined by the conceptualization of DHL (2006),¹⁷ the official proclamation of the Fourth Industrial Revolution (2016),²⁶ and the declaration of COVID-19 as a global pandemic (2020)²⁷—this study provides a granular view of DHL’s development and highlights pivotal shifts in research focus.

The study aims to analyze DHL-related research trends across three historical events using text network analysis and topic modeling. Specifically, its objectives are to:

1. Identify research trends related to DHL,
2. Perform a comparative analysis of core keyword networks across different time periods, and
3. Examine topic model themes over distinct historical periods.

The significance of this study lies in its multifaceted contributions. Through the application of advanced methodologies, including text network and topic modeling analysis, it systematically delineates key research themes shaped by evolving societal dynamics, offering critical insights into future scholarly trajectories. Beyond its methodological rigor, the study provides profound academic and practical implications by elucidating the developmental trajectory of DHL. These findings serve as a foundational resource for policymakers, educators, and healthcare professionals striving to navigate the challenges of an increasingly digitalized and health-oriented global landscape.

Methods

Data collection

This study was designed as a trend analysis incorporating text network analysis and topic modeling to examine

research trends in DHL. The study period spanned from 1977 to 2022, with data extracted from the Scopus database, focusing on articles within the domain of public health. Scopus is one of the most comprehensive academic databases, encompassing over 25,000 scholarly journals, more than 7000 publishers, and approximately 90 million records.²⁸ The year 1977 was selected as the starting point of the analysis, as it marks the earliest appearance of DHL-related studies in the Scopus database.

To systematically analyze the evolution of DHL research, the timeline was divided into four distinct periods, delineated by historical milestones. Core keyword trends were examined across these periods: period I (emerging era) for 1977–2006; period II (establishment era) for 2007–2016; period III (diffusion era) for 2017–2020; and period IV (post-pandemic era) after 2021. This methodological framework provided a comprehensive view of the progression of DHL research, highlighting shifts in focus and emerging trends across different historical contexts.

We conducted a comprehensive search of the Scopus database to identify articles in the field of public health using the terms “digital health literacy,” “e-health literacy,” and “electronic health literacy” in titles, keywords, and abstracts. This search yielded a total of 4054 articles containing at least one of the specified terms. After screening for relevance to public health and eliminating duplicates, 2645 articles were selected for further analysis.

The distribution of analyzed articles across the defined periods was as follows: Period I comprised 89 articles, Period II included 599 articles, Period III comprised 884 articles, and Period IV included 1073 articles (Figure 1).

To enhance the systematicity and transparency of reporting, the checklist of Bibliometric Reviews of Biomedical Literature (BIBLIO) was employed.²⁹ This checklist comprises 20 essential items that outline the minimum requirements for conducting and reporting bibliometric reviews. Its application ensures methodological rigor and consistency, facilitating reproducibility and the comprehensive assessment of bibliometric analyses (Supplementary material 2).

As this study analyzed publicly accessible literature, it did not involve human participants and was therefore exempt from ethical approval requirements.

Analysis methodologies

Text network analysis. Text network analysis provides enhanced insights by examining relationships among keywords to uncover hidden meanings and can be effectively utilized through visualization. Keywords were extracted for each major period from the abstracts using morpheme analysis. Data analysis was performed using NetMiner 4.0 software (Cyram, Seongnam-si, South Korea).

Stopwords were identified and excluded through an iterative review process conducted by three researchers. Words that appeared consistently across all papers were excluded

from the research trend analysis, as they lacked specificity as keywords. For example, commonly used terms in DHL-related papers, such as “digital,” “health,” and “literature,” were omitted. Similarly, generic terms like “discussion” and “methods” were excluded due to their lack of relevance to the research question (Supplementary 1).

The rankings were assessed based on the centrality of keywords (nodes) in the text network to select the core keywords. Among various centrality indices, this study adopted the “degree centrality” as it was more appropriate for keyword analysis. The degree centrality index measures how many links (connections) the nodes have in the network, identifying keywords that appear simultaneously with other keywords.²³ Degree centrality is expressed as the value between 0 and 1, with a value close to 1 representing large centrality. This study analyzed the frequency and degree centrality of the top 30 keywords for each of the four defined periods, providing a detailed exploration of keyword trends and their evolution over time.

Topic model analysis. Latent Dirichlet Allocation (LDA) is a probabilistic generative model widely utilized for identifying latent topics within a corpus of text. By analyzing the co-occurrence patterns of words, LDA extracts underlying topics and represents each document as a probabilistic distribution over these topics, while each topic is modeled as a distribution over words. This dual representation enables LDA to effectively uncover the thematic structure of large text datasets.^{30–32} Additionally, the model simulates the generative process of creating documents, offering insights into the relationships between words, topics, and documents.

In this study, the LDA algorithm was implemented with 100 iterations, employing the Markov Chain Monte Carlo (MCMC) method for model training. These parameters align with the default configuration of the applied software. The number of topics, which can be adjusted based on the analytical objectives, was set to three to balance interpretability with the granularity of topic differentiation. This choice allowed for a focused and concise analysis of thematic shifts over time.

Results

DHL-related research trends

Table 1 presents an analysis of the top 30 countries and journals contributing to the field of DHL. The United States dominated DHL research in terms of publication ($n = 36.00\%$), with the United Kingdom ($n = 9.38\%$) and Australia ($n = 8.76\%$) following as the second and third most prolific contributors. Notably, the *Journal of Medical Internet Research* ($n = 5.86\%$) and the *International Journal of Environmental Research and Public Health* ($n = 3.67\%$) were identified as the leading journals publishing the highest volume of DHL-related studies.

