MEDINFO 2021: One World, One Health – Global Partnership for Digital Innovation
P. Otero et al. (Eds.)

doi:10.3233/SHTI220210

The Gap Between Technology and Ethics, Especially in Low- and Middle- Income Country Health Information Systems: A Bibliometric Study

James Thomas[a], Rebecca Carlson[b], Michelle Cawley[b], Quan Yuan[a], Victoria Fleming[c], Fei Yu[b]

[a] Gillings School of Global Public Health, Carolina Population Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA,
[b] Health Sciences library, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA,
[c] Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

Abstract

Digital health information systems are quickly replacing paper systems worldwide. This study examined scholarly work reflecting how ethical considerations for health information systems (HISs) have evolved over time, with particular interest in the contributions from low- and middle-income country (LMIC) institutions. We systematically searched four research databases for terms related to HISs, ethics, and LMICs; and identified 601 relevant articles published from 1975 through 2019. The included publications were produced by 1,000 authors from more than 700 institutions. However, only 30 publications were co-authored by researchers from both an LMIC and a high-income country (HIC). Most publications pertained to data security. There is an acute need to address a wider array of HIS ethics topics, including those that may be unique to LMIC resource constraints. We recommend more collaborations between LMIC and HIC institutions to address the full range of HIS ethical concerns in LMICs.

Keywords:

Ethics, Information Systems, Developing Countries

Introduction

Providing evidence for policy and program decisions and supporting better health outcomes, health information systems (HISs) encompass all health data sources required by a country to plan and implement its national health strategy [1]. Although HISs are defined at different levels of scope by countries and organizations, one of the core components is information on the health needs of a patient, and on the health services provided. For generations this information was recorded on paper and stored in files kept by the service provider. If a patient visited several providers, each of them kept their own records. To monitor disease trends, providers would have to transcribe information from one piece of paper to another and then send it to a district or national office to be tabulated with other reported information.

Digital HISs began to replace paper-based HISs in the 1970s, as commercially available electronic medical record (EMR) systems became available. In the 1990s, the rapid growth of the internet facilitated both the sharing of patient data within a network of providers through electronic health records (EHRs) and the sharing of data summaries with health agencies (e.g., District Health Information System), enabling disease trend monitoring. However, the risks to patients in having their personal data shared broadly were addressed in the United States in 1996 with the federal Health Insurance Portability and Accountability Act (HIPAA), which regulated how data are to be legally shared, including the patient's rights in what is shared. In 2014, the American Health Information Management Association published eight Information Governance Principles for Healthcare including accountability, transparency, integrity, protection, compliance, availability, retention, and disposition [2]. The Principles for Digital Development, endorsed by more than 250 organizations globally in 2016, specified nine ethical principles for HIS development, including designing with the user, designing for sustainability, addressing privacy and security; and using open standards, open data, open source, and open innovation [3]. In 2019, a scoping review was conducted for a broader range of ethical concerns on HISs and identified 34 principles, demonstrating that the range of concerns extends well beyond data security, the most commonly recognized ethical concern for digital data [4]. A few of these 34 principles were transparent processes in designing the system, collecting the minimum amount of data necessary, employing a data steward, and providing access to the source community.

While the ethical concerns were being identified, development agencies such as the United States Agency for International Development's (USAID) began to invest in the creation and strengthening of digital HISs in low-and-middle-income countries (LMICs) [5]. As HISs are implemented in LMICs, perspectives from those countries are needed to identify and address their ethical issues because they will have concerns that are particular to their context [6]. For example, the digital literacy of their populations (the ability to collect, analyze and use digital data) might be less than in developed countries, which can affect how the managers of an HIS should most effectively practice transparency [7].

To the best of our knowledge, no previous work has systematically investigated the research activities and evidence concerning HIS ethics, especially in LMICs. Therefore, this study sought to understand how ethical considerations for HISs in the scholarly literature have evolved overtime, and where ethical thinking may need to catch up with quickly evolving technologies. We were particularly interested in the participation of LMIC institutions in these discussions.

