

Bringing Genome Tests into Clinical Practice

I.N. Sarkar^{1,2,3}

¹ Center for Clinical and Translational Science

² Department of Microbiology and Molecular Genetics

³ Department of Computer Science

University of Vermont, Burlington, VT USA

Summary

Objective: To provide a review of recent progress in the use of genomic data in clinical contexts.

Method: Survey of key articles selected from the previous year (2012).

Result: Two major themes emerged as areas of focus in 2012: (1) integration of genomic data into electronic health records; and (2) determining the clinical utility of genomic tests.

Conclusion: Significant advances are being made towards the integration of genomic data such that they may be useful for clinical decision making. While the short term advances will likely still be seen in clinically valid genotype-phenotype research, there have been promising advances towards developing genomic tests that may be of clinical utility in the near future.

Keywords

Translational Bioinformatics, Electronic Health Records, Genomics, IMIA Yearbook

Yearb Med Inform 2013;172-4

Introduction

The completion of the human genome, at a cost of US\$2.7 billion, more than a decade ago promised to usher in a new era of medicine. As the cost to acquire entire human genomes continues to decline, on pace to cost US\$1000 or less within a decade, the question still remains: Will genomic tests revolutionize the practice of medicine? In the last year, there has been much progress towards addressing this query, in particular with regards to: (1) considering the technical challenges of integrating genomic and clinical data, and (2) understanding the clinical implications of genomic based testing.

Neither of these areas should be considered necessarily “new” research areas for the biomedical informatics enterprise. However, the recent maturity in these areas suggests that medical decisions may increasingly be guided by genomic information. This is not to imply that genomic medicine will completely replace conventional medical decision making. Instead, the incorporation of genomic medicine may offer a complementary cadre of decision support tools that may provide clinicians with crucial information to enable personalized medicine. From a genomic perspective, clinical decisions are generally guided by phenotypes (e.g., as collections of observations such as signs and symptoms) that are often the *result* of dysfunction that may be of genomic origin, whereas genomic-based decision tools offer a detailed perspective to the underpinning *cause* of genetic diseases and disorders. It is thus important to note that ultimate success of genome-based decision making will require not only genome sequence based knowledge, but also understanding how ge-

nomics-based clinical decisions complement classical clinical tools of high diagnostic reliability or those that implicitly incorporate genomic knowledge (e.g., family history).

This survey will first consider the advances made towards the integration of genomic data into electronic health record (EHR) systems. Next, the progress towards developing genomic tools that may be of clinical utility will be discussed. Throughout both of these topics, particular emphasis is given to studies and discussion from the last year (2012).

EHRs with Genomes

The integration of genomic information into electronic health records (EHRs) has been the substrate of many discussions: How much genomic data need to be included? Will the inclusion of genomic data result in information overload in already complex clinical environments? Can the EHR infrastructure accommodate the voluminous genomes and associated annotations? The integration of genomic information alongside the array of information that is available within electronic health records will be an essential step towards enabling clinical decisions that leverage genomic information.

Masys, *et al.*[1] provide a list of seven technical “desiderata” that should be considered for the inclusion of genomic data into EHRs. These desiderata provide direction to include genomic data while balancing their potential clinical and research utility. This balance is one that has not often been sought between clinicians and biomedical researchers, but is one that is suggested as essential for the future success of genomic-based

medicine. Altman [2] also highlights the data integration challenge to meld genomic data alongside classic clinical data within EHRs. The incorporation of genomic data into machine accessible form, he suggests, will enable unprecedented unification of population and individual level data such that they can be incorporated into a Bayesian paradigm for personalizing medicine (e.g., for providing truly personalized drug therapies based on population and individual genetics).

