Not Your Father's PDF: New forms of Genomic Sequence Knowledge Representation to Support Learning Healthcare Systems

Daniel Masys, MD

Affiliate Professor

Biomedical & Health Informatics
University of Washington, Seattle

NHGRI Future Opportunities for
Genome Sequencing and Beyond
July 28, 2014

Most common current method for delivery of DNA analysis into clinical operations

DEPARTMENT OF PATHOLOGY
Nashville, Tennessee 37232

Michael Laposata, M.D., Ph.D.
Medical Director, Diagnostic Laboratories
CLIA #44D0659066

Name:	Sex:	Laboratory Number:	VUH#:
[redacted]	F	[redacted]	[redacted]

Referral Source: Dr. Kim Ely

Reason for Request: DNA Analysis for *KRAS* Mutations

Type of Specimen: Paraffin-Embedded Tissue (Block #: [redacted])

Date Received: 2/12/10

Date of Report: 2/18/10

Interpretation: *KRAS* Mutation NOT Detected

Mutations Tested Include:
G12A, G12C, G12D, G12R, G12S, G12V, G13C, G13D

The *KRAS* gene (12p12) is a member of the Ras family of proto-oncogenes, and encodes a protein containing guanosine nucleotide triphosphate hydrolysis activity (known more commonly as a GTPase). These proteins are active when bound to guanosine triphosphate (GTP) and inactive when bound to guanosine diphosphate (GDP). *KRAS* is membrane bound, is activated by growth factor receptors, and through BRAF, stimulates the MAPK/ERK pathway resulting in transcription and cell proliferation. *KRAS* mutations are observed in colon cancer (40-50%), lung cancer (20-30%) and pancreatic cancers (90%). Conserved missense mutations in codons 12 and 13 result in prolonged binding of GTP and constitutive activation of RAS proteins, thereby leading to uncontrolled cell proliferation.

Progressive and/or metastatic non-small cell lung adenocarcinomas are often treated with inhibitors of the EGFR receptor as a second line therapy. However, it has been shown that tumors, which harbor mutations in codons

Fancy color version of the same problem

FOUNDATIONONE™ | FOUNDATION MEDICINE®

How to Read the FoundationOne™ Report

THE FIRST PAGE OF THE REPORT SHOWS
the patient and ordering physicians information...

The genomic alterations identified in the hundreds of genes assayed...

The targeted therapies that may be relevant based on those genomic alterations both in the patients tumor type and in other tumor types...

and whether there are relevant clinical trials.

Genomic Alterations Detected	FDA Approved Therapies (in patient's tumor type)	FDA Approved Therapies (in another tumor type)	Potential Clinical Trials
EGFR N771_P772>KFP	Erlotinib Gefitinib	Cetuximab Panitumumab	Yes, see clinical trials section
CCND1 amplification	None	None	Yes, see clinical trials section
ARID1A	None	None	None

How to ensure that genomic medicine has little or no impact on healthcare:

rely on
clinicians
Reading
and
Remembering
clinical reports and the published literature

Problems with treating genomic analysis in same fashion as other professionally interpreted clinical data

- Lossy compression: many DNA features observed, only a few clinically relevant reported, remainder discarded
- Interpretation inextricably bound together with primary observations in a document format
- Document reporting format not amenable to parsing for automated machine interpretation and decision support
- Much more unknown than known about genomic effects, and science changing rapidly

Journal of Biomedical Informatics xxx (2012) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

Technical desiderata for the integration of genomic data into Electronic Health Records

Daniel R. Masys [a,*], Gail P. Jarvik [b,c], Neil F. Abernethy [a], Nicholas R. Anderson [a], George J. Papanicolaou [d], Dina N. Paltoo [e], Mark A. Hoffman [f], Isaac S. Kohane [g], Howard P. Levy [h]

