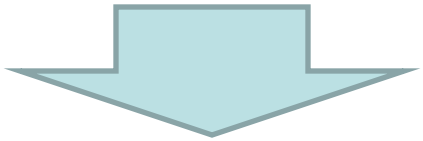
Challenges in Developing Healthcare IT Data Standards for Genomics

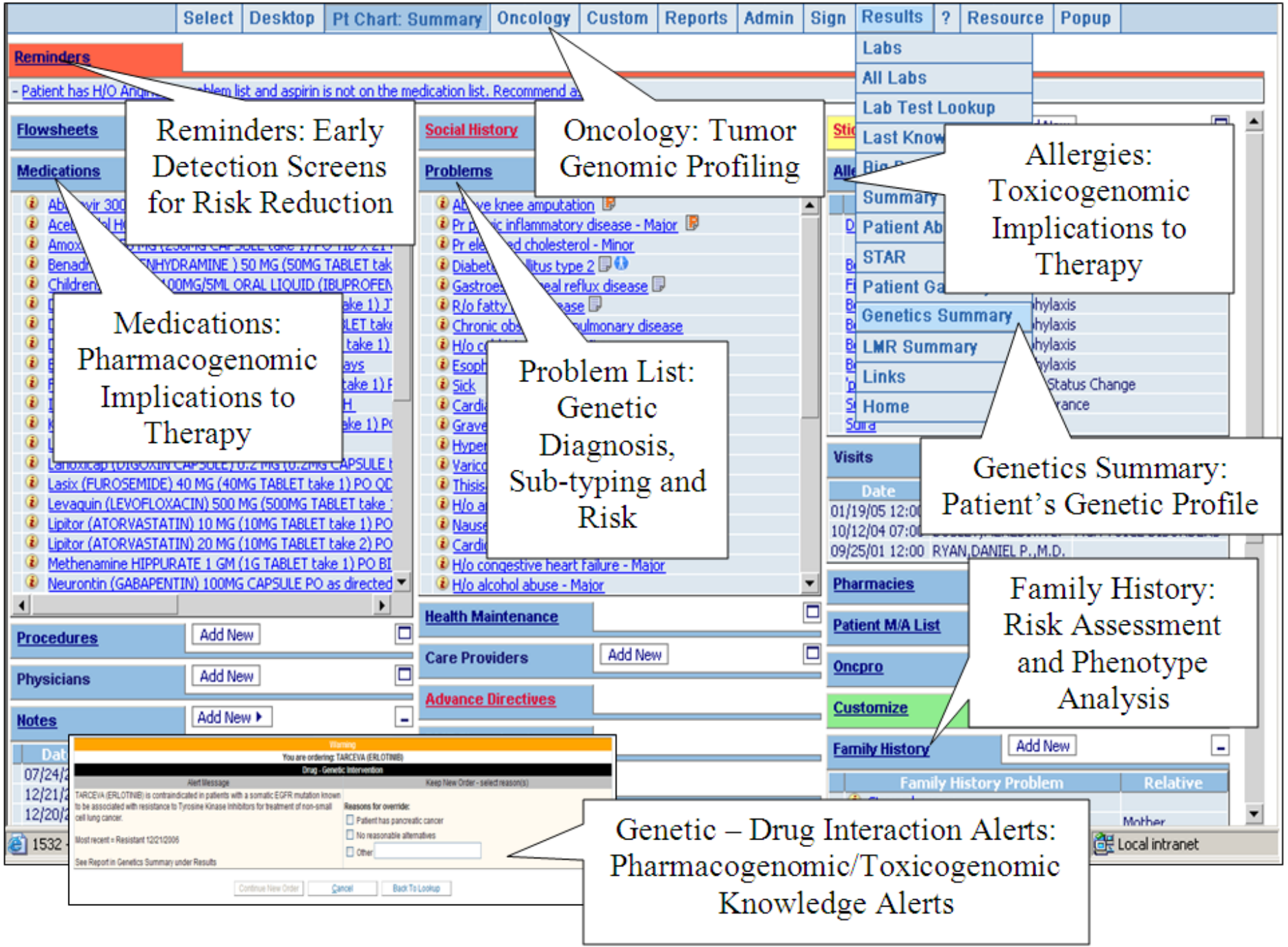
Mollie Ullman-Cullere
Dana-Farber Cancer Institute and Partners Healthcare
NHGRI Workshop
Genomics and HIT systems: Exploring the issues
April 27-28, 2011

