Present and Future Trends in Consumer Health Informatics and Patient-Generated Health Data

A. M. Lai1, P.-Y. S. Hsueh2, Y. K. Choi3, R. R. Austin4

1 Institute for Informatics, Washington University in St. Louis, USA
2 Computational Health Behavior and Decision Science, Center for Computational Health, IBM T.J. Watson Research Center, USA
3 Department of Biomedical Informatics and Medical Education, University of Washington, USA
4 School of Nursing, University of Minnesota, USA

Summary

Objectives: Consumer Health Informatics (CHI) and the use of Patient-Generated Health Data (PGHD) are rapidly growing focus areas in healthcare. The objective of this paper is to briefly review the literature that has been published over the past few years and to provide a sense of where the field is going.

Methods: We searched PubMed and the ACM Digital Library for articles published between 2014 and 2016 on the topics of CHI and PGHD. The results of the search were screened for relevance and categorized into a set of common themes. We discuss the major topics covered in these articles.

Results: We retrieved 65 articles from our PubMed query and 32 articles from our ACM Digital Library query. After a review of titles, we were left with 47 articles to conduct our full article survey of the activities in CHI and PGHD. We have summarized these articles and placed them into major categories of activity. Within the domain of consumer health informatics, articles focused on mobile health and patient-generated health data comprise the majority of the articles published in recent years.

Conclusions: Current evidence indicates that technological advancements and the widespread availability of affordable consumer-grade devices are fueling research into using PGHD for better care. As we observe a growing number of (pilot) developments using various mobile health technologies to collect PGHD, major gaps still exist in how to use the data by both patients and providers. Further research is needed to understand the impact of PGHD on clinical outcomes.

Keywords
Consumer health information/methods; patient-generated health data; mHealth; user-computer interface; consumer participation in delivery of health care

Introduction

In recent years, the widespread adoption of personal computing technology, availability of personal health records (PHRs), and utilization of various forms of social media has invigorated an interest in consumer health informatics (CHI) and created an explosion of interest in the potential of patient-generated health data (PGHD). Internet adoption by adults in advanced economies reached a median of 87% and, similarly, 68% of adults reported owning a smartphone in 2015 [1].

Another factor pushing forward the potential of PGHD is not only the widespread availability of accelerometers embedded in smartphones and wearable devices that can collect physical activity [2], but also affordable biometric sensors that can collect and transmit weight, blood pressure, heart rate, temperature, and even blood glucose information from patients to their healthcare providers [3]. This recent ability for consumers to increase their participation in their own healthcare by recording and sharing health data through the use of affordable sensors and PHRs has inextricably tied together the topics of CHI and PGHD.

While there have been previous reviews on the topic of CHI, the objective of this survey paper is to generate a sense of the recent, past, and current foci in CHI and PGHD using a semi-structured scoping review process.

A recent general review of the overall trends of CHI is discussed by Demiris, in which a majority of trends involve the use of patient-generated health data [4]. In fact, in the past decade, quite a large amount of evidence has been accumulated with regards to the use of active and passive monitoring data from home telehealth and mobile health technologies, as well as from PHRs [5]. The role of PHRs has expanded in recent years, either as parts of electronic health records (EHRs) offered by providers, or through the enablement of patients themselves on some technology-based platforms for data collection and sharing, or for integration purposes.

In the United States, with the implementation of the federal government Meaningful Use incentive program, the proportion of consumers accessing their records increased from 27% in 2014 to 45% in 2016 [6]. This adoption has further enabled more frequent patient-provider communication, with 64% of United States physicians in 2015 having an EHR with the capability to exchange secure messages with patients, an over 50% increase since 2013 [7]. In addition, the collection of patient-reported outcomes (PROs) and their contexts have been further enabled by the advance of ecological momentary assessment tools, which focus on the collection of symptoms and behaviors close in time to their experience.

Under such a healthcare landscape shift, recent major research foci are the implementation and understanding of the value of providing clinical notes through PHRs and incentivizing patients to participate in patient-provider communication [8]. In addition, with the availability of patient-generated health data from the Quantified Self (QS) movement through a plethora of consumer devices, Demiris called for an infrastructure...
that supports clinicians to become engaged in the processes that QS facilitates and the technologies that can help make sense of data through massive data retrieval and trending, as well as modeling and prediction [4].