[a] Division of Biomedical and Health Informatics, Department of Medical Education and Biomedical Informatics, University of Washington, Seattle, WA 98195-7240, United States
[b] Division of Medical Genetics, Department of Medicine, University of Washington, Seattle, WA 98195-7720, United States
[c] Department of Genome Sciences, University of Washington, Seattle, WA 98195-7720, United States
[d] Division of Prevention and Population Sciences, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD, United States
[e] Advanced Technologies and Surgery Branch, Division of Cardiovascular Sciences, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD, United States
[f] Cerner Corporation, Kansas City, MO, United States
[g] Harvard-MIT Division of Health Sciences and Technology, Bioinformatics & Integrative Genomics, Cambridge, MA, United States
[h] Division of General Internal Medicine and McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University, Baltimore, MD, United States

ARTICLE INFO

ABSTRACT

Output of workshop on “Integration of Genetic Test Results into Electronic Medical Records” convened by the National Heart Lung and Blood Institute, Bethesda, MD August 2-3, 2011

7 desiderata for genomic sequence data in EHRs

1. Lossless data compression from (high volume) primary observations to clinically relevant subsets.
2. Since methods will change, molecular lab results carry observation methods with them (LOINC model)
3. Compact representation of clinically actionable subsets for optimal performance (clinician thinkspeed = 250msec)
4. Simultaneously support for human-viewable formats (with links to interpretation) and formats interpretable by decision support rules.
5. Separate primary sequence data (remain true if accurate) from clinical interpretations of them (will change with rapidly changing science)
6. Anticipate the boundless creativity of Nature: multiple somatic genomes, multiple germline genomes for each individual over their lifetime.
7. Support both individual care and discovery science