Requirements for Genetic/Genomic Data in Clinical Care

- Accessible in the EHR as Structured Data
- Integrated into Clinical Workflows
- Available to Clinical Decision Support
- Integrated into Data Warehouses
  - patient panel management, outcomes analysis, quality assurance, reporting and discovery research
- Maintained Up-to-Date Interpretations

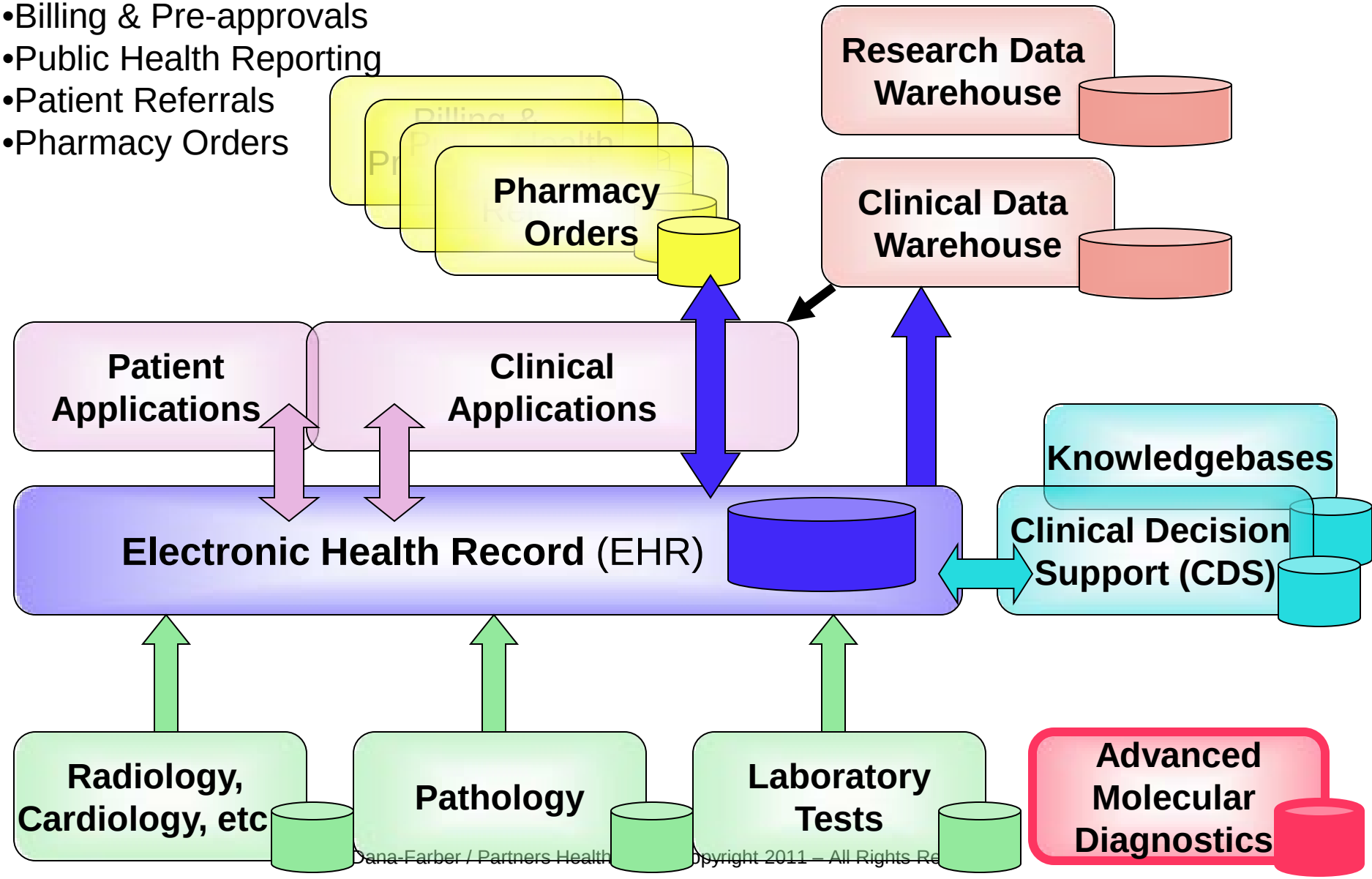


Function Like Other Laboratory Test Results


Select
Desktop
Pt Chart: Summary
Oncology
Custom
Reports
Admin
Sign
Results
?
Resource
Popup
Reminders
Labs
All Labs
Lab Test Lookup
Last Know
Summary
Patient Ab
STAR
Patient Ga
Genetics Summary
LMR Summary
Links
Home
Flowsheets
Medications
Social History
Problems
Visits
Date
Pharmacies
Patient M/A List
Oncpro
Customize
Family History
Add New
Family History Problem
Relative
Procedures
Physicians
Notes
Health Maintenance
Care Providers
Advance Directives
Reminders: Early Detection Screens for Risk Reduction
Oncology: Tumor Genomic Profiling
Allergies: Toxicogenomic Implications to Therapy
Medications: Pharmacogenomic Implications to Therapy
Problem List: Genetic Diagnosis, Sub-typing and Risk
Genetics Summary: Patient's Genetic Profile
Family History: Risk Assessment and Phenotype Analysis
Genetic – Drug Interaction Alerts: Pharmacogenomic/Toxicogenomic Knowledge Alerts

Electronic Health Record (EHR) – Components & Data Flow

External Communications:
- Billing & Pre-approvals
- Public Health Reporting
- Patient Referrals
- Pharmacy Orders



How do we <u>not</u> make this a picture of Genomic Medicine and Healthcare?

Butler, D. Translational Research: Crossing the Valley of Death. Nature 453, 840-842 (2008)

Current HIT Clinical Genomics Data Standards

