Public Health and Epidemiology Informatics: Recent Research and Trends in the United States

B. E. Dixon1,2, H. Kharrazi4,5, H. P. Lehmann6

1 Richard M. Fairbanks School of Public Health, Indiana University—Purdue University Indianapolis, Indianapolis, IN, USA
2 Center for Biomedical Informatics, Regenstrief Institute, Indianapolis, IN, USA
3 Center for Health Information and Communication, Department of Veterans Affairs, Veterans Health Administration, Health Services Research and Development Service CIN 13-416, Richard L. Roudebush VA Medical Center, Indianapolis, IN, USA
4 Johns Hopkins University, Baltimore MD
5 Center for Population Health IT, Johns Hopkins School of Public Health, Baltimore MD

Summary

Objectives: To survey advances in public health and epidemiology informatics over the past three years.

Methods: We conducted a review of English-language research works conducted in the domain of public health informatics (PHI), and published in MEDLINE between January 2012 and December 2014, where information and communication technology (ICT) was a primary subject, or a main component of the study methodology. Selected articles were synthesized using a thematic analysis using the Essential Services of Public Health as a typology.

Results: Based on themes that emerged, we organized the advances into a model where applications that support the Essential Services are, in turn, supported by a socio-technical infrastructure that relies on government policies and ethical principles. That infrastructure, in turn, depends upon education and training of the public health workforce, development that creates novel or adapts existing infrastructure, and research that evaluates the success of the infrastructure. Finally, the persistence and growth of infrastructure depends on financial sustainability.

Conclusions: Public health informatics is a field that is growing in breadth, depth, and complexity. Several Essential Services have benefited from informatics, notably, “Monitor Health,” “Diagnose & Investigate,” and “Evaluate.” Yet many Essential Services still have not yet benefited from advances such as maturing electronic health record systems, interoperability amongst health information systems, analytics for population health management, use of social media among consumers, and educational certification in clinical informatics. There is much work to be done to further advance the science of PHI as well as its impact on public health practice.

Keywords

Epidemiology, disease outbreaks, public health practice, Public Health Informatics, health services needs and demand

Introduction

Public health informatics (PHI) was defined at the beginning of the century, as “the systematic application of information and computer science and technology to public health practice, research, and learning.” [1] and an agenda to advance the field was proposed by a group meeting at the American Medical Informatics Association (AMIA) 2001 Spring Congress [2]. The definition was refined in 2014 [3]: The scope of public health informatics includes the conceptualization, design, development, deployment, refinement, maintenance, and evaluation of communication, surveillance, information, and learning systems relevant to public health. Subsequently, this definition and the agenda [4] have been updated to reflect four key drivers: 1) a global reduction in infectious disease rates in combination with a growth in non-communicable chronic diseases; 2) the broader implementation of health information systems across the world; 3) continued rising costs of health care in nearly every nation; and 4) a greater emphasis on patient involvement in care decision-making processes. The result is a complex scope for public health informatics that intersects with similarly expanding fields of Global Health Informatics (public health and health care in low-resource settings) and Population Health Informatics (healthcare and social services to clinical populations), as depicted in Figure 1.

During the first decade of the twenty-first century, PHI efforts around the world were characterized by a focus on the core public health functions of monitoring populations: early detection of bioterrorism [5], such as Anthrax attacks in the U.S. [6] or the Tokyo subway attacks [7], as well as global health threats such as SARS [8], and the H1N1 pandemic [9]. While the threat of a large-scale epidemic has not diminished in recent years, as evidenced in 2014 by MERS [10, 11] and Ebola [12] changes in national policies and funding priorities have steered public health informatics in new directions.

Today public health informatics not only addresses public health functions in low-resource settings, but also contributes to national implementations of health information systems such as electronic health records (Delivery of healthcare in low-resource settings), the measurement of population health indicators within and across jurisdictions (Health planning), and conversations about the role of the citizen/patient in self-monitoring and self-management of both disease and well-being (health) (Care of clinical populations). Data science [13] comprising predictive analytics (chosen as the special topic of the 2014 edition of the Yearbook [14]), and statistical and epidemiological methodology, is incorporated into the practice and delivery of health services to better manage populations with complex and chronic illnesses and to provide a methodology common to all 3 domains of PHI, Global Health Informatics and Population Health Informatics [15].

