Leveraging Genomic Data for Advancing Research

Christopher G. Chute, MD DrPH
Professor Biomedical Informatics
Chair, Data Governance
Mayo Clinic

Chair, International Classification of Disease Revision, WHO
Chair, ISO Technical Committee 215 on Health Informatics

IOM: Genomics-Enabled Learning Health Care Systems

8 December, 2014

Go Hopkins!

Sources and Types of Genomic Data Research

Genomic

- Genomic studies and protocols
  - GWAS, Next Gen sequencing
  - eMERGE, Pharmacogenomics Research Network
- Shared Repositories
  - dbGaP
  - MayoGC (enterprise version of dbGaP)

Clinical

- High-throughput Clinical Phenotyping of EHRs

Mayo's Local "dbGaP"

MAYO CLINIC
Mayo Clinic Proceedings
A Peer-Reviewed Medical Journal Sponsored by Mayo Clinic and Authored by Physicians Worldwide

Mayo Clin Proc. Jul 2011; 86(7): 606–614.
doi: 10.4065/mcp.2011.0178
PMCID: PMC3127556

Mayo Genome Consortia: A Genotype-Phenotype Resource for Genome-Wide Association Studies With an Application to the Analysis of Circulating Bilirubin Levels

Suzette J. Bielinski, PhD, High Seng Chai, PhD, Jyotishman Pathak, PhD, Jayant A. Talwalkar, MD, Paul J. Limburg, MD, Rachel E. Gullerud, BA, Hugues Sicotte, PhD, Eric W. Klee, PhD, Jason L. Ross, MBA, Jean-Pierre A. Kocher, PhD, Iftikhar J. Kullo, MD, John A. Heit, MD, Gloria M. Petersen, PhD, Mariza de Andrade, PhD, and Christopher G. Chute, MD, DrPH

Author information ▶ Copyright and License information ▶

Sources and Types of Genomic Data Clinical

- Prospective Genotyping/ Sequencing
  - Piloted in PGx project (PGRN and eMERGE)
  - Used PGRN Seq platform (84 pharmacogenes)
  - Increased use of sequencing
    - Variable policies on verification (CYP2D6)
- Clinical Diagnostic Genotyping
  - Small sampling frames
- Data linkage with clinical information
  - Must be preserved for genome to phenome

PGx Project at Mayo

MAYO CLINIC

Mayo Clin Proc. 2014 Jan;89(1):25-33.

ORIGINAL ARTICLE

Preemptive Genotyping for Personalized Medicine: Design of the Right Drug, Right Dose, Right Time—Using Genomic Data to Individualize Treatment Protocol

Suzette J. Bielinski, PhD, MEd; Janet E. Olson, PhD; Jyotishman Pathak, PhD;
Richard M. Weinshilboum, MD; Liewei Wang, MD, PhD; Kelly J. Lyke; Euijung Ryu, PhD;
Paul V. Targonski, MD, PhD; Michael D. Van Norstrand, MD, PhD;
Matthew A. Hathcock, MS; Paul Y. Takahashi, MD; Jennifer B. McCormick, PhD, MPP;
Kiley J. Johnson, MS, CGC; Karen J. Maschke, PhD; Carolyn R. Rohrer Vitek, MS;
Marissa S. Ellingson, MS, CGC; Eric D. Wieben, PhD; Gianrico Farrugia, MD;
Jody A. Morrisette, MBA; Keri J. Kruckeberg, BA; Jamie K. Bruflat, MS; Lisa M. Peterson, BS;
Joseph H. Blommel, MS; Jennifer M. Skierka, BS; Matthew J. Ferber, PhD;
John L. Black, MD; Linnea M. Baudhuin, PhD; Eric W. Klee, PhD; Jason L. Ross, MA;
Tamra L. Veldhuizen, BS; Cloann G. Schultz, MS; Pedro J. Caraballo, MD;
Robert R. Freimuth, PhD; Christopher G. Chute, MD, DrPH; and Iftikhar J. Kullo, MD

Learning Healthcare System Leveraging Genomic Information

- Must establish and maintain patient trust
  - Expectation that research is the norm
  - Establish and atmosphere of confidentiality
- Rigorous adherence to good data practices
  - Establish and enforce confidentiality policies
  - Maintain encryption, including at rest
  - Separate identifiers after dataset linkage
  - Prohibit any clinical data on portable devices
  - Invoke multi-factor authentication

Will Big Data Save Us?