Methods

This study took a bibliometric approach to systematically retrieve and identify research publications that addressed ethical issues associated with HISs. Particularly, we investigated the research output, researchers, research organizations, organizational collaborations, and research topics involved in HIS ethics studies from both LMICs and high-income countries (HICs).

Literature search

Two medical librarians (RC & MC) developed comprehensive search strategies for related concepts using subject headings and keywords, and then executed them in PubMed, Embase, Web of Science (WoS), and Scopus for literature published until December 31, 2019. The search terms related to ethics (e.g., ethic, ethical, bioethics, moral, medical ethics), ethics concepts (e.g., transparency, justice, accountability, equity), information systems (e.g., health information exchange, information policy, data policies), electronic health records (e.g., EHR, electronic medical record, EMR), LMICs (e.g., developing country, less developed country, less developed nation, middle income nation), and the name of each LMIC. We excluded terms specific to certain research types (e.g., clinical trial, controlled trial, RCT).

Citation prioritization & screening

The citation records of retrieved publications were downloaded, aggregated, and then prioritized using machine learning [8] with the DoCTER software [9] before they were imported to Covidence [10] for manual screening (Figure 1). To be included, a publication must have explicitly, not tangentially, addressed HIS and an ethical consideration, and not focus on biomedical research ethics. A study was excluded if its citation record has missing crucial data fields (i.e., title, year, or abstract) and its full text was not accessible with the current institutional journal subscription license and interlibrary loan service. JT and VF independently screened the title and abstract for each article using Covidence.

Data extraction & analysis

The citation records of included articles were exported from Covidence, and then matched back to Scopus using digital object identifier (DOI), title, and year. Scopus has a much broader journal coverage compared to PubMed, Embase, and WoS [11]. If an included record can't be identified in Scopus, it would be matched to PubMed, WoS, or Embase subsequently. The citation fields of matched records were downloaded for publication year, title, author name, author affiliation, source title, and abstract. The downloaded citation fields from these four databases were then standardized using the citation format of Scopus to facilitate data analysis. For example, "Bevanda, M; Colakovic, M" from WoS was converted to "Bevanda M., Colakovic M." (Scopus format). In addition, the author affiliations were converted to the general level. For example, "Harvard Medical School" and "Harvard School of Public Health" were standardized as "Harvard University."

Several bibliometric measures and tools were applied to the standardized data for analysis. **Research output** was measured by the number of publications by year, the number of publications by document types, and the number of publishing venues (source titles). The statistics and visualization were produced in Microsoft Excel. **Researcher** analysis focused on the total number of unique authors contributing to the included studies and the top authors who published the most. This analysis was conducted using Microsoft Excel and the author analysis function in VOSviewer [12]. **Research organizations** were investigated at two levels: (1) the number of all unique author affiliations, the top organizations that the authors affiliate with contributing the most to the included studies, and organization collaboration network; (2) the number of countries that the authors were from. The organizational analysis function in VOSviewer was used for this investigation. **Research landscape** was explored by extracting key terms with high occurrence from the title/abstract and keywords/MeSH of the included studies using both VOSviewer and Biblioshiny [13].

Results

Research output

After citation prioritization, 2,896 out of 6,914 citations produced from citation database search were determined as high likehood of relevance and proceeded with manual screening. Finally, a total of 601 publications were included in this study (Figure 1).

