

A 21st Century Embarrassment of Riches: The Balance Between Health Data Access, Usage, and Sharing

John H. Holmes¹, Lina F. Soualmia^{2,3}, Brigitte Séroussi^{3,4}

¹ University of Pennsylvania Perelman School of Medicine, Philadelphia, USA

² Normandie Université, Univ Rouen, LITIS EA 4108, Rouen, France

³ Sorbonne Université, Université Paris 13, Sorbonne Paris Cité, INSERM UMR_S 1142, LIMICS, Paris, France

⁴ AP-HP, Hôpital Tenon, Paris, France

Summary

Objectives: To provide an introduction to the 2018 International Medical Informatics Association (IMIA) Yearbook by the editors.

Methods: This editorial provides an overview and introduction to the 2018 IMIA Yearbook which special topic is: "Between access and privacy: Challenges in sharing health data". The special topic editors and section are discussed, and the new section of the 2018 Yearbook, Cancer Informatics, is introduced. Changes in the Yearbook editorial team are also described.

Results: With the exponential burgeoning of health-related data, and attendant demands for sharing and using these data, the special topic for 2018 is noteworthy for its timeliness. Data sharing brings responsibility for preservation of data privacy, and for this, patient perspectives are of paramount importance in understanding how patients view their health data and how their privacy should be protected.

Conclusion: With the increase in availability of health-related data from many different sources and contexts, there is an urgent need for informaticians to become aware of their role in maintaining the balance between data sharing and privacy.

Keywords

Informatics, Medical, Health Information Technology, IMIA Yearbook of Medical Informatics, health data sharing, health data privacy

Yearb Med Inform 2018;5-6

<http://dx.doi.org/10.1055/s-0038-1641213>

Between Access and Privacy: Challenges in Sharing Health Data

The theme for the 2018 International Medical Informatics Association (IMIA) Yearbook could not have been more timely. There truly is a tension between the access to, and the sharing of, health-related data, and this tension encompasses several concerns. First among these is that health-related data are being produced at an exponential rate, and this has been expressed in some quarters as a "gushing fire hydrant" with a nearly uncontrollable flow of water (data). This is a rather accurate metaphor of the challenges faced by informaticians, analysts, clinicians, and administrators in their quest to understand and act on these data.

Second, health-related data come from an ever-expanding universe of sources. No longer are we content to generate and work with only electronic health record or other health information systems data. Rather, our attention is turning to "non-traditional" data sources, such as geographic information systems, climatologic surveys and monitoring, wearable technology, and persistent environmental and behavioural monitoring systems such as one finds in "smart homes". Part of the reasons for this increased attention to these sources is availability, of course, but it is also driven by the developing domain of precision medicine and researchers' interest in the exposome and deep phenotyping,

which demands access to large volumes of highly heterogeneous data from previously unknown or unexplored data sources.

Third, as a function of these concerns, is the need to constrain the availability and use of these data in order to preserve the privacy and confidentiality of the people or corporate entities they surrogate. With the increasing availability of supplemental, health-related data, it is becoming much easier to identify an individual, not only as a discrete entity, but to track that person's trajectory in space and time. For example, a person's health record could be linked to her fitness tracking wearable or to an application on a smart phone that records not only her physical activity but also her location and the time spent in each activity. While such linkage helps to serve the goal of precision medicine to some extent, it also reveals much about the patient's behaviour- and possibly that of others- that would have otherwise been unavailable.

Finally, and perhaps most in need of attention by the informatics community, is the issue of sharing these data. Everyone should be familiar with the constraints on sharing data from the electronic health record at will. In the US, protections through the Health Insurance Portability and Accountability Act (HIPAA) restrict the sharing of identifiable health data, even from one health provider to another. More recently, the EU General Data Protection Regulation (GDPR) is even more restrictive, and applies to all types

of personal data, even those that are not health-related. However, there is a palpable tension between these regulations and the desire, if not need, to share sensitive health-related data. In the clinical domain, general practitioners who need to refer a patient for specialty care may find the transfer of the patient's health data to a specialist delayed or hampered by such regulations. Researchers, in their quest to create a dataset as complete as possible to examine the interactions between genome, phenome, and environment, may likewise encounter constraints on not only data linkage but also in obtaining needed data from a variety of sources that may themselves be hampered by law or other regulation in releasing or sharing data with a researcher who requests them.