In a review of the definitions of CHI, Flaherty et al. have identified five quality assessment criteria to evaluate on the definitions in a set of 23 selected CHI papers [9]. Despite the fact that none of the definitions used in these papers fulfilled all five quality criteria identified, this review identified some success stories, including the use of Comprehensive Health Enhancement Support System (CHESS) for helping individuals deal with health crises or medical concerns, as well as the use of CHI devices for improving clinical and behavioral outcomes. The varying definitions have demonstrated the multidisciplinary nature of the field of CHI. This nature has also reflected on the varying efforts of PGHD evidence collection.

This article differentiates itself from previously written reviews on the topic of CHI and PGHD by providing an update and summary of the recently published scientific literature on these topics through the conduct of a scoping review. We also examine not only the biomedical literature, but also that of computer science and engineering.

**Methods**

A scoping review of the literature was conducted using a search strategy to generate an overview survey analysis of the last three years of the biomedical, computer science, and engineering literature focusing on CHI and PGHD. The authors selected PubMed and the ACM Digital Library to understand recent literature and up to date conference proceedings. In this paper, the rationale for the database selection is not intended to be a full all-inclusive systematic review but the aim is to propose a survey paper to highlight and provide an overview of current trends and the future trajectory of CHI and PGHD.

Our initial search strategy for PubMed was to use a variety of MeSH headings to identify the articles of interest, but an appropriate set of MeSH headings that yielded the desired articles could not be identified. The list of MeSH headings included Consumer Health Information/methods, User-Computer Interface, Health Services Needs and Demand, Physician-Patient Relations, Consumer Participation, Medical Informatics, and Delivery of Health Care. This resulted in the use of key search terms for any mentions of the terms “patient-generated health data” or “consumer health informatics” restricted to the papers published in 2014, 2015, and 2016, and written in the English language. As a result, the following query was used:

```
```

The search strategy for the ACM Digital library used key search terms: consumer, patient, self-quantifier, citizen, “patient-generated data,” and “consumer health,” restricted to the years 2014–2016, and limited to English language. The team iteratively refined the search terms in order to focus solely on consumer health informatics and patient-generated data and to reduce the number of articles that were not related to these topics and did not address the search criteria. The key search terms allowed the team to narrowly focus on a specified time frame and topic for the scoping review. This resulted in the following query:

```
acmdTitle:(consumer, patient, self-quantifier, citizen -mlearning, -learning, -education, -vision, -speech, -preliminary, -3d, -flash, -API, -search, -retrieval, -seeking, -gaming, -simulation, -robot, -image, -virtual, -invasive, -scratch, -middleware, -immune, -vision, -children, -eye, -VLSI) AND ("patient generated data", "consumer health")
```

After retrieving the references from these two queries, the authors screened the articles based on titles for relevance to CHI and PGHD. Articles that were selected by two or more of the authors were placed on the list for further review. The authors then categorized the articles on the list into themes. The articles were divided amongst the authors to retrieve, summarize, and perform a thematic analysis. A small subset of the articles (n=5) could not be retrieved by any of the authors due to limitations in publication availability at our combined institutions. Due to time constraints and because this is not an exhaustive literature review the authors decided to exclude this small subset of articles from the survey. Articles that were reviewed by the authors and determined to be general review or survey articles were also excluded from this analysis.

**Results**

The team executed the queries described above on November 4th, 2016. From the searches, a total of 97 articles (65 PubMed and 32 ACM Digital Library) were included for screening (Figure 1.). Three of the authors (AL, PH, and RA) screened the articles for relevance to CHI or PGHD. After the review and screening of titles, 47 articles were included in the survey of the activities in CHI and PGHD. The authors (AL, PH, RA, and YC) categorized the remaining 47 titles into three themes: 1) Patient-Generated Health Data; 2) Mobile Technology; 3) Human-Computer Interaction. The articles have been summarized below.