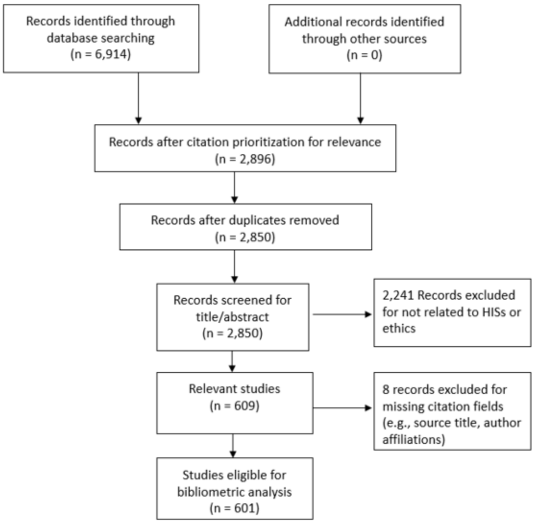

Figure 1– Study Screening and Selection Diagram

Generally, the number of relevant publications increased steadily by an average of 1.2 papers per year from 1975 to 2018. Then in 2019 the number doubled from 52 (2018) to 110 papers (Figure 2).

Figure 2– Publication Distribution by Year (1975-2019)

In rank order, the publication types were peer-reviewed journal research articles (273, or 45%), conference abstracts (162), conference reviews (60), peer-reviewed journal reviews (52), book chapters (33), journal commentaries (11), books (7), and journal editorials (3). The most common publication source was *Studies in Health Technology and Informatics* (45). The next most frequent source, *International Journal of Medical Informatics* published a third as many (15); followed by *Lecture Notes in Computer Science* (11), and *Communications in Computer and Information Science* (10).

Researcher

The publications listed 1,000 authors from 790 unique institutions in the world. Most authors appeared only once, in a single publication; 11 authors had three publications, and two authors from LMICs had four publications.

Research organization

Among the institutional affiliations, the most frequently named was Harvard University (14). Indiana University was cited as an institutional affiliation for 9 publications. An additional 7 cited the Regenstrief Institute which is affiliated with Indiana University. The US Centers for Disease Control and Prevention was cited for 9 articles.

Of the 78 author countries named, 49 countries were high income, 10 were middle income, and 19 were low income. The country with the most authors was the United States (149 publications), followed by the United Kingdom (UK) (49), Australia (43), and Canada (30), all of which are HICs. India was the middle-income country with the most publications included in this study (27), followed by China (18), Malaysia (15), and South Africa (13). The low-income countries with the most publications were Pakistan (6), Ghana (3), Nigeria (3), and Colombia (3).

About 23% of LMIC institutions published findings on HIS ethics more than once. Of the 21 institutions in middle-income countries that published more than once, most were in Malaysia (International Islamic University: 4 publications; Universiti Teknologi: 4; Universiti Kebangsaan: 3; Asia Pacific University of Technology and Innovation: 2; and University of Malaya: 2). The middle-income country with the next largest number of institutions with multiple publications was South Africa (University of Pretoria: 3 publications; Cape Peninsula University of Technology: 3; University of Cape Town: 2; and University of The Western Cape: 2). Two institutions in low-income countries published more than once: United Nations Population Fund in Pakistan (2 publications) and the University of Ghana (2).

The organization collaboration network (Figure 3) revealed that, among more than 700 identified organizations, only 218 of them collaborated (co-authorship) while many (72%) worked alone. The institutions publishing the most on HIS ethics were in HICs and collaborated with other HICs (Figure 3).