Highlights in this Year's Yearbook

This year's keynote paper, contributed by Carolyn Petersen, amply addresses all of the concerns raised above. She provides a call to action on the part of informaticians to participate fully in the discussion about data privacy, particularly for the protection of patients. This call is made all the more urgent by the increasing realization of precision medicine, which relies on that expanding universe of data discussed above. Her article is made all the more compelling by focusing on privacy concerns from the perspective of patients, all of which are of great importance to our profession. One of these concerns, that is her contention that "privacy is global", deserves a special mention. When it comes to health-related data, we live in a borderless world, although there exist regulations that constrain the use and sharing of such data, as noted above. It behoves us informaticians to become aware of, and understand, the various legal and privacy-preserving constraints as they affect data sharing.

In keeping with the theme of this year's edition of the Yearbook, Professors Shankar and Anderson offer a survey in the Special Section about the issues

of data access, sharing, and privacy in a specialized domain, clinical decision support. This is an area which has not been considered by many as one where data sharing is an issue. Yet, in the age of health information exchanges and health data portability, clinical decision support systems and the data they rely on deserve considerable attention. This is especially true in situations where patients may move from one healthcare provider to another, unrelated one, perhaps in another country. In the case of such patients who also have a severe chronic condition that requires considerable management expertise, where decision support is a critically important adjunct to appropriate care, the sharing of data could be life-saving. In addition to data, however, Shankar and Anderson point to the sharing of algorithms, which could be proprietary or perhaps inappropriate for one or more clinical settings. This poses a potential roadblock for researchers as well, as they might seek to translate a decision support algorithm from one setting to another in the service of software or clinical outcome evaluation.

As will be seen from the best papers and surveys in each of the remaining sections, there is much to learn about health-related data availability, sharing, and privacy protections. It is incumbent on all informatics practitioners and researchers to become fluent in the issues raised in these sections, with the hope that they will contribute to this important dimension of our profession.

IMIA President's Statement

Attention should be drawn to Dr. Christoph U. Lehmann's Yearbook contribution for his inaugural IMIA President's Statement. In addition to his moving reflection on taking the gavel from now-Past President Hyeoun-Ae Park, he offers important news about IMIA. This includes the induction of the founding class of 119 Fellows of the International Academy of Health Sciences Informatics at the Medical Informatics Europe (MIE) conference in Gothenburg, Sweden, in April 2018. Congratulations to the new Fellows!

Changes in this Year's Yearbook and in the Yearbook Editorial Team

Readers should be pleased to see the new section of the Yearbook on Cancer Informatics. Although it was announced in last year's Yearbook, this is the first year that it appears as a full section, with its own survey and synopsis. Congratulations are due to the two section editors, Jeremy L. Warner, Associate Professor of Medicine and Biomedical Informatics at Vanderbilt University (TN, USA), and Debra Patt, Vice President of Texas Oncology (TX, USA).

Thierry Lecroq, Professor of Bioinformatics at Normandy University (Rouen, France), and Hélène Dauchel, Associate Professor of Biology at Normandy University (Rouen, France) have left the Yearbook editorial team in 2017 after being the editors of the section Bioinformatics and Translational Informatics during a few years. This section is not included in this year's Yearbook, but two new section editors have been enrolled and will serve for this section in the 2019 edition of the Yearbook.

In 2018, Thomas Deserno (Technische Universität Braunschweig and Medizinische Hochschule Hannover, Braunschweig, Germany), Alexander Hörbst (eHealth Research and Innovation Unit, UMIT-University of Health Sciences, Medical Informatics and Technology, Hall in Tirol, Austria), Johanna Kaipio (Department of Computer Science, Aalto University, Espoo, Finland), Dipak Kalra (University College London, London, UK), and Annie Lau (Centre for Health Informatics, Australian Institute of Health Innovation, Macquarie University, Australia) have joined the editorial team to serve as section editor of Sensors, Signals and Imaging Informatics, Clinical information Systems, Human Factors and Organizational Issues, Clinical Research Informatics, and Consumer Health Informatics sections, respectively.

From 2019, Dr. Kate Fultz Hollis is replacing Dr. John H. Holmes, who is stepping down after two years of service to the Yearbook. Congratulations are due to Kate as she takes on this responsibility; please join the Yearbook editors in welcoming her to the team.