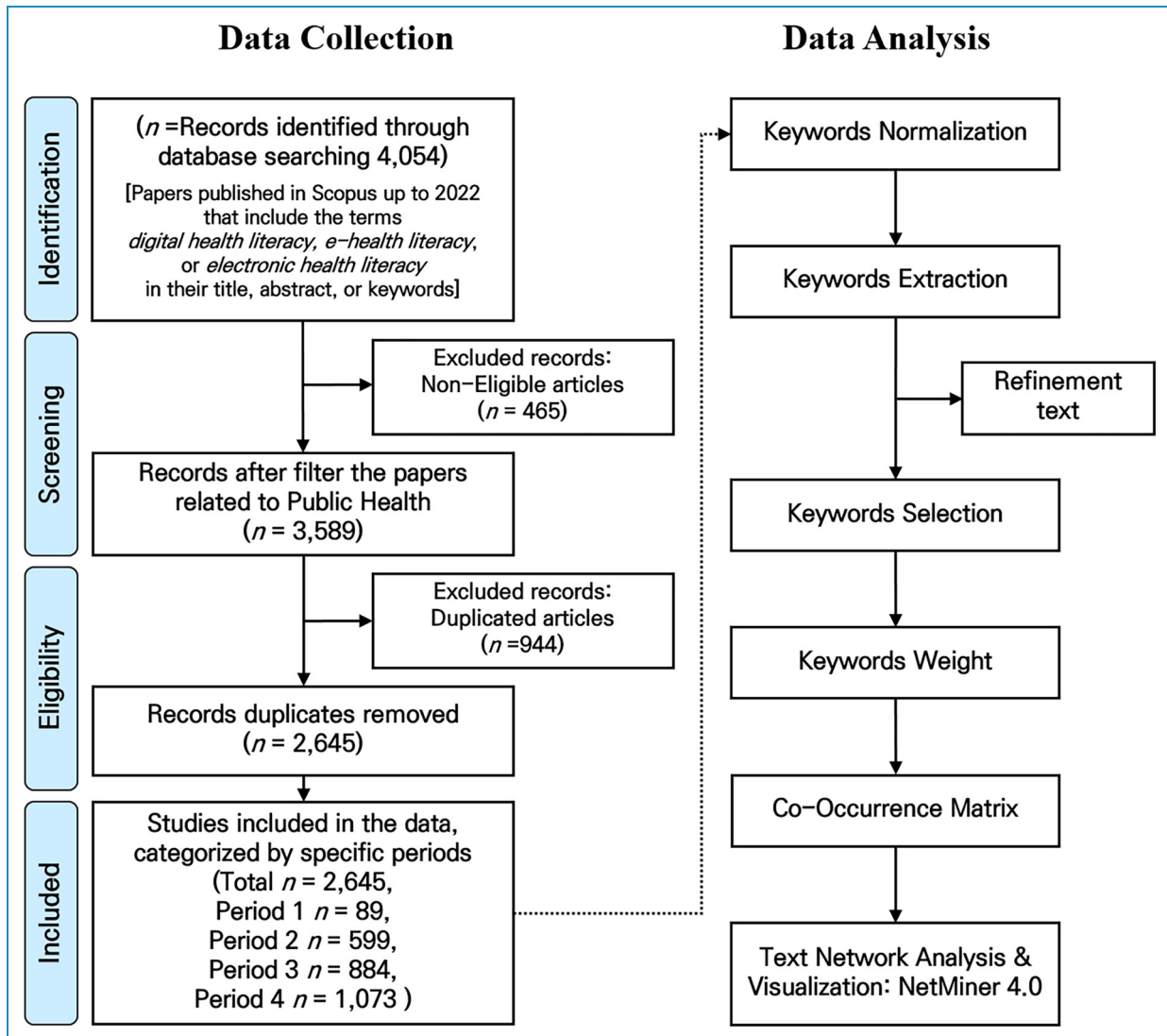


Figure 1. Flow diagram for the data collection (literature search and article selection) and data analysis process.

Figure 2 illustrates the distribution of academic fields related to DHL research. The majority of studies were in the field of Medicine (56%), followed by Social Sciences (16%) and Nursing (9%).

Figure 3 delineates the annual publication trends in academic papers pertaining to DHL. Prior to 2006, the volume of publications remained relatively static, indicating minimal research activity in this domain. A modest yet steady upward trajectory becomes discernible from 2006, transitioning into a pronounced acceleration beginning in 2016, a period conceptualized as the “diffusion era” (Period III). This growth trajectory was further catalyzed during the “post-pandemic era” (Period IV), reflecting the increasing scholarly attention to DHL as a critical framework for addressing global challenges and leveraging technological innovation.

When examining trends by specific terms—“digital health literacy,” “e-health literacy,” and “electronic health literacy”—the trajectory of digital health literacy closely mirrors the overall pattern, with a steeper growth compared to the other terms. In contrast, e-health literacy and electronic health literacy show less pronounced increases, with e-health literacy consistently recording the lowest number of publications among the three categories. Notably, during Period III (diffusion era), the upward trajectories of “digital health literacy” and “electronic health literacy” reverse, reflecting a significant shift in scholarly focus and a pronounced preference for the term “digital health literacy” within academic discourse.

Table 2 presents a comparative analysis of the 30 most frequently appearing keywords in DHL-related papers

Table 1. Leading 30 countries and journals by number of publications (n = 2645)

Rank	Country	Documents	Percent	Source title	Frequency	Percent
1	United States	1106	36.00	Journal of Medical Internet Research	155	5.86
2	United Kingdom	288	9.38	International Journal of Environmental Research and Public Health	97	3.67
3	Australia	269	8.76	JMIR Formative Research	40	1.51
4	Canada	176	5.73	BMC Public Health	37	1.4
5	Germany	146	4.75	BMJ Open	33	1.25
6	Netherlands	102	3.32	International Journal of Medical Informatics	31	1.17
7	China	79	2.57	Patient Education and Counseling	30	1.13
8	India	69	2.25	Journal of the American Medical Informatics Association	28	1.06
9	Denmark	62	2.02	JMIR Research Protocols	27	1.02
10	Spain	60	1.95	JMIR mHealth and uHealth	26	0.98
11	South Korea	56	1.82	BMC Medical Informatics and Decision Making	24	0.91
12	Switzerland	55	1.79	PLoS ONE	23	0.87
13	Italy	49	1.60	BMC Health Services Research	22	0.83
14	Sweden	48	1.56	Frontiers in Public Health	21	0.79
15	Iran	43	1.40	JMIR Human Factors	21	0.79
16	Norway	40	1.30	Health Information and Libraries Journal	20	0.76
17	Brazil	37	1.20	CIN – Computers Informatics Nursing	19	0.72
18	Taiwan	37	1.20	Telemedicine and e-Health	19	0.72
19	France	33	1.07	Contemporary Clinical Trials	17	0.64
20	South Africa	33	1.07	Digital Health	17	0.64
21	Hong Kong	32	1.04	Journal of General Internal Medicine	14	0.53
22	Japan	30	0.98	Trials	13	0.49
23	Belgium	29	0.94	Applied Clinical Informatics	11	0.42
24	Ireland	29	0.94	Journal of Medical Systems	11	0.42
25	Malaysia	28	0.91	Pediatrics	11	0.42
26	Saudi Arabia	28	0.91	Frontiers in Digital Health	10	0.38
27	Israel	27	0.88	JMIR Mental Health	10	0.38
28	Portugal	27	0.88	Journal of Cancer Education	10	0.38
29	Singapore	27	0.88	Journal of Consumer Health on the Internet	10	0.38
30	Turkey	27	0.88	Library Philosophy and Practice	10	0.38

across different periods. Commonly recurring keywords across all periods included “care,” “student,” and “education,” with consistent frequency trends. Notably, the keyword “barrier” exhibited a remarkable increase in prominence, advancing from the 30th position in Period 1 to the 8th position in Period 4.