Genomic data may be considered in a similar manner as other clinical laboratory tests, as discussed in the viewpoint from Biesecker, *et al.* [3]. Towards this end, as a practical step towards integrating genomic data within an EHR, Jing, *et al.* [4] demonstrated the feasibility to leverage the Continuity of Care Record (CCR) standard. In their study, they were able to extend the CCR standard to represent genomic test results as an extended laboratory test. Their study highlights the possibility of leveraging existing standards used in EHR implementations that might provide a truly integrated genomic clinical record.

Whilst the incorporation of genomic data into EHRs will position the clinical enterprise to be prepared for a future of genome guided medicine, the most significant short term result will be the research potential. The ability to perform population level genotype-phenotype studies will enable the deep insights into the genomic etiology of many diseases and disorders. As reviewed by Kohane [5], EHR-Driven Genomic Research (EDGR) will be an essential endeavor to bring together clinical phenotypes (as represented in EHRs) and genotypic information (as acquired from contemporary sequencing technologies or genome wide scans). The increased adoption of comprehensive EHR systems alongside improvements in biobanking and sequencing technologies also suggest the EDGR projects will provide a cost-effective means to provide insights into clinical phenotype-genotypes.

Recent years have seen a growing number of EHR-linked biobanks. One of the most successful examples of an EDGR initiative, the electronic Medical Records and GENome (eMERGE) project, has been developing strategies for robust clinical phenotyping from EHR data alongside genomic data for the purposes of identifying clinically mean-

ingful genotype-phenotype correlations. As of 2012, eMERGE consists of nine sites across the United States. The eMERGE consortium has recently developed a clearinghouse for clinical phenotyping algorithms (<http://pheKB.org>) that is publicly accessible and includes performance values for the algorithms across the eMERGE consortium institutions. In the last year, the eMERGE consortium have described potential applications related to areas such as studying genetic risk factors associated with Type 2 Diabetes Mellitus [6] and identification of cataract cases from EHR data [7]. A particularly promising eMERGE EDGR study identified a putative protective *ApoE* variant in African Americans [8] associated with Low Density Lipoprotein Cholesterol.

Pharmacogenomics, which focuses on identifying genetic variation correlations that influence drug response, continues to be an important area of growth in biomedical informatics. Studies from the last year have demonstrated how EHR-linked biobanks could be used to as a source of pharmacogenomics knowledge to enable real-time, preemptive drug dosing decisions [9]. There are a growing number of medical centers that have started to make significant investments in support of pharmacogenomic knowledge driven clinical decisions. These programs are starting to directly address the challenges of pharmacogenomics data integration such that clinicians can better leverage more complete personalized medicine information available at the point of care [10-12].

In summary, the recent progress towards the integration of genomic and clinical data are helping enable a new generation of studies that uniquely leverage relationships between genomic and clinical data. In support of the EDGR paradigm, we have seen significant advances in disease focused studies – such as those that are the subject of the eMERGE consortium. The true vision of the EDGR paradigm will be realized as increasing numbers of EHRs are able to incorporate genomic data, thus supporting researchers in their quest to develop approaches for identifying genotype-phenotype correlations across a broader spectrum of diseases and clinicians in their need to have clinically meaningful, personalized information at the point-of-care.

Clinically Actionable Genomic Analysis

Parallel to the discussion about how to integrate genomic data into EHRs, many in the clinical and research community continue to question the potential utility of genomic data. As described in the previous section, perhaps the only notable exception is pharmacogenomic information, where there is a direct path from genetic information to clinical action. There are many research implications for the integration of clinical and genomic data to understand the underlying cause of genetic diseases (e.g., to help develop targeted treatment regimens), but it is not yet entirely clear whether the identification of genomic correlations with disease will provide better diagnostics.

Do, *et al.* [13] compared the predictive power of standard of care family history versus genomic based single nucleotide polymorphism (SNPs, which are single point mutations that may be correlated with disease susceptibility) for predicting complex diseases. Their findings suggest that neither family history nor SNP-based analyses are by themselves uniform predictors of disease. However, their results do suggest that family history and SNP-based analyses together provide complementary, clinically informative information that could potentially be used in differential diagnoses of diseases of a genetic origin. This finding further underscores the 2004 plea from Guttmacher, *et al.* [14] to clinicians for gathering better quality family history data that can be used in complement with genomic data.