**Patient-Generated Health Data**

A large fraction of the papers on the topic of PGHD and its applications to various aspects of healthcare that we encountered in our literature search could be categorized as reviews or opinion papers [10-15]. In these articles, the definitions of PGHD that were used were slightly different, but many were based on a broad definition of PGHD, which included any observation, result, device finding, confirmation, change correction, or addition of data to a patient’s record that was created, recorded, gathered, or inferred by or from patients or their designees [10, 12]. In order to cast a broad net with regards to the activity in this emerging focus area, we have used this broadest definition.

The most prominent difference separating PGHD from data generated during clinical settings is that with PGHD, it’s the patient and not the provider that takes the ownership of generating, capturing, and sharing the data. The concept of PGHD is not new in clinical care as patients have always been...
recording and sharing information on their health in a variety of ways [12]. However, it is becoming increasingly common with the proliferation of more affordable consumer devices (e.g., smart phones, wearables, and sensors) and the advent of technologies such as patient portals and personal health records. We discuss the use of PGHD with mobile devices in the following section.

The recent literature points to the many potential benefits of integrating PGHD into health care delivery. Cohen et al. conducted semi-structured interviews to examine the experiences of health care professionals who use PGHD in an outpatient clinical setting [16]. The identified benefits included a deeper understanding of a patient’s conditions, the availability of more accurate patient information, and the ability to monitor patient’s health between clinic visits. PGHD through monitoring devices could be used to collect a more comprehensive view of a patient’s physiology [14].

A number of articles discussed different ways PGHD can be incorporated into clinical trials and discussed some of the issues [10, 3, 17]. Wood et al. also described how PGHD could be used to monitor medication adherence and to remind patients to take study drugs or follow study protocols [10]. The authors acknowledged the concern that PGHD streams have been primarily from consumer-grade devices rather than calibrated research-grade devices, and that further studies are needed to understand the data quality from these devices.

Furthermore, ubiquitous monitoring of a patient allowed through PGHD could not only advance our knowledge of a patient’s cancer experience but also enhance the efficiency and productivity of cancer clinical trials [3, 17]. In addition, PGHD empowers patients by engaging them as key players in their care [3, 18, 19]. Petersen discussed how PGHD could improve cancer survivorship through “support of survivor autonomy, improvement of survivor health, and promotion of survivor population health” [11].

Another area of interest for the research community is how to extract and process meaningful information from unstructured
PGHD from online health communities. Hartzler et al. generated health interest profiles by extracting health-related terms from users’ past posts from an online health community and conducted a user study to examine the validity of such profiles [20]. Their findings demonstrated that the PGHD-generated health profiles were matched with the users’ health interests. Furthermore, online health community members showed favorable views on using the profiles to find a match and connect with other members for peer support.

Additional challenges of widespread adoption of PGHD reported in the literature include: lack of randomized controlled trials (RCTs) to examine the associations between PGHD and longitudinal health outcomes [3, 17], PGHD data standardization and device interoperability [3, 17, 21], integration of PGHD with EHRs [3, 17], concerns for security and privacy [3, 17, 21], and building data analysis and visualization tools that facilitate informed health decision making [3, 21-23]. Future research is needed to address the barriers and gaps identified in the literature to harness the true potential of PGHD, bringing value to both patients and health care providers.

Mobile Technology

Mobile technology or mHealth is transforming the way patients, providers, and researchers interact with data. For the purposes of this review, mHealth includes mobile technologies such as smartphones, tablets, wearable devices, and sensor technologies. M-Health is fundamental to healthcare transformation since it shifts the ability to collect data anywhere, at any time, and it is becoming seamlessly integrated into our lives. This allows for a gradual shift of healthcare closer to the patients’ daily living and away from the traditional clinical environment [24].

One specific application of mHealth includes monitoring chronic disease and the potential for enhancing self-management of chronic disease. Chronic diseases are in the center of mHealth developments as they require the continuous and active involvement of not only healthcare professionals but also patients, both of whom can be empowered through the use of specialized mobile applications and the analysis of data from modern miniaturized and wearable sensing devices [24]. Important considerations such as system design of these technologies can increase long-term sustained use of consumers but also reduce the burden for healthcare providers.

An example of the use of mobile health technologies to help patients with chronic conditions to collect and monitor their physiological signs is WearCOPD, a mobile app developed by Liaqt et al. that uses a smartwatch and a smartphone to collect chronic obstructive pulmonary disease (COPD) related symptom data to detect acute exacerbations of COPD [25]. Patients complete a daily health questionnaire prompted by the app to monitor their status. The app also collects patients’ peak expiratory flow by having them blow into the microphone.