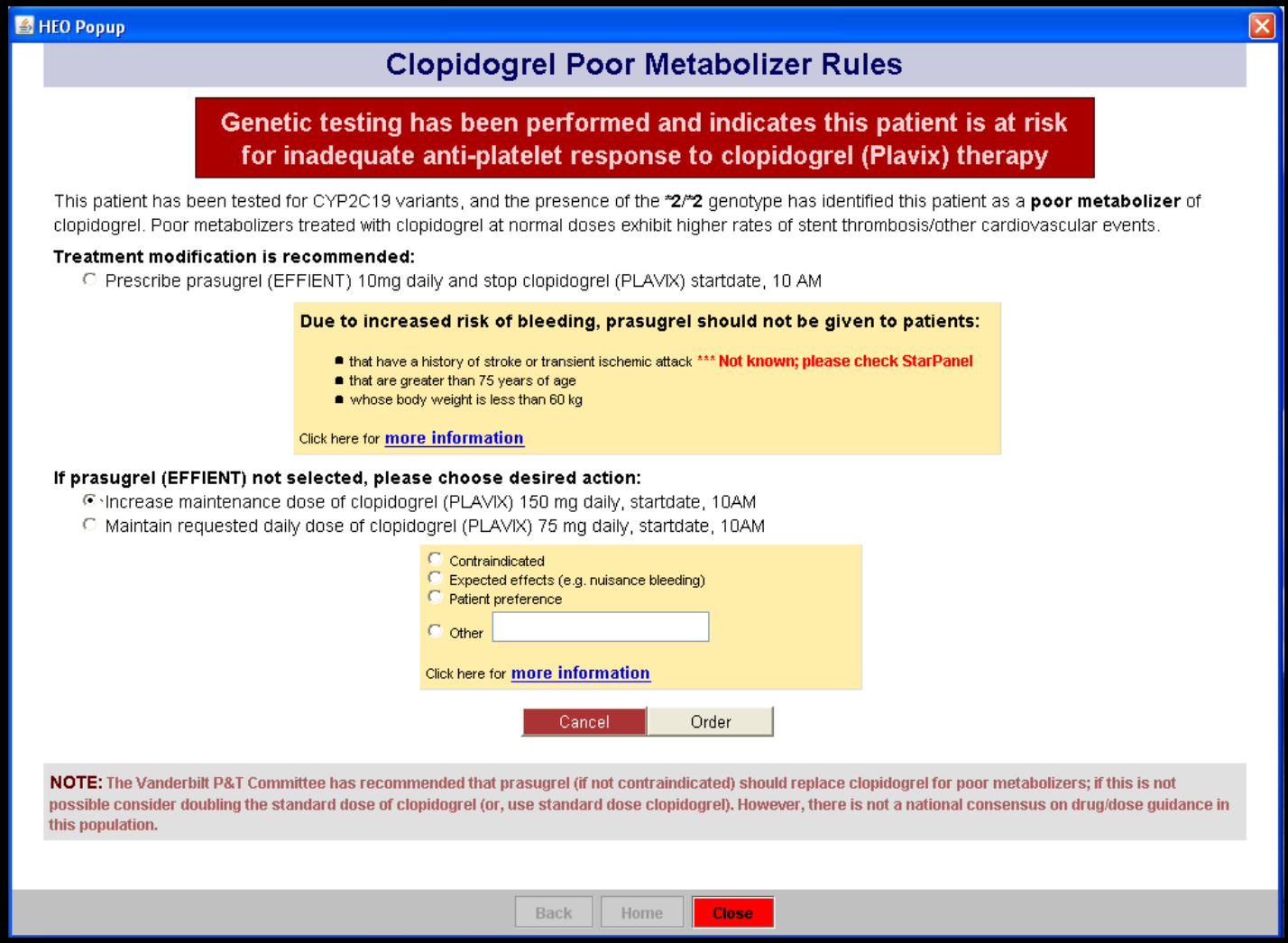
The opportunity for NHGRI

- Create a scaleable national capacity for genomically enabled clinical decision support by:
  - Supporting creation of a (Wikipedia-like) **closed loop** public computing infrastructure for guiding clinical care based on molecular variation, that
  - Learns and Improves **whether or not** clinicians follow the guidance.

The scope of decision support

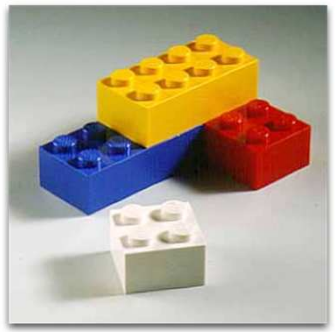
- “Rule based systems” do not mean providers must follow rules. Rules in informatics context = computerized approach to identification of characteristics.
- Examples of interventions
  - Educational prompts: here is additional general information to consider in this setting.
  - Data gathering prompts: given what is known about this {genotype|phenotype|genotype+phenotype}, it would be helpful to get this additional observation or testing.
  - Guidance that improves certainty of diagnosis given data currently available.
  - Guidance for best-evidence-based therapy selection
  - Information relevant to prevention and/or prognosis.

Example of Patient-specific decision support as seen by providers at the moment of prescribing:



Key computer technology: event monitor

Building Blocks for the infrastructure



- Standards for electronic “decision support packages” containing:
  1. Recognition logic for conditions of interest as represented in EHR systems (both genotype and phenotype)
  2. Guidance for target users (clinician, patient, family)
  3. Recognition logic for “closed loop decision support”: process or outcome measure to monitor, along with record of whether user accepted or rejected guidance
- Decision support authoring systems: tools to enable local ‘best practice rules committees’ to easily import, review, and implement decision support packages received from a Decision Support Library