The clinical utility of genomic tests still remains in question – Roberts, *et al.* [15] suggest that the vast majority of genomic test results do not yield clinically useful information. Nonetheless, they do suggest that identification of genetic predispositions may help guide lifestyle changes or more regular screenings to prevent or lessen the severity of disease. However, it must be emphasized that the reliability of genomic tests to produce robust models for predicting disease can vary. Morgan, *et al.* [16] highlight this type of variability a study where a significant difference was found between disease risk models developed from commercial genome

arrays (e.g., those used by direct-to-consumer genomics services) versus those from full genome sequence based analyses. A review by Bodurtha and Strauss [17] describes a range of potentially insightful perinatal genomic tests that can be done; however, the clinical utility of these tests remains unclear.

The challenges in determining the clinical utility of genomic tests is largely due to the limited number of randomized controlled trials (RCTs), which are the accepted “gold standard” mechanism to assess the clinical utility of a given clinical tool, associated with determining the utility of genomic tests. Thus, while the biomedical informatics and medical genetics communities may continue to discover clinically valid genotype-phenotype correlations, until the clinical utility of the correlations can be determined, their adoption will remain minimized. To address this specific point, Goddard, *et al.* [6] have proposed the use of a Comparative Effectiveness Research (CER) framework to assess the clinical utility of genome tests. In the context of determining the clinical utility of genomic tests for cancer, they suggest an approach that builds on the four criteria established in 2004 by the United States Center for Disease Control Office of Public Health Genomics framework for evaluating a genetic test, ACCE [18]: (1) analytic validity, (2) clinical validity, (3) clinical utility, and (4) the ethical, legal, and social implications. Goddard, *et al.* suggest that the clinical utility be evaluated through the use of observational studies from analysis of EHR-based data. They also note the challenges of clinical phenotyping from EHRs are described, and therefore further support the importance of the promising EDGR work that is being done through initiatives like eMERGE as described in the previous section.

Overall, it may still be early to see a direct impact of the increasing number of genotype-phenotype correlations on clinical care. Genomic analyses may very well complement current genome based analyses that are part of the standard of care (e.g., acquisition of family history), but there is enough variability in disease prediction models that it may be difficult to interpret genomic test findings clinically. Nonetheless, the realization of the need for the demonstration of clinical utility has led to the development of

promising frameworks that may ultimately lead to greater adoption and use of genomic tests in clinical practice.

Conclusions

Collectively, we have seen great progress towards the clinically meaningful integration of genomic information into EHR systems, especially in the last year. The integration of genomic and clinical data has catalyzed retrospective and prospective genotype-phenotype studies, which further supports the positive outlook of EDGR initiatives like eMERGE. The development of an EDGR framework also better enables the development of real-time, personalized pharmacogenomic clinical decision support systems. However, aside from pharmacogenomics contexts, the clinical implications of genomic integration with clinical data remain promissory. Nonetheless, the increased integration of genotypic and more robust clinical phenotyping strategies suggest that the synergy between genomic and classical clinical data will continue to increase in the very near future. At the minimum, we will see an increased number of genomic markers that provide unprecedented insights into the clinical etiology of many diseases and disorders. Some of these findings may lead to genome-informed clinical decision support tools that will complement the current arsenal of diagnostics.