Another example is a mobile app that uses wearable sensors and biochips to facilitate self-management of Inflammatory Bowel Disease (IBD) [26]. Pernencar et al. argue that automatic data collection with the use of wearables and biochips could reduce the burden of self-report and should allow objective physiological data to be captured and analyzed for self-management of IBD [21]. The previous two studies [20, 21] were ongoing at the time of the publication and therefore relevant evaluation results could not be assessed.

Another study presented a privacy-preserving video surveillance system using a network of multiple IP cameras and motion sensors to monitor patients in their real-life settings. The system stores and sends a “ghost” form of an image, which only contains the information about the object’s movement in blue color on a black background to preserve the privacy of the monitored individual [27]. The proposed surveillance system demonstrated effective patient monitoring detection algorithm while preserving the privacy of monitored patients.

Kumar et al. described their approach to the integration of continuous glucose monitor data and the transmission of the data to a Personal Health Record [28]. The glucose monitor data was collected onto an Apple mobile device via Bluetooth, stored locally on the iOS device using the Apple HealthKit interface, and subsequently transmitted from HealthKit to an Epic EHR through the Epic MyChart app running on the mobile device.

Nundy et al. discussed the value of PGHD gathered using mobile technology for diabetes self-management [29]. They explored provider perceptions of a summary report based on the data extracted from a text message-based diabetes self-management program. The patients texted back responses to automated text messages consisting of reminders and educational messages. From these responses, summary reports were generated to inform providers of clinically relevant patient-reported health data. The authors used Likert-like response surveys and interviews with primary care physicians (PCPs) and endocrinologists to assess their responses to summary reports. Only three out of the 12 providers felt that the report changed the care they provided. However, nine of the 12 providers were willing to use the summary report.

Another growing area of mHealth is monitoring and tracking mental health conditions with mobile devices and sensing technology. Mobile technology makes it possible to extend mood self-assessment, from lab to real life, by the collection of mood data, frequently, over a long time, and in different life situations [30]. Improving data collection for conditions such as mental health increases the opportunity for early intervention and improves outcomes. The utilization of pervasive technology, including a mobile phone and its sensors, could potentially provide a way to make therapies more personalized and accessible at any time [31].

Telerehab is another area that is capitalizing on the expansion of mobile technologies. Rehabilitation sensing and tracking allow for pattern recognition through data analysis and visually represent the data needed to share the patient’s story. Conventional physical rehabilitation in stroke and other physical disorders is provided by sparse and home-based systems that suffer from issues of compliance, low patient engagement, and lack of personalization [32]. Another program, Back on Bike, a cycling rehab program, showed tremendous promise and is expanding the possibilities of traditional rehabilitation for cardiac patients [33]. Dynamic approaches to rehab used in the patient’s home for closer...
self-monitoring have the potential to increase adherence and to enhance secondary prevention of cardiac disease.

Despite the potential benefits of collecting PGHD using various mobile health technologies, challenges remain with how to analyze and interpret large amounts of high-dimensional PGHD to detect trends and to make meaningful inferences [21, 23]. To address this issue, researchers have developed a new data analysis methodology and a new framework to analyze PGHD. Liang et al. proposed a new method based on association rule mining techniques and successfully demonstrated its validity to discover correlations between lifestyle factors and sleep patterns [22]. Gollamudi et al. proposed a hypothesis testing the framework for analyzing unstructured time series data, a common data type of PGHD, to discover unique trends and associations in the data [23]. Using the framework, they have analyzed blood pressure data collected by patients at home and they demonstrated the efficacy of a comprehensive smartphone-based health monitoring intervention.

While some initial mHealth studies are being conducted, more large-scale studies are however needed to develop best practices and create a foundation of evidence to support the use of mHealth across care settings. Barriers and opportunities exist that incorporate new technology for both providers and consumers. Since the locus of healthcare is shifting to the home and community setting, there is an increasing need to adopt a broader approach across the traditional boundaries of health and social care in order to operationalize a more integrated and personalized healthcare service provision [34].