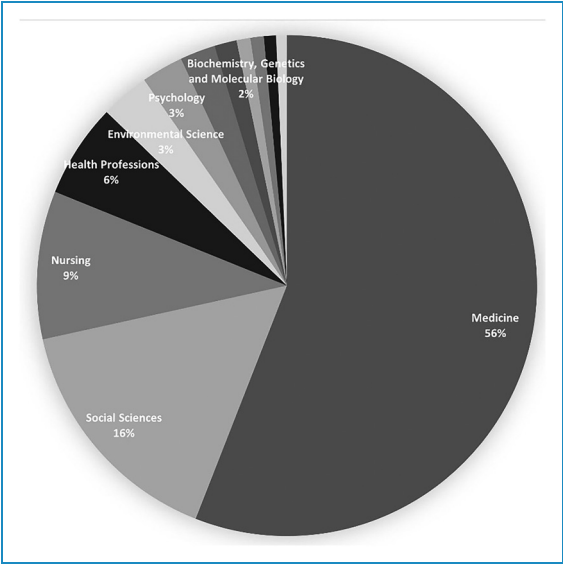


Figure 2. Academic fields related to digital health literacy.

Comparative analysis of core keyword networks based on periods

Table 3 presents the 30 major keywords with the degree centrality according to periods. Words that appeared in all periods, included “care,” “computer,” “education,” “internet,” “service,” “training,” and “student,” which could be considered the identity of the DHL-related research.

Figures 4–7 presents the network map of the top 50 keywords in DHL-related research. Words with high degree centrality are considered the most frequently occurring within the network, representing core concepts of the subject.³³ In the figure, node size corresponds to word frequency, with larger nodes indicating higher prominence.

During Period I (Emerging Era), seven large nodes—such as “care,” “student,” and “internet”—stand out as dominant. Over time, node sizes become more uniform, and the connections between nodes grow increasingly intricate, reflecting the expanding diversity and complexity of concepts related to DHL.

Certain terms emerged exclusively in specific periods, such as “copyright” in Period II, “medicine” in Period III, and “pandemic,” “license,” “platform,” and “device” in Period IV. Notably, “gap” and “disparity” appeared for the first time during the post-pandemic era (Period IV), emphasizing their relevance to recent research trends.

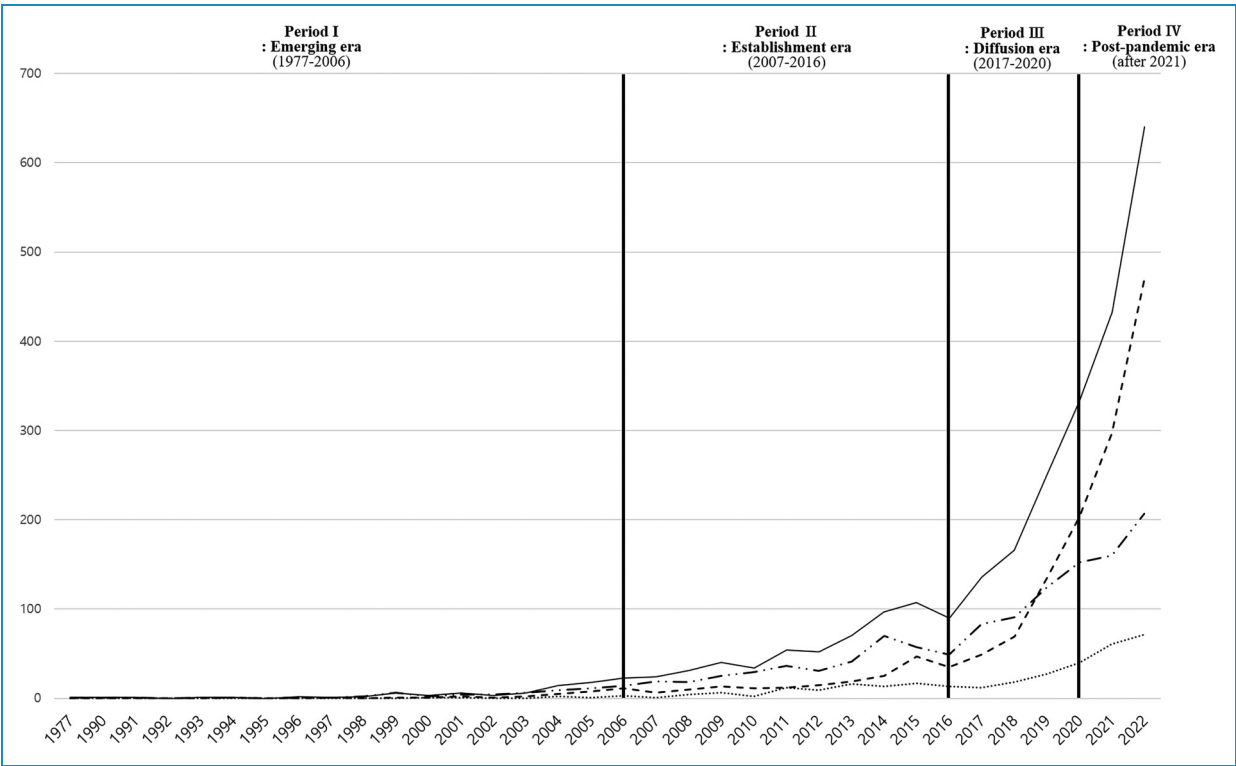


Figure 3. Annual publication trends in digital health literacy. Note: Solid lines represent publications used in this analysis; dotted lines represent e-health literacy; dashed lines represent digital health literacy; and dashed-dotted lines represent electronic health literacy.

Table 2. Frequency of keywords according to the four periods.