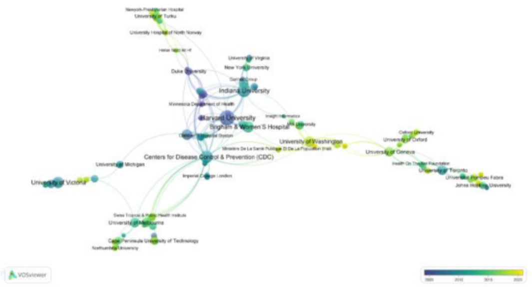

Figure 3. Closely Collaborated Institutions

Out of 601 publications, 30 (5%) were co-authored by institutions from both LMICs and HICs (Figure 4). The majority of these LMIC-HIC collaborated articles (23) were produced from institutions in one LMIC and one or two institutions in an HIC. For example, the University of Sindh (Pakistan) collaborated with Brunel University (UK) [14]. In three instances authors from one or two institutions in an LMIC collaborated with authors from institutions in multiple HICs. For example, Peking University Medical Informatics Center (China) collaborated with University of Washington (US), Fred Hutchinson Cancer Research Center (US), and University of Edinburgh (UK) [15]. In three cases, institutions from multiple different LMICs collaborated with one HIC institution [16–18]. One publication involved multiple LMICs and HICs, in which co-authors were from African Field Epidemiology Network (AFENET) (Kenya), University of Witwatersrand (South Africa), United Nations in Ethiopia, US Centers for Disease Control and Prevention, World Health Organization, and Umeå University in Sweden [19].

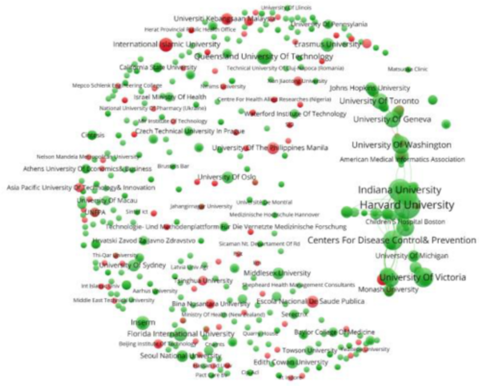

Figure 4. Collaboration Network Between Institutions in LMICs (red) and HICs (green)

Research landscape

The most common key terms (occurring more than five times) pertained to electronic health record(s) (similarly, electronic medical records and medical records) and health information system(s) (along with health information, health informatics, hospital information systems, and health information technology). The ethical concerns addressed by virtually all the publications were data security (with data protection and cyber security), privacy (with data privacy), or confidentiality. The only ethics keyword not tied to these concepts was human rights, used in only one publication.

Discussion

Occurrences and patterns

Overall, the number of publications explicitly addressing ethics in HISs was relatively small, which resonates with findings from two previous studies. Phillips and colleagues constructed a bibliometric map using index keywords from 17,655 publications (1990-2014) retrieved from Scopus about the health systems in sub-Saharan Africa [20]. Heesmink et al. took a bibliometric mapping approach to an overview of the field of computer and information ethics (C&IE) by visualizing the key terms co-occurred in 1,027 publications (2003-2009) using VOSviewer [21]. First, neither of these two studies conducted systematic literature search for their topics. Phillips only searched Scopus while Heesmink's study searched six journals and two conference proceedings. Second, neither studies screened search results, however, which likely would have reduced their included publications. Even so, in Phillips' study, none of their identified high-density keywords, visualized clusters of themes, or summarized topical foci was related to HIS ethics or similar concepts; Heesmink et al. referred to their included C&IE studies as “relatively limited” and cautioned against over-interpretation. Compaired to them, our study covered a much longer period (1975-2019) and exaustively searched more citation databases. Therefore, our longer review period and expansive source of data further underscores the paucity of research on HIS ethics, which may be due in part to a narrower focus (e.g., not just health or information systems, but health information systems). It is also likely that ethics in HISs is a neglected topic.

Although our results showed a general growing trend of relevant studies, we can’t predict whether the recent spike in publications related to HIS ethics will continue. Presently we have only one data point that stands in contrast to the trend of the last few decades. Moreover, the COVID-19 pandemic may affect research productivity in 2020. Generally, the global trend in the number of scholarly publications in the health system research [20] or the biomedical and health sciences field has been exponential growth [22].

Gaps

Arguably more important than increasing the number of HIS ethics publications is the need for a greater diversity of topics addressed. We found few people addressed ethical concerns beyond security, confidentiality, and privacy. Several other ethical principles pertaining to information systems did not appear in our included studies at all. For example, in their bioethical framework for health systems, Krubner and Hyder [23] included accountability to the public who are affected by decisions based on health system data. Furthermore, the extracted key terms from the title and abstract of included studies made no mention of open standards, open-source programming, or open innovation that are part of the Principles for Digital Development [3].