Emerging Research in Human-Computer Interaction

Historically, a cross-cutting focus in CHI has been human computer interaction (HCI) and this trend has continued. While many existing efforts are focused on the collection of evidence on existing online health tools, more researchers are looking for ways to improve patient-provider communication, and patient-led, shared experiences, using methods ranging from theory-informed models, to surveys, interviews, and personal constructs.

Jacobs et al. compared health information sharing practices between patients, providers, and healthcare navigators. They identified gaps between the groups with implications on design and adoption of technology [35]. Another study revealed limitations of existing approaches to support patient-provider communication and identified challenges for the design of systems that honor patient needs and preferences [36].

Chung et al. surveyed 211 patients, interviewed 18 patients, and re-analyzed interviews of 21 providers [37]. They found that there exist needs of support for collaboration in every stage of self-tracking and that patients and providers create boundary negotiating artifacts to support the collaboration. These are important findings to further support the design of PGHD tools based on stage-based models of personal informatics.

Briggs et al. have further supported that patients and caregivers frequently get their health information and advice from websites containing patient-led shared health experiences [38]. This means that they often engage in a very idiosyncratic selection process in order to determine which websites have personally-resonant material. In this paper, a Repertory Grid (repgrid) technique was applied on websites for patients with chronic asthma and caregivers of people with multiple sclerosis (MS), presenting each patient/caregiver with a set of health websites relevant to the condition for which they were looking for information. Hyper-personal representations of those constructs provided new insights on the ways individual patients can use the online health tool in their own contexts.

Finally, Maniam et al. applied a more traditional approach to assess literature, derive an HIT adoption model, and validated it with surveys in the case of diabetic patients [39]. The results are used to inform the design of Diabetes Self-Management Applications (DSMA) as impactful patient-centered tools, enabling diabetic patients to manage their health conditions and thereby prevent complications. The findings indicate that perceived financial risk, perceived privacy and security risk, technology anxiety, and facilitating conditions have significantly positive relationships with the intention to adopt DSMA. This will help to design the next-generation DSMA applications.

Discussion

Studies in CHI and PGHD have begun to show positive results that impact care delivery, improve patient-provider communication, and enhance health outcomes. Despite early positive results, significant barriers and challenges exist related to the use of PGHD, mobile technologies, and digital sensors within clinical settings.

A key finding from this review is the overall increase in acceptance of PGHD to be used by consumers, healthcare providers, and researchers. Polling data shows that the adoption of mobile apps for health has doubled in two years, reaching 33% in 2016, and, that patient data is believed to be beneficial for maintaining health [40]. Another poll from 13,000 consumers in Europe also reveals that 32% of the general populations in Europe are considering they are using wearable devices [41]. In fact, the survey has further shown that more than 78% of customers are willing to wear technology for health tracking. This can be attributed to the availability of smartphones and to the increasing prevalence of patient-reported outcomes and ecological momentary assessment technologies that allow for data collection [42]. While this signifies an enormous opportunity, it can also be stated as a concern for some segments of the population that may not have access to these technologies [29]. This disparity was not addressed in this current review but is something that should be addressed in future reviews.

Despite the increased use of PGHD, there exist gaps for how to use the data by both patients and providers. The first step was to collect and store data. Now, we need additional guidance regarding how to operationalize and interpret the data. In the following discussion, we will discuss the gaps existing in the two latter aspects.

First, operationalizing data can include adjusting clinical workflow, improving interoperability, and providing guidelines for
responsibility of the data. Data interpretation could include the use of a clear language, enhanced data visualization techniques, and built-in summaries of the data. The major barriers commonly faced by healthcare professionals in routine care include a lack of proper incentives and of a supportive infrastructure to integrate PGHD into the clinical workflow [42, 43]. These barriers limit the utilization of PGHD in clinical decision-making and make it hard to assess how much extra time it would take the healthcare professionals in a consultation, a home visit, or a care management call. As a result so far, the adoption by health professionals to incorporate PGHD into routine care is still slow. Given that 78% of US consumers would be willing to wear fitness and lifestyle and/or vital sign monitoring technology [6], the disparity between the interest of consumers to provide PGHD and the ability of providers to incorporate the data into their workflow is alarming. On the consumer side, current self-monitoring tools also lack flexibility, standardized formats, and mechanisms to share data with providers.