Ensure transfer of data between systems …	Health Level Seven – HL7
Ensure standard description of tests, results, and interpretations …	LOINC, Logical Observation Identifiers Names and Codes HGVS Nomenclature, Human Genome Variation Society HGNC, Human Gene Nomenclature Committee RefSeq, Reference Sequences NCBI dbSNP. Single Nucleotide Polymorphism ISCN, International Society for Cytogenetics Nomenclature
Ensure standard context for interpretations (i.e. associations) …	SNOMED & RxNORM
Other References	LRG, OMIM, COSMIC, PubMed…

Portion of LOINC Panel for DNA Variant Details

LOINC Code	Name	Example value	source
51958-7	Transcript Reference Sequence Identifier	**NM_005228.3**	NCBI DB
48018-6	Gene identifier	**EGFR**	HGNC Nomenclature
48004-6	DNA Sequence Variation	**c.2573T>G**	HGVS Nomenclature
48003-8	DNA Sequence Variation identifier	**rs121434568**	NCBI dbSNP
48002-0	Genomic source class	**Somatic**, Likely Somatic, Unknown Origin, Likely Germline, Germline	LOINC Answer List
51961-1	Drug efficacy sequence variation interpretation	Resistant, **Responsive**, Presumed Resistant, Presumed Responsive, Unknown Significance, Benign, Presumed Benign, Presumed Non-Responsive	LOINC Answer List

Portion of HL7 Genetic Results v2 Message

Header

63-7^**Medication assessed**^LN||337525^**Erlotinib**^RxNorm|||||F|20080702100909|||||||||Laboratory for Molecular
22D1005307^^^CLIA&2.16.840.1.113883.4.7&ISO|1000 Laboratory Lane^Ste. 123^Cambridge^MA^99999^USA^B

02-0^**Genomic source class**^LN||LA6684-0^**Somatic**^LN|||||F|20080702100909|||||||||Laboratory for Molecular
Medicine^L^22D1005307^^^CLIA&2.16.840.1.113883.4.7&ISO|1000 Laboratory Lane^Ste. 123^Cambridge^MA^99999^USA^B