Heesmink et al. [21] identified three clusters of key terms for their included C&IE literature, which they labeled “privacy,” “ethics,” and “internet.” Their ethics cluster consisted of subclusters related to types of moral reasoning, ethics education, professionalism, and information technologies such as robotics. The topic gaps that we noted above were also gaps disclosed in their study.

In addition to the gap of research topics, we observed another gap in research collaborations. The personnel in Kenya who were charged with creating and maintaining HISs explicitly expressed a desire to understand the ethical issues and address them, but they were lack of resources to do so [24]. This interest and lack of resources could both be addressed through collaborations with better resourced countries. However, collaboration between LMIC institutions and those of HICs was rare (Figure 4). We found some LMICs (i.e., Nigeria, Philippines, Egypt, Morocco, and Ukraine) did not collaborate with any researcher in HICs in spite of their shared interests in HIS ethics. Even researchers in India who contributed to 27 publications in this study only collaborated with two HICs. Through collaboration, the LMIC institutions can explore with others the ethical concerns in their digital environments and learn about the concerns in more advanced systems that may soon be implemented in their country. Moreover, researchers from HICs can become more informed of the challenges and opportunities presented in LMICs. Collaboration between HIC and LMIC institutions is perhaps the fastest route to mutual learning. For example, regarding the principle of sustainability, almost by definition, low resource countries will find it challenging to upgrade their hardware every few years and keep their personnel up to date on new software releases. If they are to avoid a dependency on donors to finance the maintenance and replacement of their systems, they will need to be proactive in a dialog with technical experts from the donor countries that are usually HICs to express their needs and identify mitigating strategies.

This study focused on ethical considerations for HISs reflected in the peer-reviewed publications. In the future, a systematic search of the grey literature not restricted to full sets of principles would likely reveal additional documents related to HIS ethics.

Conclusions

We found that the literature on HIS ethics is sparse and, with the exception of the most recent year, did not grow at a particularly fast speed. This stands in contrast to the speed at which HISs are being developed and implemented, not just in LMICs, but worldwide. The technical advances are outpacing progress in the recognition of ethical concerns, much less the implementation of policies and procedures to ensure the ethical collection and use of data for populations. Our knowledge and practices regarding ethics in HISs are nascent and slowly growing globally, but the implementation of digital HISs in LMICs and the small number of authors from LMICs addressing the ethical issues is a gap that can and should be closed.

Acknowledgements

This study was supported by USAID Cooperative Agreement GHA-A-00-08-00003-00. USAID played no role in the design of the study; collection, analysis, and interpretation of data; or in writing the manuscript.

References

[1] MEASURE Evaluation, Defining Health Information Systems, https://www.measureevaluation.org/his-strengthening-resource-center/his-definitions/defining-health-information-systems

[2] S. Empel, Way Forward: AHIMA Develops Information Governance Principles to Lead Healthcare Toward Better Data Management, *Journal of AHIMA* **85** (2014), 30–32.

[3] A. Waugaman, *From Principle to Practice: Implementing the Principles for Digital Development*, The Principles for Digital Development Working Group, Washington, DC, 2016.

[4] J.C. Thomas, A. Reis, and V. Fleming, Principled Health Information Systems: Ethics Beyond Data Security, (2019), 1–22, https://www.measureevaluation.org/resources/publications/tr-19-360

[5] USAID, USAID’s Vision for Health Systems Strengthening,http://www.usaid.gov/sites/default/files/documents/1864/HSS-Vision.pdf

[6] J. Berens, U. Mans, and S. Verhulst, Mapping and comparing responsible data approaches, *SSRN Electronic Journal* (2016). doi: 10.2139/ssrn.3141453

[7] S. Montoya, Tracking Literacy in an Increasingly Digital World, http://uis.unesco.org/en/blog/tracking-literacy-increasingly-digital-world-0

[8] A. Varghese, M. Cawley, and T. Hong, Supervised clustering for automated document classification and prioritization: a case study using toxicological abstracts, *Environment Systems and Decisions* (2017).