In this article, we discuss the evolving scope of PHI through recent trends observed principally in the United States, and then we highlight advancements in the field evidenced in recent literature.
Methods

Our work is based on efforts by the American Medical Informatics Association (AMIA) Public Health Informatics Working Group to summarize recent literature for presentation at an annual “Year in Review” [16]. For that review, we performed multiple searches of the biomedical informatics literature published in MEDLINE between January 1, 2012 and December 31, 2014. Supplemental articles were gathered by reviewing the bibliographies of selected articles and by soliciting suggestions from PHI working group members using AMIA listservs. We used the following MEDLINE Subject Headings (MeSH) keywords in various combinations to maximize sensitivity: “public health informatics”, “public health”, “informatics”, and “information systems”. Use of MeSH headings was preferred, although articles were also identified using keywords to check article titles and abstracts. The search strategy was similar to that used by the IMIA Yearbook in Informatics to select candidate “best papers” [17].

We used a variety of inclusion and exclusion criteria to narrow the list of candidate articles. To be included, an article needed to (a) focus on a topic related to public health practice or research; and (b) include informatics, information science, or computing as the primary subject of the study or a main component of the study methodology. We used a broad lens to assess each article’s relevance to public health practice and research, including: 1) activities conducted by, with, or involving a local, state, or federal health agency; 2) assessment and monitoring of diseases and health outcomes; 3) primary and secondary prevention of diseases; 4) social determinants of health as well as health disparities; and 5) development of the public health workforce including PHI education and competencies. Articles were excluded if they did not constitute original research such as letters, editorials, perspectives, opinions, whitepapers, comments, and study protocols. We further excluded articles outside the United States to limit the scope.

Selected articles were synthesized using a thematic analysis to create a broad survey of advances in the field of PHI. Articles were grouped into themes aligned with a framework based on the Essential Services of public health, a hierarchical typology defined by the National Academy of Sciences, whose top level comprises Assessment, Policy Development, and Assurance [18].

Results

From a policy perspective, the disease focus of public health is changing. Multiple epidemiological studies provide clear evidence that modern health care systems, even beyond the developed world, must less focus their attention on infectious diseases and more on chronic (non-communicable) diseases [19]. For example, worldwide tuberculosis incidence has fallen at the average rate of 1.3% per year over the past twenty years [20], while the incidence of diabetes mellitus continue to rise and is projected to be as high as 33% by 2050 [21]. This change requires a broadening of public health information systems to focus on the routine, systematic collection of a wider array of data on population health indicators, including social determinants of health [22, 23, 24]. Many of these data have been collected in the past using health surveys, and managed as distinct silos of information distributed by specific health agencies. Newer health policies, such as the Patient Protection and Affordable Care Act in the U.S., require the integration of chronic disease data with clinically-derived indicators to paint a broader picture of health status for a given community. In the United States, these efforts are sometimes referred to as population health informatics [25], despite a semantic overlap in purpose and activities with public health indicators that were long performed by ministries of health in other countries. Such efforts further seek to incorporate “data science” [13] as well as predictive analytics (discussed in last year’s Yearbook [14]) into the practice and delivery of health services to better manage populations with complex and chronic illnesses [15].

Policy now also addresses health information technology. In the United States, the 2010 HITECH Act introduced “meaningful use” (MU) incentives to encourage hospitals and general practitioners to adopt EHR systems [26]. In the past four years, EHR adoption rates among non-federal acute care hospitals increased from 15% to over 90% [27]. We shall see the several roles these incentives play in American PHI.

These public health and clinical policies have led to an investment in health informatics by public health authorities, which rests on the promise that information and com-
Communication technology (ICT) would support health systems in creating efficiencies, improving care quality, and lowering health care spending [28]. Globally, health-related costs are rising [29] due in large part to a globally aging population combined with the rapid growth of chronic disease burden. Ministries of health seek to implement integrated health informatics programs as part of a broader strategy to better coordinate care, reduce waste and inefficiencies while supporting better management of costly diseases [30].

In the US, these population-health goals are represented by the policy push to implement patient-centered care models, which aim to achieve better, coordinated care, while reducing costs. Patient-centered or patient-focused care models seek to apply patient preferences, self-management, and self-reported outcomes into clinical decision-making. These efforts attempt to build patient-provider trust and communication. The establishment of the Patient-Centered Outcome Research Institute as part of the Affordable Care Act provides a novel funding stream and aims to build that trust from the ground up through patient engagement [31].

Another patient-centered effort concerns social services. The UK and other countries are further implementing “social care informatics” [32] where the human-services community is working towards integration and interoperability of social services into health systems. For example, the Administration for Children and Family (ACF), a division of the U.S. Health and Human Services agency, funded 7 state-based interoperability projects in 2013 to provide models for, and demonstrate the value of, interoperability beyond Medicaid eligibility, and include case management [33]. A major goal is providing “no wrong door” services [34], where services are focused on the client’s and family’s needs, rather than on the bureaucracy’s, which the family may have contacted in no particular order.