Rank	Overall		Period I (1977–2006)		Period II (2007–2016)		Period III (2017–2020)		Period IV (2021)	
	Words	Frequency	Words	Frequency	Words	Frequency	Words	Frequency	Words	Frequency
1	care	2990	computer	98	care	645	care	698	care	1556
2	student	1723	care	91	education	406	education	336	student	963
3	education	1607	student	67	student	368	student	325	education	817
4	service	1240	nurse	65	computer	367	medication	289	service	741
5	adult	1180	internet	50	internet	363	service	261	adult	705
6	internet	1168	education	48	communication	288	internet	248	app	528
7	communication	942	record	46	record	268	communication	234	internet	507
8	barrier	937	library	45	adult	246	age	223	barrier	502
9	age	912	nursing	44	nurse	228	app	218	disease	497
10	medication	822	service	42	medication	204	adult	218	pandemic	493
11	disease	787	training	38	age	200	barrier	217	age	483
12	app	784	staff	34	barrier	198	provider	216	score	449
13	provider	770	woman	30	service	196	risk	206	community	423
14	score	763	hospital	28	diabetes	181	record	187	interview	418
15	risk	759	doctor	28	provider	172	diabetes	183	risk	411
16	interview	721	clinician	28	hospital	166	treatment	168	communication	395
17	community	721	pdas	25	library	165	interview	165	provider	372
18	computer	720	consumer	25	woman	162	child	160	treatment	342
19	record	702	communication	25	score	159	woman	156	woman	328
20	woman	676	web	23	physician	151	training	153	university	318

(continued)

Table 2. Continued.

Rank	Overall	Period I (1977–2006)			Period II (2007–2016)			Period III (2017–2020)			Period IV (2021)		
		Frequency	Words		Frequency	Words		Frequency	Words		Frequency	Words	
21	treatment	644	child	23	adoption	148	disease	151	medication	316			
22	diabetes	631	science	22	adherence	148	score	150	challenge	308			
23	training	621	physician	22	training	142	adherence	149	mhealth	291			
24	adherence	570	software	21	community	142	community	142	training	288			
25	hospital	568	center	21	consumer	135	portal	140	visit	277			
26	child	557	safety	20	disease	132	hospital	136	family	274			
27	challenge	536	practitioner	20	risk	127	application	128	diabetes	267			
28	family	532	family	20	interview	126	mhealth	125	child	266			
29	university	519	email	20	treatment	124	professional	124	adherence	257			
30	nurse	516	barrier	20	professional	124	family	122	application	254			

Table 3. Comparative analysis of core keyword networks according to the four periods

Rank	Period I (1977–2006)		Period II (2007–2016)		Period III (2017–2020)		Period IV (after 2021)	
	Words	Degree centrality	Words	Degree centrality	Words	Degree centrality	Words	Degree centrality
1	care	0.425926	care	0.402597	care	0.427711	care	0.358389
2	computer	0.37963	education	0.329004	education	0.292169	education	0.283893
3	education	0.287037	record	0.262626	age	0.233434	age	0.232215
4	nurse	0.268519	computer	0.217893	communication	0.213855	service	0.219463
5	record	0.240741	age	0.212121	record	0.194277	adult	0.197315
6	library	0.231481	internet	0.199134	barrier	0.177711	internet	0.188591
7	internet	0.203704	communication	0.194805	provider	0.173193	disease	0.187248
8	service	0.194444	adult	0.177489	adult	0.162651	pandemic	0.185906
9	training	0.194444	barrier	0.167388	interview	0.162651	interview	0.177852
10	student	0.175926	provider	0.152958	service	0.158133	barrier	0.177181
11	barrier	0.148148	service	0.151515	internet	0.150602	communication	0.171812
12	environment	0.138889	student	0.128427	disease	0.14006	community	0.163087
13	web	0.138889	training	0.126984	relationship	0.129518	score	0.155034
14	nursing	0.12963	interview	0.122655	challenge	0.128012	challenge	0.15302
15	opportunity	0.12963	community	0.121212	risk	0.128012	risk	0.148993
16	science	0.12963	hospital	0.121212	benefit	0.123494	student	0.141611
17	staff	0.12963	copyright	0.119769	hospital	0.120482	university	0.134899
18	hospital	0.12037	disease	0.119769	score	0.11747	provider	0.128188
19	librarian	0.12037	challenge	0.112554	community	0.115964	treatment	0.118792
20	life	0.12037	nurse	0.111111	treatment	0.112952	life	0.113423

(continued)

Table 3. Continued.

Rank	Period I (1977–2006)		Period II (2007–2016)		Period III (2017–2020)		Period IV (after 2021)	
	Words	Degree centrality	Words	Degree centrality	Words	Degree centrality	Words	Degree centrality
21	professional	0.12037	treatment	0.106782	lack	0.111446	training	0.108054
22	woman	0.12037	question	0.105339	professional	0.10994	relationship	0.107383
23	communication	0.111111	lack	0.103896	application	0.106928	license	0.10604
24	consumer	0.111111	benefit	0.10101	training	0.106928	family	0.099329
25	practitioner	0.111111	relationship	0.10101	student	0.105422	hospital	0.098658
26	software	0.111111	clinic	0.099567	question	0.103916	platform	0.097315
27	clinician	0.101852	professional	0.098124	family	0.100904	application	0.092617
28	interview	0.101852	family	0.096681	university	0.099398	device	0.084564
29	question	0.101852	university	0.096681	computer	0.09488	woman	0.084564
30	university	0.101852	score	0.095238	medicine	0.09488	clinic	0.071141