OBX|3|CWE|51964-5^**Drug efficacy analysis overall interpretation**^LN||^LA6677-4^**Responsive**^LN|||||F|2008070210
Molecular Medicine^L^22D1005307^^^CLIA&2.16.840.1.113883.4.7&ISO|1000 Laboratory Lane^Ste. 123^Cambridge^MA^99999^USA^B

OBX|4|FT|51969-4^**Genetic analysis summary repor**t^LN||\... **...<Report Text>**

Findings (repeat panel as needed)

5^Genetic analysis discrete result panel^LN|||20080702000000|||||||||||||20080703

OBR|4||PM-08-Q00228-3^HPCGG-LMM^2.16.840.1.113883.3.167.1^ISO|55208-3^DNA analysis discrete sequence variation panel^LN|||20080702000000|||||||||||2
Q00228-2&HPCGG-LMM&2.16.840.1.113883.3.167.1&ISO

OBX|1|CWE|48018-6^**Gene identifier**^LN||3236^**EGFR**^2.16.840.1.113883.6.281|||||F|20080702100909|||||||||Laboratory for Molecular
Medicine^L^22D1005307^^^CLIA&2.16.840.1.113883.4.7&ISO|1000 Laboratory Lane^Ste. 123^Cambridge^MA^99999^USA^B

OBX|3|CWE|51958-7^**Transcript reference sequence identifier**^LN||**NM_005228.3**^^2.16.840.1.113883.6.280|||||F|20080702100
Medicine^L^22D1005307^^^CLIA&2.16.840.1.113883.4.7&ISO|1000 Laboratory Lane^Ste. 123^Cambridge^MA^99999^USA^B

OBX|4|CWE|48003-8^**DNA sequence variation Identifier**^LN|1|**rs121434568**^^99HPCGG-LMM-MARKER|||||F|20080702100909|||||||||Labo
Medicine^L^22D1005307^^^CLIA&2.16.840.1.113883.4.7&ISO|1000 Laboratory Lane^Ste. 123^Cambridge^MA^99999^USA^B