**The Public Health Informatics Stack**

Despite a rapidly evolving context, the Essential Services of Public Health [18, 35] — Assessment (of public health situations and threats), Policy Development (to address what was assessed), and Assurance (to implement the policies) — remain the core activities performed by public health authorities. In order to depict and systematize the analysis of the current international trends in PHI, we reinterpreted high-level concerns outlined in the updated agenda for PHI[4] into a “stack” of layers (like the OSI stack for networking [36]) that provide a supporting infrastructure for the Essential Services (see Figure 2).

From the informatics perspective, public health services are supported by a socio-technical infrastructure [37] “delivered” to public health, relying on governance policies and ethical principles of local public health systems. In turn, that infrastructure depends on education and training of the public health workforce, on research and development that results in a novel infrastructure, adapts the existing infrastructure to the public-health domain, or evaluates the success of the infrastructure. Finally, both the persistence and growth of the infrastructure depend on financial sustainability.

In the remainder of the paper, we highlight recent literature that demonstrates the evolving PHI stack (see Figure 2) as influenced by the public health context. We start at the top of the stack.

**Public Health Essential Services**

In the PHI “Year In Review” presented at the 2014 AMIA, the US Essential Services covered by the literature were presented:
Notifiable Disease Surveillance (clinician lab results or diagnoses reported to state agencies), Syndromic Surveillance (clinical data reported to state agencies on an ongoing basis), Population Health (broadly conceived), Immunization Information Systems (reporting of vaccinations; access to clinicians; immunization forecasting), Environmental systems (non-clinical sources), and Other [38]. Each of the first four topics has a relationship to the MU regulations mentioned earlier. For Stage 1, hospitals were obliged to demonstrate their capability to submit reportable lab results and data for syndromic surveillance to a state agency [39]. In MU Stage 2, submission of immunization information is now a core requirement, and two further mhealth menu options have been added: reporting cases to cancer registries and reporting cases to other registries [40].

(1) Socio-Technical Infrastructure & Research

The notion of “technical” is broad, since informatics is a “socio-technical” activity [41], whether in hospitals in the US [37] or pediatric care in the UK [32] or mobile health in India [42]. PHI involves a complex social-technical infrastructure composed of technical systems, people, and of organizational systems in which both people and technology are deployed to achieve an aim [43, 44].

Merrill and colleagues performed a socio-technical infrastructure evaluation of a health information exchange (HIE) and its system dynamics, mixing both technical concerns (“RHIO” aggregation of electronic health information in standardized form”), and social issues (“pressure to increase resources”) [45]. They found a number of “loops” to capture key dynamics in resolving the gap between the desired level of health information exchange and the clinical information electronically available: the Sliding Goals Loop, the Project Rework Loop, and the Maturity of Resources Causal Loop. The consequence is that the states planning to implement HIE should be mindful of these dynamics proactively, or at least in the event of a persistent failure.

The following sections provide specifics of recent PHI literature that document advancements in both public health practice and methods as well as the broader use of interoperable HIE, mHealth, and social media technologies to support the public health workforce and its mission.

(1-a) Evaluation of Surveillance Practice

Surveillance is a key aspect of monitoring population health, an Essential Service under Assessment. A decade ago, Buehler et al. challenged the PHI community not just to build better surveillance systems, but to evaluate them rigorously as they are developed and deployed in service for health agencies [46]. Recent published evaluations demonstrate some progress both in terms of surveillance system performance as well as methods used to conduct surveillance. In the U.S., surveillance systems received a boost from MU incentives for the adoption and use of EHR systems. In the latest requirements of the MU program, hospitals and physicians are required to submit data to state public health agencies in support of syndromic and infectious disease surveillance as well as immunization registries [40]. Recent literature in PHI from the U.S. has largely focused on measuring the adoption of these technologies [47, 48] and evaluating the impact of these technologies on the Essential Services [38].

Overall adoption of surveillance systems, especially electronic laboratory reporting (ELR) and syndromic surveillance (SyS), continues to grow. A data brief from Wu and colleagues highlights that nearly two-thirds of U.S. states report adoption of ELR due in part to MU regulations [49]. A systematic review by Gluskin and colleagues found that the implementation of ELR systems has improved the timeliness and workflow associated with infectious disease surveillance in U.S. public health agencies [50].