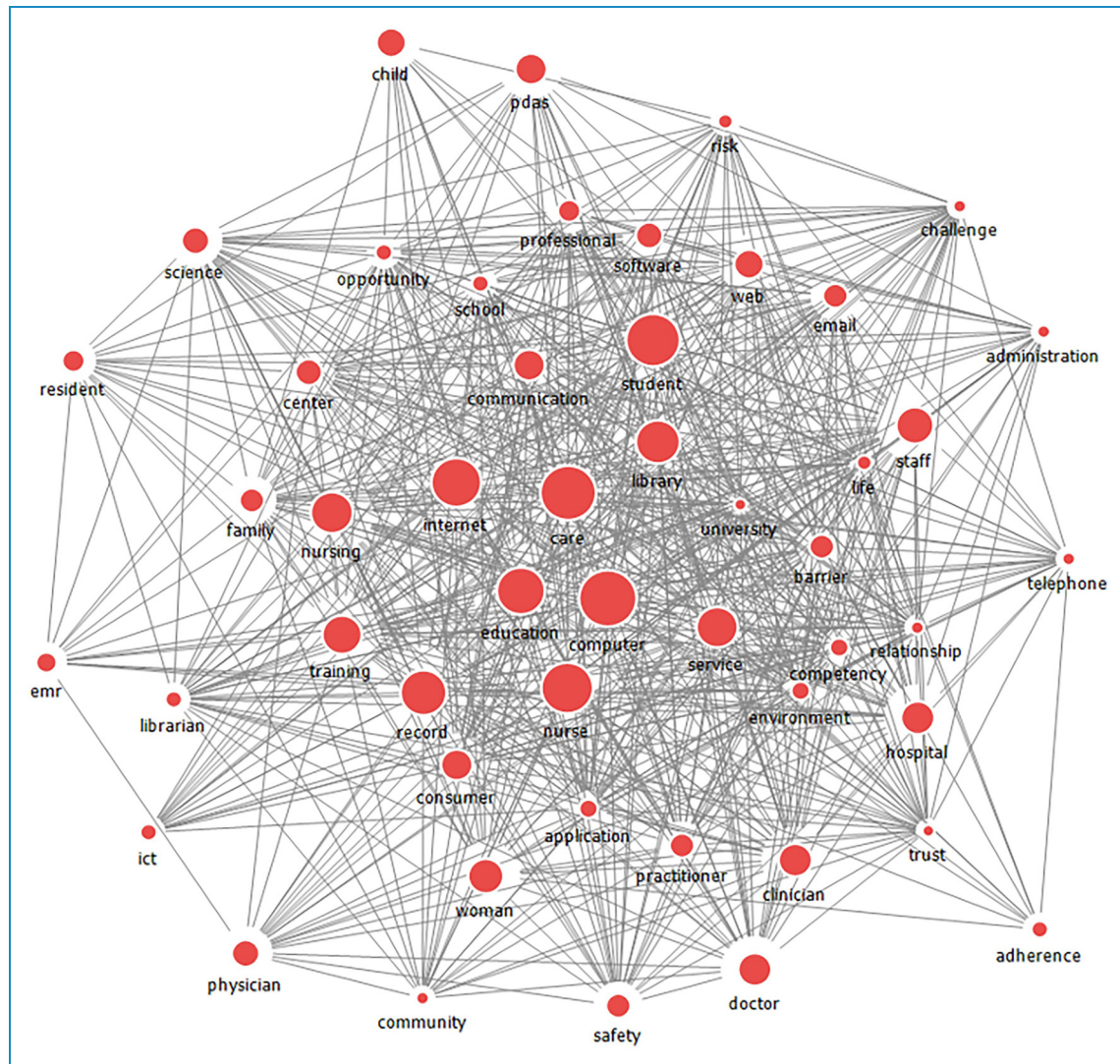


Figure 4. Network map of the top 50 keywords in Periods I.

Comparison of topic model topics by period

Table 4 presents the findings of the topic modeling analysis, providing a comprehensive overview of the evolution of research themes across four distinct periods. Although the overarching emphasis on DHL remained consistent, thematic shifts aligned with changing priorities and societal needs were discernible. During Period I (1977–2006), research primarily centered on hospital-related concepts, with predominant clusters including terms such as “electronic medical records (EMR),” “hospital,” and “doctor.” In Period II, clusters evolved to include terms such as “system,” “computer,” “physician,” and “EHR,” reflecting a transition toward technological advancements and digitalization in healthcare. By Period IV, thematic clusters incorporated terms such as “healthcare,” “provider,” “barrier,” and “implementation,” signaling an expanded focus on public health, community involvement, and the broader societal implications of DHL.

Discussion

Milestones in the evolution of DHL research

Norman and Skinner's foundational conceptualization (2006). The analysis of research trends before and after 2006, defined in this study as the year of the first conceptualization of DHL, reveals distinct patterns. Prior to 2006, the volume of publications remained relatively static, indicating minimal research activity in this domain. However, from 2006 onward, a modest yet steady increase in scholarly output becomes evident, reflecting the growing recognition of DHL as an important research focus. This observation is further supported by the text network analysis results presented in Figure 4. During Period I (Emerging Era), a small number of dominant nodes—such as “care,” “student,” and “internet”—emerged as central terms. Over time, the network structure evolves, with nodes becoming more evenly distributed and interconnections becoming increasingly