Genetics inMedicine | REVIEW

Genet Med 15: 802-809;

Oct, 2013

Some experiences and opportunities for big data in translational research

Christopher G. Chute, MD, DrPH[1], Mollie Ullman-Cullere, MS, MSE[2], Grant M. Wood, BS[3], Simon M. Lin, MD[4], Min He, PhD[4] and Jyotishman Pathak, PhD[1]

Health care has become increasingly information intensive. The advent of genomic data, integrated into patient care, significantly accelerates the complexity and amount of clinical data. Translational research in the present day increasingly embraces new biomedical discovery in this data-intensive world, thus entering the domain of "big data." The Electronic Medical Records and Genomics consortium has taught us many lessons, while simultaneously advances in commodity computing methods enable the academic community to affordably manage and process big data. Although great promise can emerge from the adoption of big data methods and philosophy, the heterogeneity and complexity of clinical data, in particular, pose additional challenges for big data inferencing and clinical application. However, the ultimate comparability and consistency of heterogeneous clinical information sources can be enhanced by existing and emerging data standards, which promise to bring order to clinical data chaos. Meaningful Use data standards in particular have already simplified the task of identifying clinical phenotyping patterns in electronic health records.

Genet Med advance online publication 5 September 2013

Key Words: clinical data representation; big data; genomics; health information technology standards

Problems and Barriers
Specific to Clinical Genomic Research

- Patient consenting
  - Full and fair consenting practices
  - Allow opt-out for genomic data linkages
- Comparable and consistent data rendering
  - * allele nomenclature is collapsing
  - Achieve consensus on clinical naming of genomic variants
- Genomic results much be machine readable
  - Textual reports are useless for clinical decision support

VIEWPOINT

JAMA
April 10, 2013, Vol 309, No. 14

Genomic Medicine, Health Information Technology, and Patient Care

Christopher G. Chute, MD, DrPH
Isaac S. Kohane, MD, PhD

CELEBRATING THE TENTH ANNIVERSARY OF COMPLETing the draft human genome sequence in 2011, authors from the National Human Genome Research Institute of the US National Institutes of Health outlined the influence of genomic understanding across 5 domains: structure, the biology of the genome, the biology of disease, medicine, and improvements in health care.[1] The authors assert that this is the era of enhanced genomic understanding of medicine, which is expected to usher in improvements in health care effectiveness by the end of this decade. It is thus fitting to explore how health information technology will contribute to or hamper the promise of genomic medicine.

pital and medical practice in the country. Given the accelerating pace of genomic discovery, this is neither efficient nor scalable. Any expectation that a clinician can or should "know" the vast permutation of emerging genomic influences on disease risk, treatment, or prognosis, as well as the interactions of these influences with drugs or other diseases or, most confusingly, their co-occurrence with other genomic or environmental factors, is unrealistic.

The state of the art for academic medical centers in 2013 is determining a small number of relatively high-profile genomic variants from some or all of their patients at risk for specific drug treatments and integrating these findings into the electronic health records (EHRs) of those patients. Then, if a drug such as warfarin, clopidogrel, mercaptopurine, or codeine is ordered and a clinically significant drug-gene interaction is known, an alert to the physician or pharmacist is made, and in some settings an alternative recommended

Policies to Advance Clinical Research on Genomic Information

- Standard consent language
  - No reason every academic medical center must reinvent clear language
  - Facilitate portability (ACO, population health)
- Data donor culture
  - Pattern after chic fashion of organ donor status
  - Patient engagement in the importance of learning what helps and what hurts