[9] DoCTER, https://www.icf.com/technology/docter

[10] Covidence, https://www.covidence.org/

[11] M.E. Falagas, E.I. Pitsouni, G.A. Malietzis, and G. Pappas, Comparison of PubMed, Scopus, Web of Science, and Google Scholar: strengths and weaknesses, *The FASEB Journal* **22** (2008), 338–342.

[12] VOSviewer, https://www.vosviewer.com/

[13] M. Aria, and C. Cuccurullo, bibliometrix : An R-tool for comprehensive science mapping analysis, *Journal of Informetrics* **11** (2017), 959–975.

[14] K. Khoumbati, and M. Themistocleous, Evaluating Integration Approaches Adopted by Healthcare Organizations, *Journal of Computer Information Systems* **47** (2016).

[15] T.H. Payne, C. Lovis, C. Gutteridge, C. Pagliari, S. Natarajan, C. Yong, and L.-P. Zhao, Status of health information exchange: a comparison of six countries., *Journal of global health* **9** (2019), 0204279.

[16] A. Meri, M.K. Hasan, M. Danaee, M. Jaber, M. Jarrar, N. Safei, M. Dauwed, S.K. Abd, and M. Al-bsheish, Modelling the utilization of cloud health information systems in the Iraqi public healthcare sector, *Telematics and Informatics* **36** (2019), 132–146.

[17] N. Daniels, W. Flores, S. Pannarunothai, P.N. Ndumbe, J.H. Bryant, T.J. Ngulube, and Y. Wang, An evidence-based approach to benchmarking the fairness of health-sector reform in developing countries., *Bulletin of the World Health Organization* **83** (2005), 534–540.

[18] T. Simbini, R. Foster, P. Nesara, and C. Hullin Lucay Cossio, Monitoring diseases across borders: African regional integrative information systems, *Studies in Health Technology and Informatics* (2010).

[19] D. de Savigny, I. Riley, D. Chandramohan, F. Odhiambo, E. Nichols, S. Notzon, C. AbouZahr, R. Mitra, D. Cobos Muñoz, S. Firth, N. Maire, O. Sankoh, G. Bronson, P. Setel, P. Byass, R. Jakob, T. Boerma, and A.D. Lopez, Integrating community-based verbal autopsy into civil registration and vital statistics (CRVS): system-level considerations., *Global health action* **10** (2017), 1272882.

[20] J.F. Phillips, M. Sheff, and C.B. Boyer, The astronomy of africa's health systems literature during the MDG era: where are the systems clusters?, *Global health, science and practice* **3** (2015), 482–502.

[21] R. Heersmink, J. van den Hoven, N.J. van Eck, and J. van den Berg, Bibliometric mapping of computer and information ethics, *Ethics and information technology* **13** (2011), 241–249.

[22] J. Oh, H. Chang, J.A. Kim, M. Choi, Z. Park, Y. Cho, and E.-G. Lee, Citation analysis for biomedical and health sciences journals published in korea., *Healthcare informatics research* **23** (2017), 218–225.

[23] C.B. Krubiner, and A.A. Hyder, A bioethical framework for health systems activity: a conceptual exploration applying 'systems thinking, *Health Systems* **3** (2014), 124–135.

[24] S. Wambugu, J.C. Thomas, D. Johnson, and C. Villella, Digital data ethics in low- and middle-income countries: The road ahead, https://www.measureevaluation.org/resources/publications/tr-17-149

Address for correspondence

Fei Yu, PhD, MPS
335 S. Columbia Street CB 7585
Chapel hill, NC 27599-7585
feifei@unc.edu, +1 919-966-7637