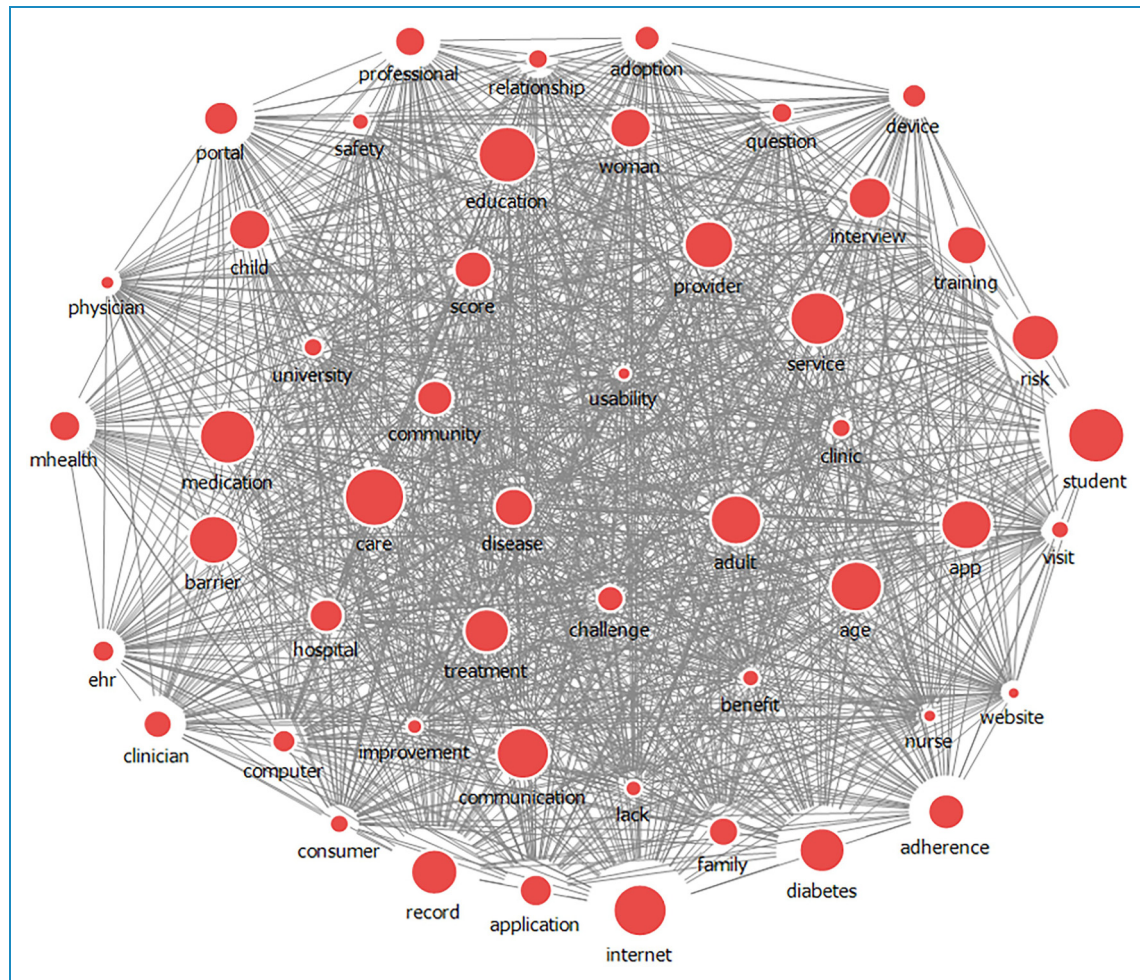


Figure 6. Network map of the top 50 keywords in Periods III.

based technologies and digital communication tools, “digital health” extends to a broader and more inclusive framework, encompassing a diverse range of advanced technologies that capture the increasingly complex and evolving digital health landscape.³⁵

The rapid evolution and integration of these technologies have profoundly influenced the conceptualization and scholarly understanding of DHL. This study reaffirms the pivotal role of technological advancements in shaping the trajectory of DHL research, highlighting how the adoption of emerging technologies continues to drive academic discourse and redefine the field’s scope and application.

The COVID-19 pandemic as a global catalyst (2020). This study highlights significant accelerations in DHL research during pivotal periods such as the Fourth Industrial Revolution and the COVID-19 pandemic. These observations align with prior literature, such as Yang et al. (2022),²⁰ which delineated distinct phases in DHL research, including an incubation period (1998–2005), a slow growth period (2006–2013), and a rapid growth period (after 2014).

The current study advances this framework by introducing the “post-pandemic era” as a discrete and pivotal phase, emphasizing the significant impact of macro-environmental disruptions on shaping research trajectories.

A comparative analysis of these periods indicates that the COVID-19 pandemic exerted a more immediate and transformative influence on DHL research compared to the gradual impact of technological advancements observed during the diffusion era (2016 onward), as illustrated in Figure 3. Furthermore, the text network analysis results from this study identify the emergence of terms such as “pandemic,” “license,” “platform,” and “device” exclusively during Period IV. For instance, individuals were compelled to discern accurate information about COVID-19 amidst an overwhelming influx of misinformation to maintain their daily lives.² Simultaneously, children and adolescents adapted to utilizing platforms and devices for health education purposes.³⁶ These circumstances significantly heightened research interest in DHL, as reflected in the increased scholarly output on DHL-related topics (e.g., platform, device) during the pandemic. Aligning

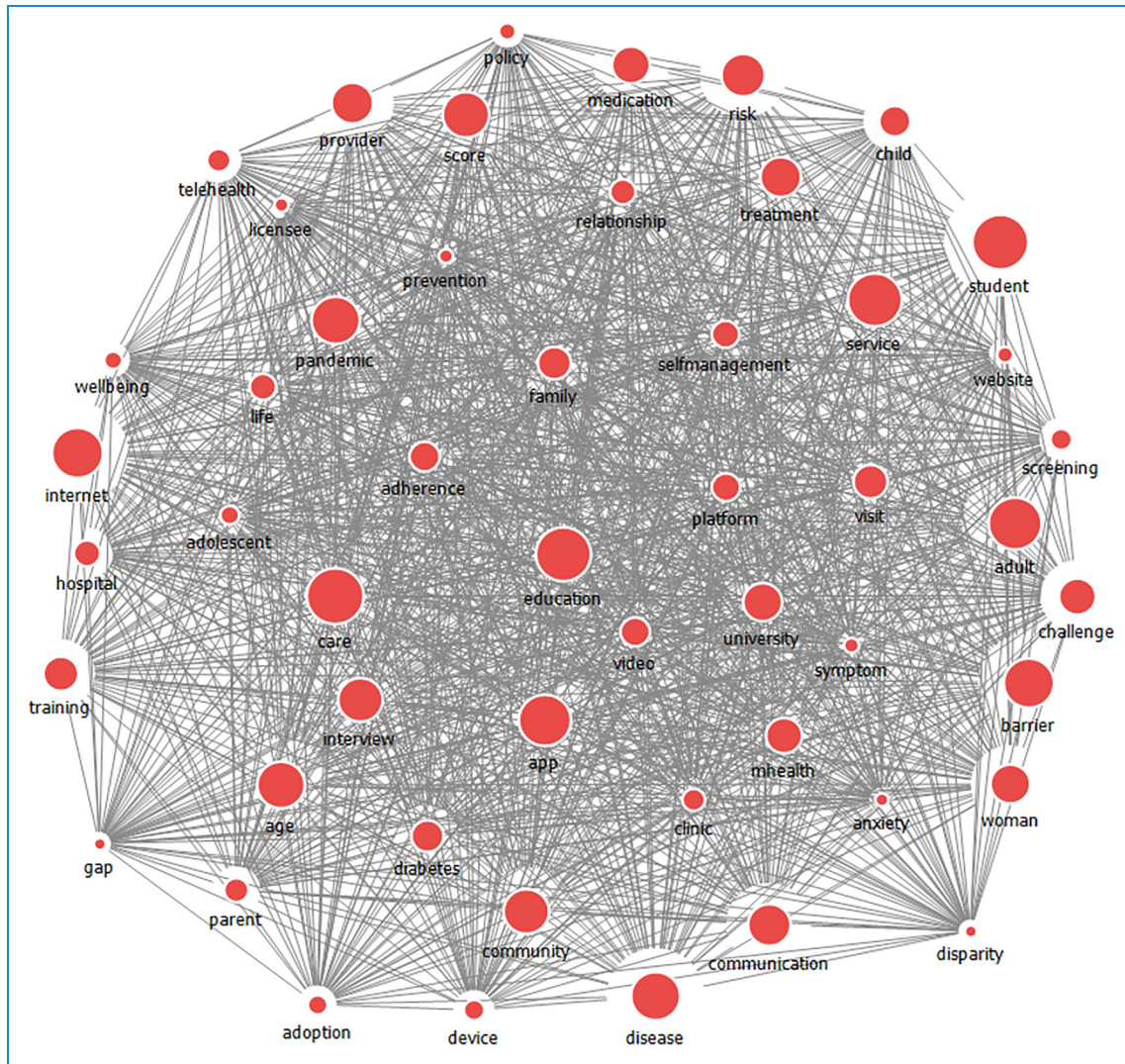


Figure 7. Network map of the top 50 keywords in Periods IV.