A Systems Approach to Scaling-up

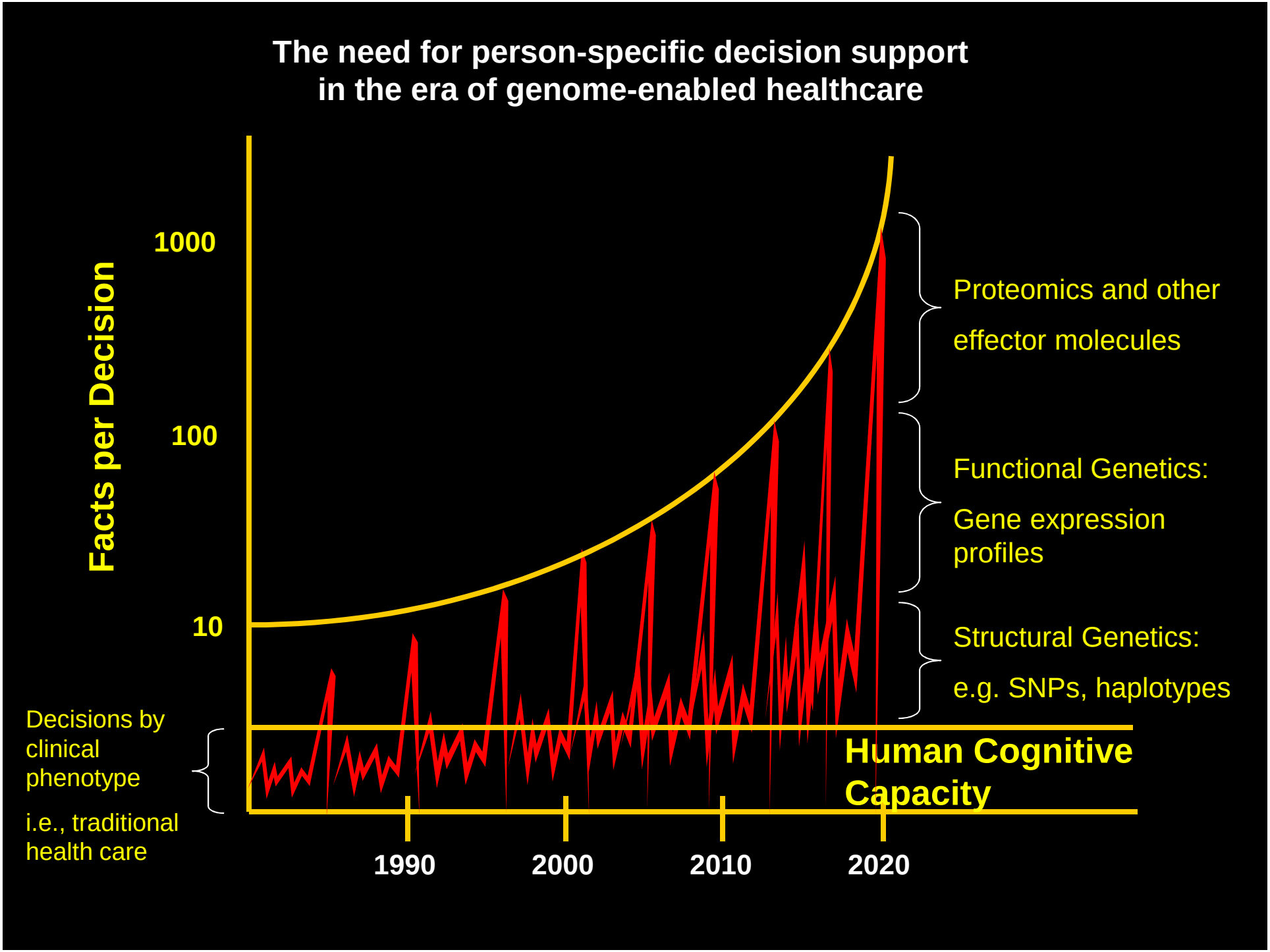
1. A continuously updated Public Library of clinical decision support 'packages' (an information commons): a Wikipedia for genome-enabled health decision support.
2. Event monitors embedded in EHR systems
3. System-generated alerts at the "teachable moment" of diagnostic testing and therapy decision making
4. Automated tracking of outcomes vs. user decisions: a learning healthcare system of providers and patients

Closing the Loop nationally

- Quid pro quo for use of public library clinical decision support packages would be automated local monitoring whether guidance was accepted or rejected, and whether subsequent clinical events (phenotypes) occurred or did not occur.
- Local uploads to the public library of aggregate local outcomes -> a Learning Healthcare System that learns from every decision support event, whether or not recommendations were accepted by clinicians.

Why NHGRI, Why Now?

- Fits NHGRI strategic plan
  - Resource generating
  - Technology-advancing
  - Scientifically/medically relevant
  - A natural consortium-requiring opportunity
- Fits the complex, high volume, evolving science challenge of genomics
- Fits NHGRI trailblazer role
- Starting now with prototypes is important since full scale up and adoption will take many years


The need for person-specific decision support
in the era of genome-enabled healthcare
Facts per Decision
1000
100
10
Proteomics and other
effector molecules
Functional Genetics:
Gene expression
profiles
Structural Genetics:
e.g. SNPs, haplotypes
Decisions by
clinical
phenotype
i.e., traditional
health care
Human Cognitive
Capacity
1990
2000
2010
2020