These process improvements are noteworthy, and they confirm previous studies that had examined the timeliness of ELR in the US [51], among other places. Yet faster and more complete reporting should lead to improved population health outcomes. Moreover, recent literature provides insights into the challenges that affect greater adoption, use, and impact of surveillance systems on population health. Studies by Vest and colleagues [52], as well as Dixon and colleagues [38] highlight that despite a more timely reporting, data quality in surveillance systems is often poor and presents a barrier to more efficient PHI practice. For example, Dixon and colleagues [38] show that ELR systems open a floodgate of data (including duplicate reports and false positives) for health agencies but limited budgets for human resources and to update technologies within health agencies constitute a barrier to the timely cast investigation and follow-up after the detection of an outbreak or the routine case of infectious disease.

(1-b) Surveillance Methods

While many existing SyS systems utilize free text chief complaints [53], there is increasing interest in PHI surveillance using more complex text sources. For example, Travers and colleagues developed a natural language text processing method to pre-process and classified the text, using both heuristic statistical methods and parameters tuned to emergency contexts [54].

Text processing should be an important enabler of mass customization of health messaging across cultures within a jurisdiction. Turner and colleagues [55] evaluated machine translation (MT) from input through post editing, using Google Translate, as in their pilot study [56]. They found human and machine translation to be equally preferred (and accurate), but MT took only 18% as much time.

Even less “traditional,” for the Essential Service Diagnose & Investigate, Sinčak and colleagues used a “neuro-fuzzy” approach (combination of neural network and fuzzy/approximate reasoning) to simulate water systems and assess Legionella risk in specific buildings [57]. Another study used wearable monitors, usually part of the “quantified self” movement, for infectious disease modeling [58].

(1-c) Interoperable HIE Infrastructure

Recent literature in PHI further highlights both advancements and barriers to interoperable health information exchange (HIE). Public health authorities often need multiple data sources to comprehensively monitor population health and examine impact of policies and interventions on health out-
comes. Historically, epidemiologists have spent a significant amount of their time and energy in the gathering, integration, and analysis of multiple datasets. Given the adoption and utilization of EHRs, as well as other informatics systems, public health authorities increasingly had access to a wider array of electronic datasets. This greater access has ushered in a “Neolithic Revolution” in PHI [59], yet few electronic datasets within a health authority or external ICT systems are easily linked. Therefore, interoperable HIE data can be an important asset for health authorities to support the essential functions of public health including community assessment.

The U.S. has witnessed a growth in interoperable HIE activities [60, 61] as well as the number of public data sets [62]. However, most public health authorities are not involved in community HIE efforts, and few possess full access to the data they need to monitor population health [63, 64]. Klann and colleagues highlighted a successful model by which a local health authority might query a federated network of health clinics to capture population health indicators [65]. Yet many health departments lack sufficient informatics capacity [66] to establish or leverage robust information infrastructures [52]. While some larger, U.S. state health departments have been able to establish robust communication networks for interconnecting health departments and supporting summary dashboards to increase the use of such networks [67], few public health authorities actually have the capacity both to receive and to send information to health care delivery systems [68, 69].

(1-d) mHealth and Social Media Infrastructure
Mobile technologies play several roles in the Essential Services. As part of Assessment, they can replace paper-based surveys that have been the stalwarts of Assessment (situational awareness). In their ability to communicate out, from the public health “center,” they can Inform, Educate & Empower, a component ascribed to Policy Development. There has been increased use of mobile technologies in all areas of PHI. Househ performed an umbrella systematic review—that is, a review of reviews—of short messaging service (SMS), the prevalent form of mobile communication in reduced-resource settings [70]. The 13 reviews, representing 299 studies, covered a range of functions and a range of contexts (both developed and developing). SMS shows promise in preventive care [71, 72] in promoting healthy behaviors [73, 74], and in promoting medication adherence among people with HIV [75]. Results were unclear for communicating lab results [76] and even less clear for “health” outcomes [77].

Social media have received attention, most famously as providing a signal for disease surveillance [78]. Denecke and colleagues reviewed the state of the field (as of 2013) and offered a proof (the M-Eco system) of the concept that multiple sources of social data can be brought to bear on the question of infectious-disease outbreak, more focused on reportable diseases [79]. The system is user-centered, in filling the pipeline from (German text) data source to dashboard. They found that only 21% (rather than about 75%) of texts were making it all the way through. They realized that in “social media, medical terms are not necessarily used to talk about medical conditions.” Synonymy, only based on medical terminologies, is not adequate.

Capurro and colleagues provide a more complete, but from 2012 data, picture of the uses of social networking sites for public health practice and research [80]. Their review comprises 73 primary papers. There were 15 articles that targeted hard-to-reach populations, such as youths and adolescents and people at risk for Sexually Transmitted Infections and HIV. Other Essential Functions addressed were general health promotion strategies, substance (not just drugs) abuse, and mental health. The review also pointed out the fickleness of some social networks: who remembers MySpace?