References

- Masys DR, Jarvik GP, Abernethy NF, Anderson NR, Papanicolaou GJ, Paltoo DN, et al. Technical desiderata for the integration of genomic data into Electronic Health Records. *J Biomed Inform* 2012 Jun;45(3):419-22.
- Altman RB. Personal genomic measurements: the opportunity for information integration. *Clin Pharmacol Ther* 2013 Jan;93(1):21-3.
- Biesecker LG, Burke W, Kohane I, Plon SE, Zimmern R. Next-generation sequencing in the clinic: are we ready? *Nat Rev Genet* 2012 Nov;13(11):818-24.
- Jing X, Kay S, Marley T, Hardiker NR, Cimino JJ. Incorporating personalized gene sequence variants, molecular genetics knowledge, and health knowledge into an EHR prototype based on the Continuity of Care Record standard. *J Biomed Inform* 2012 Feb;45(1):82-92.
- Kohane IS. Using electronic health records to drive discovery in disease genomics. *Nat Rev Genet* 2011 Jun;12(6):417-28.
- Goddard KA, Knaus WA, Whitlock E, Lyman GH, Feigelson HS, Schully SD, et al. Building the evidence base for decision making in cancer genomic medicine using comparative effectiveness research. *Genet Med* 2012 Jul;14(7):633-42.
- Peissig PL, Rasmussen LV, Berg RL, Linneman JG, McCarty CA, Waudby C, et al. Importance of multi-modal approaches to effectively identify cataract cases from electronic health records. *J Am Med Inform Assoc* 2012 Mar-Apr;19(2):225-34.
- Rasmussen-Torvik LJ, Pacheco JA, Wilke RA, Thompson WK, Ritchie MD, Kho AN, et al. High density GWAS for LDL cholesterol in African Americans using electronic medical records reveals a strong protective variant in APOE. *Clin Transl Sci* 2012 Oct;5(5):394-9.
- Schildcrout JS, Denny JC, Bowton E, Gregg W, Pulley JM, Basford MA, et al. Optimizing drug outcomes through pharmacogenetics: a case for preemptive genotyping. *Clin Pharmacol Ther* 2012 Aug;92(2):235-42.
- Hicks JK, Crews KR, Hoffman JM, Kornegay NM, Wilkinson MR, Lorier R, et al. A clinician-driven automated system for integration of pharmacogenetic interpretations into an electronic medical record. *Clin Pharmacol Ther* 2012 Nov;92(5):563-6.
- Delaney JT, Ramirez AH, Bowton E, Pulley JM, Basford MA, Schildcrout JS, et al. Predicting clopidogrel response using DNA samples linked to an electronic health record. *Clin Pharmacol Ther* 2012 Feb;91(2):257-63.
- Pulley JM, Denny JC, Peterson JF, Bernard GR, Vnencak-Jones CL, Ramirez AH, et al. Operational implementation of prospective genotyping for personalized medicine: the design of the Vanderbilt PREDICT project. *Clin Pharmacol Ther* 2012 Jul;92(1):87-95.
- Do CB, Hinds DA, Francke U, Eriksson N. Comparison of family history and SNPs for predicting risk of complex disease. *PLoS Genet* 2012 Oct;8(10):e1002973.
- Guttmacher AE, Collins FS, Carmona RH. The family history—more important than ever. *N Engl J Med* 2004 Nov 25;351(22):2333-6.
- Roberts NJ, Vogelstein JT, Parmigiani G, Kinzler KW, Vogelstein B, Velculescu VE. The predictive capacity of personal genome sequencing. *Sci Transl Med* 2012 May 9;4(133):133ra58.
- Morgan AA, Chen R, Butte AJ. Clinical utility of sequence-based genotype compared with that derivable from genotyping arrays. *J Am Med Inform Assoc* 2012 Jun;19(1e):e21-7.
- Bodurtha J, Strauss JF, 3rd. Genomics and perinatal care. *New Engl J Med* 2012 Jan 5;366(1):64-73.
- <http://www.cdc.gov/genomics/gtesting/ACCE/index.htm>.

Correspondence to:

Dr. Indra Neil Sarkar
Center for Clinical and Translational Science
University of Vermont
89 Beaumont Avenue, Given Courtyard N309
Burlington, VT 05405 USA
Tel: +1 802 656 8283
Fax: +1 802 656 4589
E-mail: Neil.Sarkar@uvm.edu