Table 4. Comparison of topics according to the four periods.

	Period I (1977–2006)		Period II (2007–2016)		Period III (2017–2019)		Period IV (2020–2022.10)	
	Topic words	Documents	Topic words	Documents	Topic words	Documents	Topic words	Documents
Topic 1	Internet, EMR, Hospital, Resident, Doctor	20	Internet, Student, Skill, Tool, Computer	178	Service, Healthcare, App, Provider, Person	173	Student, Behavior, Adult, Practice, Disease	295
Topic 2	User, Healthcare, Database, Event, Child	19	System, Practice, Computer, Physician, EHR	188	Medication, Person, Diabetes, Adherence, Provider	136	App, Adult, Disease, Medication, Adherence	506
Topic 3	Student, Clinician, Intervention, PDAs, Hospital	50	Intervention, Medication, Adult, Tool, Adherence	233	Student, Internet, Person, Provider, Service	243	Healthcare, Provider, Barrier, Implementation, Practice	604

with findings from prior studies,³⁷ this analysis underscores the catalytic role of global crises in accelerating research priorities and directing scholarly focus toward addressing critical gaps in DHL.

Finally, during this period, novel terms such as “gap” and “disparity” emerged, though they did not rank among the top 30 keywords. Furthermore, the topic modeling analysis identified significant clusters, including “barrier” and “implementation,” reflecting an increasing focus on addressing systemic challenges and enhancing health outcomes within community and population health contexts. This thematic evolution underscores the growing recognition of DHL as a critical determinant of health equity and a fundamental mechanism for mitigating the digital divide and reducing health disparities in an increasingly interconnected global landscape.

Recent scholarly discourse has highlighted DHL’s potential to become a prominent social determinant of health, contributing to future health inequalities and exacerbating the digital divide.^{12,13,38} Given that reducing health disparities remains a central objective of the global health agenda,³⁹ research on the inequities associated with DHL is anticipated to grow rapidly, aligning with efforts to promote equitable access to digital health resources worldwide.

Geographical trends and future directions in DHL research

This study identified the United States as the leading contributor to research on DHL, with the United Kingdom and Australia ranking as the second and third most prolific contributors, respectively. These findings are consistent with prior studies that have highlighted the United States, Australia, the United Kingdom, and Canada as dominant contributors to advancements in DHL research.²⁰ This underscores the significant role of Western nations in shaping the field and driving innovation.

However, the development and characteristics of digital health systems within healthcare sectors exhibit substantial variability across countries.⁴⁰ This highlights the necessity of expanding research efforts to include diverse geographical and cultural contexts. Future research should prioritize examining DHL among populations in Asia and developing regions, focusing on identifying context-specific influencing factors and addressing unique challenges. Such efforts are essential to expanding the scope and inclusivity of DHL studies.

Limitations of this study

First, limitations related to text network and topic model analysis must be considered. While these methods effectively identify research trends by analyzing a large volume of studies in a short period,^{23–25} they rely on key terms and their connection strengths to infer results. This approach may

fail to capture deeper conceptual meanings, potentially leading to lower concreteness in the findings.⁴¹

Second, the exclusion of certain keywords as stopwords presents a methodological limitation. To enhance the precision of keyword analysis and prevent overrepresentation, commonly used terms such as “digital” and “health” were removed. However, this decision may have unintentionally excluded critical terms that contribute to nuanced interpretations within specific subfields. As a result, the omission of these terms may have influenced the formation of thematic clusters and the structure of keyword networks by altering the co-occurrence patterns of domain-specific terminology.

Third, in this study, the number of topics was set to three based on the researchers’ judgment, considering the interpretability of the extracted topics and the research context. Future studies should consider applying quantitative evaluation metrics, such as Perplexity or Coherence Score, to enhance the objectivity and robustness of topic modeling results.

Fourth, limitations related to inclusion criteria and study scope should be noted. Since this study analyzed only English-language publications, publication bias may have occurred, leading to an overrepresentation of research conducted in English-speaking countries and limiting the generalizability of the findings to non-English contexts. Additionally, a significant portion of the analyzed literature originates from Western countries, which may further restrict the global representativeness of DHL research trends.

Finally, limitations regarding publication types must be acknowledged. This study focused exclusively on journal articles, excluding gray literature and conference proceedings, which often contain important research findings and emerging trends. The exclusion of such sources may have influenced the comprehensiveness of the study’s findings. Therefore, recognizing these limitations and clearly defining the study’s scope is essential for a more holistic understanding of the field.

Conclusions

This study systematically analyzed digital health literacy (DHL)-related literature using text network and topic modeling approaches. The findings revealed the evolution of the DHL concept in alignment with three significant historical events. Notably, the diversity of DHL-related research has increased over time compared to its initial emergence (Period I), underscoring its dynamic and evolving nature shaped by historical contexts. Future research is anticipated to expand DHL-related studies, focusing on specific patient populations and diverse groups, including low-income and vulnerable populations, to address their unique needs and challenges.

ORCID iD

Mingee Choi  <https://orcid.org/0000-0001-9383-6402>

Statements and declarations

Author Contributions/CRedit

Jiyoung Park: conceptualization, funding acquisition, investigation, project administration, supervision, validation, writing—original draft, writing—review and editing; Seohyun Won: writing—original draft, writing—review and editing; Mingee Choi: software, data curation, methodology, validation; Chul Hee Kang: supervision, conceptualization, writing—review and editing, Han Shi Jocelyn Chew: supervision, writing—original draft, writing—review and editing

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Research Foundation of Korea (grant number 2020R1A2C4096046, 2022R1C1C1009609).

Conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Data availability

Derived data supporting the findings of this study are available from the corresponding author Mingee Choi on request.

Supplemental material

Supplemental material for this article is available online.