(2) Education and Workforce Development
Education and Workforce Development are components of Assurance in the Essential Services typology. In addition, recent activities, primarily in the U.S., highlighted major advances in the education of the PHI workforce. The U.S. Office of the National Coordinator for Health IT provided grants in 2011 to undergraduate level institutions to provide training in PHI. In 2014, the National Association of Schools in Public Health recognized informatics as a core competency for Masters of Public Health (MPH) degrees in the 21st century, paving the way for MPH programs to revise their curriculum and include informatics competencies as well as courses. The U.S. Centers for Disease Control and Prevention (CDC) Informatics Fellowship has been recognized as a Registered Apprenticeship program [81]. Combined, these different activities create opportunities and a pipeline of PHI workers to serve the needs of local, state, and federal agencies.

Despite these advancements, PHI remains undersubscribed when compared with other informatics branches, including clinical informatics and broader health informatics. The U.S. recognized clinical informatics as an official medical sub-specialty with a Board certification exam available to medical doctors [82], sparking a growth in medical schools interested in teaching and hosting biomedical informatics programs. Furthermore, the recent agreement among AMIA, the American Health Information Management Association, and the Commission on Accreditation for Health Informatics and Information Management Education to align health informatics competencies and program accreditation has sparked growth in health informatics programs and courses. Many students in these programs seek broader degrees in health or biomedical informatics, creating uncertainty about the sustainability of PHI programs and tracks. Although individual certification is available in clinical informatics, there is no equivalent exam in PHI.

(3) Governance and Ethics
Samuel and Zaïane [83] laid out a framework for health informatics ethics, based on a review of nine organizations’ codes of ethics. Earlier, in 2002, a consensus by the public health community expressed a 12-point code of ethics [84]. Only one informatics paper relates informatics issues to this code, but it is in the realm of bioinformatics, rather than the primary functions we consider in PHI [85]. IMIA has made a more official effort in 2013 [86].
(4) Sustainability
A perennial concern in informatics is whether funding might come to sustain PHI efforts in practice and research. While little new funding has been made available for PHI research, the Institute of Medicine (IOM), the US equivalent to a National Academy of Medicine, has pushed forward the concept of a Learning Health System [87]. The essentials of the Learning Health System are familiar to anyone in public health: Use aggregate data to learn, and project that learning to the care of the individual. Whether the imprimatur of the IOM will result in increased funding for the basic research needed to accomplish this goal—which essentially is PHI—remains to be seen.

Conclusion

PHI is a field that is growing in breadth, depth, and complexity. In this essay, we examined the roles PHI contributes in each of the major Essential Services typology. Yet many specific, or component, Essential Services still have not benefited from the sea changes in the socio-technological contexts described above. While social media have been to for monitor the health of the population and to deliver healthcare, we have just seen the beginnings of the use of technologies for “Informing and Educating.”[55] (2 Beyond machine translation, there are entire areas of mass customization of health messaging that should take advantage of the same technologies that for-profit organizations currently use to identify which advertising should be displayed to a specific user.

The full power of modern simulations and analysis is rarely made available to departments or ministries of health to “Develop Policies.” The ability to monitor wearable technologies, or to monitor Web-page access, provides a window to possibilities to “Enforce Laws,” although this needs to be limited, in keeping with Ethical principles that still need to be defined.

Finally, the degree to which ICT can support non-informatics public health training and life-long learning—and to what extent PHI should be an expected competency in both contexts—remains to be worked out.

PHI, then, in its extended form (see Figure 1) has as many Grand Challenges as clinical informatics. We hope future IMIA reviews document continued, sustained progress in both research and practice.

Acknowledgements

We wish to thank Louis Imperiale and Fardad Gharahbel for their help in identifying and reviewing recent public health informatics articles. We further thank Anne Turner, Jamie Pina, and Janise Richards for their help and support with the AMIA Year in Review sessions in both 2013 and 2014, which provided the foundation for these comments.

References

19. Chioloaro A, Paccaud F, Fornerod L. How to conduct public health surveillance? The exam-


87. Strengthening the evidence base for a learning health system - Inspirations from good practice for capacity building in health services research and public health research [Internet]. 2011. Available from: http://eprints.hta.ac.uk/908/1/HTA-Projektericht_NR48.pdf Item Resolution URL.

Correspondence to: Harold Lehmann 2024 E Emmanuel St Baltimore MD 21209 Tel: +1 410 502 7569 E-mail: lehmann@jhmi.edu