Second, for interpreting data, the major barriers are the lack of standard data models for integration and of a supportive infrastructure to make sense of the collected data for patients themselves. From the accumulated evidence in the past, self-monitoring has been shown to be good for activation, but not enough for sustaining behavior [44]. More frequent feedback has been expected as an effective counter strategy to address barriers such as stress level, lack of social support, and discomfort with recording that can affect adherence to self-monitoring [45, 46]. However, as pointed out in [47], in the area of mobile health, only a limited number of mHealth apps integrated health behavior theory and left room for future work.

In order to start addressing the missing support for healthcare professionals and the patients themselves, researchers and policy makers in the field need to re-examine evidence to select the most important opportunity areas and prioritize them. This review thus serves as a status update for the emerging evidence for the promising opportunity areas to pursue. In particular, the review points out the potential of secondary use of PGHD for a variety of applications, particularly with respect to the application of data analytics and mining techniques for the generation of new knowledge and to help create effective learning health systems. However, many open questions still exist and we anticipate an increased focus of research in interpreting the context of collected PGHD, with better automated techniques for cleaning PGHD collected from sensors, understanding the reliability of PGHD sensors, and PGHD interoperability.

One such gap relates to standard vocabularies and data models. There have been quite a few consortium-based and commercial efforts that have attempted to address these issues. For example, the new Fast Healthcare Interoperability Resources (FHIR) specification from the HL7 organization is designed to allow health consumers to share their data with clinical systems [48]. Patient-Centered Outcomes Research Institute (PCORI) is starting to collect cross-site patient ambulatory assessment data to enable patient-centered effectiveness study [49]. The creation of ecologically-valid tools when using ambulatory assessment can then lend support to the understanding of bio-psychosocial processes as they unfold naturally in time and in context. Emerging mobile health data and research platforms such as Apple HealthKit and ResearchKit have also added support to the interchangeability of PGHD for future clinical integration needs [50, 51].

Another open research question is the sense-making of PGHD. Many have observed the problems but so far no strong evidence has emerged for the right solutions yet. While only 10% of the population are using health-tracking devices/apps to help manage chronic conditions [52], the proportion of US health consumers accessing their health records has increased to 45% in 2016 (an increase of 67% in two years) [53]. The increasing access and connection between clinics and homes have indicated new opportunity areas to pursue. In particular, the review points out the potential of secondary use of PGHD for a variety of applications, particularly with respect to the application of data analytics and mining techniques for the generation of new knowledge and to help create effective learning health systems. However, many open questions still exist and we anticipate an increased focus of research in interpreting the context of collected PGHD, with better automated techniques for cleaning PGHD collected from sensors, understanding the reliability of PGHD sensors, and PGHD interoperability.

In the future, we expect more research to be conducted to identify the association between the top barriers and the possible intervention designs to address the different opportunity areas of PGHD in clinical care. Opportunity areas include diagnosis support (e.g., clinical decision support systems for better reliability and resolution), interaction support (e.g., interactive feedback apps for collaborative reflection and communication), focus on education goals (e.g., automatic synthesis of data logs for content absorbing), and a better understanding of patient values (e.g., care management support for better communication of contextual information). Only when a better assessment is done, then we can start addressing the health consumer demands. Research is also needed to extend the standardized consumer health vocabulary, in which more up-to-date concepts are included and similar symptoms can be grouped together to provide standardization. Research has started to build open-source Consumer Health Vocabulary (CHV) [57] to deal with this issue.

Conclusion

Current evidence indicates that the technological advancement and widespread availability of affordable consumer-grade devices are fueling research into using PGHD for better care. As we see a growing number of (pilot) developments using various mobile health technologies to collect PGHD, major gaps still exist in how to use the data by both patients and providers. Further research is needed to understand the impact of PGHD on clinical outcomes.

References
3. Chung AEA, Basch EM. Potential and Challenges


47. Borrelli B, Ritterband LM. Special Issue on eHealth and mHealth : Challenges and Future Directions for Assessment, Treatment, and Dissemination. Heal Psychol 2015;34:1205–8.

48. HL7 FHIR (Fast Healthcare Interoperability Resources) Specification [Internet]. Available from: https://www.hl7.org/fhir/