References

1. Ahsan MM and Siddique Z. Industry 4.0 in healthcare: a systematic review. *Int J Info Manage Data Insights* 2022; 2: 100079.
2. Wang C, Wu X and Qi H. A comprehensive analysis of e-health literacy research focuses and trends. *Healthcare (Basel)* 2021; 10: 66.
3. Araújo NMF. Impact of the fourth industrial revolution on the health sector: a qualitative study. *Healthc Inform Res* 2020; 26: 328–334.
4. OECD. *OECD Digital Economy Outlook 2020. Report*. Paris: OECD Publishing, 2020.
5. Dib F, Mayaud P, Chauvin P, et al. Online mis/disinformation and vaccine hesitancy in the era of COVID-19: why we need an eHealth literacy revolution. *Hum Vacc Immunother* 2022; 18: 1–3.
6. World Health Organization. *Global strategy on digital health 2020–2025. Report*. Geneva, Switzerland: World Health Organization, 2021.
7. Ban S, Kim Y and Seomun G. Digital health literacy: A concept analysis. *Digital health* 2024; 10: 20552076241287894.
8. World Health Organization. *Global diffusion of eHealth: making universal health coverage achievable: report of the third global survey on eHealth*. Geneva: World Health Organization, 2017.
9. Arias López MDP, Ong BA, Borrat Frigola X, et al. Digital literacy as a new determinant of health: a scoping review. *PLOS Digital Health* 2023; 2: e0000279.
10. Le C, Finbråten HS, Pettersen KS, et al. *Health Literacy in the Norwegian Population. Report*. Oslo, Norway: The Norwegian Directorate of Health, 2021.
11. The HLS19 Consortium of the WHO Action Network M-POHL. *International Report on the Methodology, Results, and Recommendations of the European Health Literacy Population Survey 2019–2021 (HLS19) of M-POHL. Report*. Vienna, Austria: The HLS19 Consortium of the WHO Action Network M-POHL. 2021.
12. van Kessel R, Wong BLH, Clemens T, et al. Digital health literacy as a super determinant of health: more than simply the sum of its parts. *Internet Interv* 2022; 27: 100500.
13. Sieck CJ, Sheon A, Ancker JS, et al. Digital inclusion as a social determinant of health. *Npj Digit Med* 2021; 4: 52.
14. Crawford A and Serhal E. Digital health equity and COVID-19: the innovation curve cannot reinforce the social gradient of health. *J Med Internet Res* 2020; 22: e19361.
15. Hsu W, Chiang C and Yang S. The effect of individual factors on health behaviors among college students: the mediating effects of eHealth literacy. *J Med Internet Res* 2014; 16: e287.
16. Yuen E, Winter N, Savira F, et al. Digital health literacy and its association with sociodemographic characteristics, health resource use, and health outcomes: rapid review. *Interact J Med Res* 2024; 13: e46888.
17. Norman CD and Skinner HA. eHEALS: the eHealth literacy scale. *J Med Internet Res* 2006; 8: 27.
18. Dunn P and Hazzard E. Technology approaches to digital health literacy. *Int J Cardiol* 2019; 293: 294–296.
19. Wang X and Luan W. Research progress on digital health literacy of older adults: a scoping review. *Front Public Health* 2022; 10: 906089.
20. Yang K, Hu Y and Qi H. Digital health literacy: bibliometric analysis. *J Med Internet Res* 2022; 24: e35816.
21. Estrela M, Semedo G, Roque F, et al. Sociodemographic determinants of digital health literacy: a systematic review and meta-analysis. *Int J Med Inf* 2023; 177: 105124.
22. Verweel L, Newman A, Michaelchuk W, et al. The effect of digital interventions on related health literacy and skills for individuals living with chronic diseases: a systematic review and meta-analysis. *Int J Med Inf* 2023; 177: 105114.
23. Drieger P. Semantic network analysis as a method for visual text analytics. *Procedia Soc Behav Sci* 2013; 79: 4–17.
24. Paranyushkin D. Identifying the pathways for meaning circulation using text network analysis. *Venture Fiction Practices* 2011; 2(4): 1–26.
25. Churchill R and Singh L. The evolution of topic modeling. *ACM Comput Surv* 2022; 54: 1–35.
26. Schwab K. *The Fourth Industrial Revolution*. 1st ed. Geneva, Switzerland: World Economic Forum, 2016.
27. World Health Organization. Archived: WHO Timeline – COVID-19. 2020, www.who.int/news/item/27-04-2020-who-timeline-covid-19 (accessed 13 January 2025).

28. Scopus. Archived: Scopus now includes 90 million + content records! 2023, blog.scopus.com/posts/scopus-now-includes-90-million-content-records (accessed 26 January 2025).
29. Montazeri A, Mohammadi S, Hesari PM, et al. Preliminary guideline for reporting bibliometric reviews of the biomedical literature (BIBLIO): a minimum requirements. *Syst Rev* 2023; 12: 39.
30. Blei DM, Ng AY and Jordan MI. Latent dirichlet allocation. *J Mach Learn Res* 2003; 3: 993–1022.
31. Griffiths TL and Steyvers M. Finding scientific topics. *Proc Natnl Acad Sci* 2004; 101: 5228–5235.
32. Steyvers M, Griffiths T, et al. Probabilistic topic models. In: Landauer TK, McNamara DS and Dennis S (eds) *Handbook of latent semantic analysis*. 1st ed. London, UK: Psychology Press, 2007, 439–460.
33. Li F and Du TC. Who is talking? An ontology-based opinion leader identification framework for word-of-mouth marketing in online social blogs. *Decis Support Syst* 2011; 51: 190–197.
34. Popov VV, Kudryavtseva EV, Kumar Katiyar N, et al. Industry 4.0 and digitalisation in healthcare. *Materials (Basel)* 2022; 15: 2140.
35. World Health Organization. Archived: Digital health. 2022, www.who.int/europe/health-topics/digital-health#tab=tab_1 (accessed 26 January 2025).
36. Pokhrel S and Chhetri R. A literature review on impact of COVID-19 pandemic on teaching and learning. *High Educ Future* 2021; 8: 133–141.
37. Wang C and Si L. A bibliometric analysis of digital literacy research from 1990 to 2022 and research on emerging themes during the COVID-19 pandemic. *Sustainability* 2023; 15: 5769.
38. Saeed SA and Masters RM. Disparities in health care and the digital divide. *Curr Psychiatry Rep* 2021; 23: 61.
39. United Nations. *COVID-19, Inequalities and Building Back Better an opportunity to build back better: policy brief by the HLCP Inequalities Task Team. Report*. New York: UN, 2020.
40. Sharma R and Kshetri N. Digital healthcare: historical development, applications, and future research directions. *Int J Inf Manage* 2020; 53: 102105.
41. Chen Y, Peng Z, Kim SH, et al. What we can do and cannot do with topic modeling: a systematic review. *Commun Methods Meas* 2023; 17: 111–130.