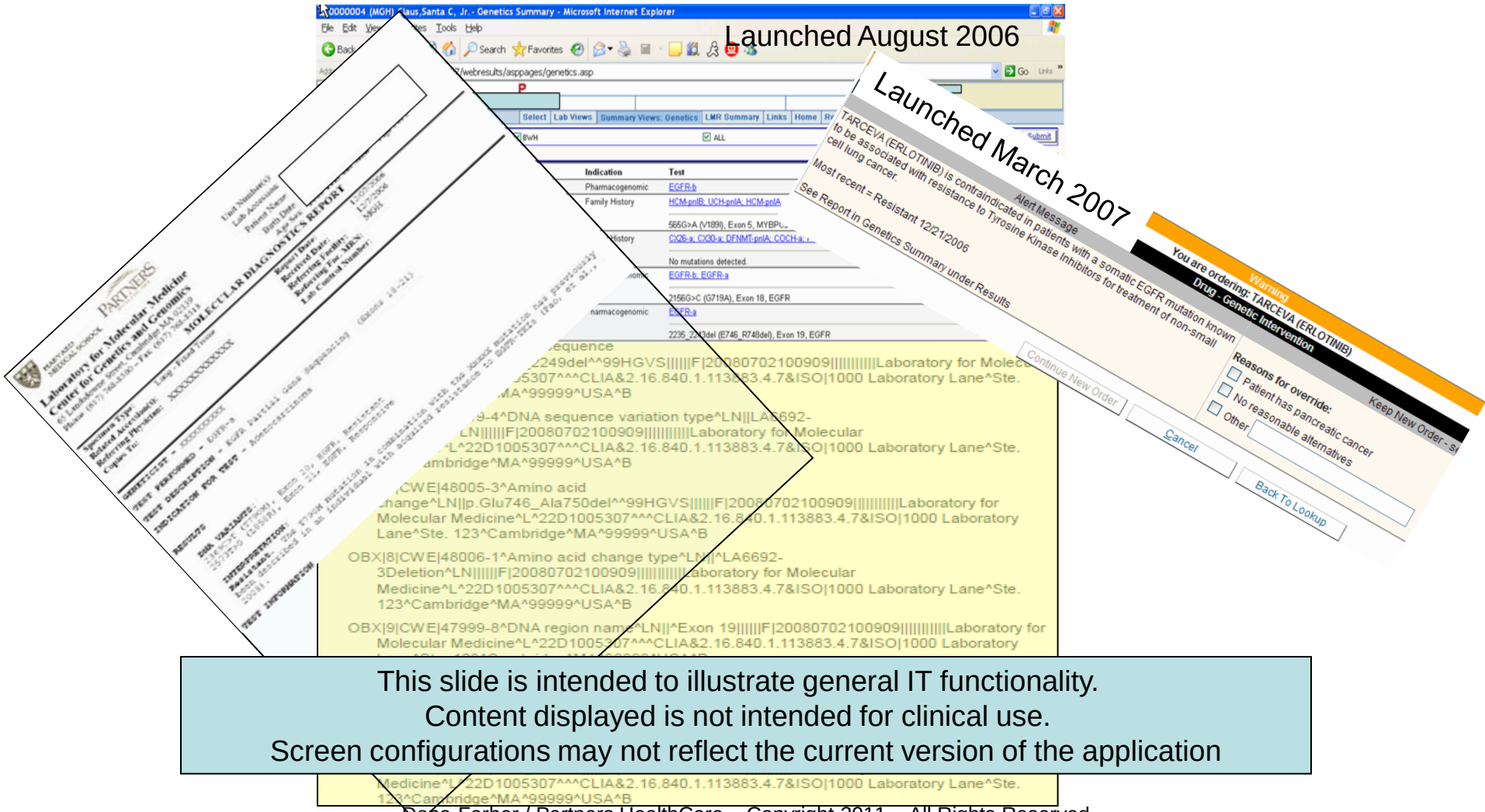
OBX|5|CWE|48004-6^**DNA sequence variation**^LN||**c.2573T>G**
^^2.16.840.1.113883.6.282|||||F|20080702100909|||||||||Laboratory for Molecular Medicine^L^22D1005307^^^CLIA&2.16.840.1.113883.4.7&ISO|1000 Laboratory Lane^Ste. 123^Cambridge^MA^9999
Exon 19|||||F|20080702100909|||||||||Laboratory for Molecular Medicine^L^22D1005307^^^CLIA&2.16.840.1.113883.4.7&ISO|1000 Laboratory Lane^Ste. 123^Cambridge^MA^99999^USA^B

OBX|10|CWE|48002-0^**Genomic source class**^LN||LA6684-0^**Somatic**^LN|||||F|20080702100909|||||||||Laboratory for Molecular
Medicine^L^22D1005307^^^CLIA&2.16.840.1.113883.4.7&ISO|1000 Laboratory Lane^Ste. 123^Cambridge^MA^99999^USA^B

OBX|11|ST|47998-0^**DNA sequence variation display name**^LN||
c.2573T>G (p.Leu858Arg), Exon 21, EGFR, Responsive|||||F|20080702100909|||||||||Laboratory for Molecular
Medicine^L^22D1005307^^^CLIA&2.16.840.1.113883.4.7&ISO|1000 Laboratory Lane^Ste. 123^Cambridge^MA^99999^USA^B

OBX|12|CWE|51961-1^**Drug efficacy sequence variation interpretation**^LN |1|LA6677-4^**Responsive**^LN|||||F|20
for Molecular Medicine^L^22D1005307^^^CLIA&2.16.840.1.113883.4.7&ISO|1000 Laboratory Lane^Ste. 123^Cambridge^MA^99999^USA^B

Co-Create Human Readable & Computer Readable Elements for the EHR



This slide is intended to illustrate general IT functionality.
Content displayed is not intended for clinical use.
Screen configurations may not reflect the current version of the application

HL7 Implementation Guides for Structuring Clinical Genetic Test Results within Established Standards

Published

- Gene Variants associated with Disease/Risk, Drug Metabolism and Drug Efficacy

In Ballot

- Cytogenetics and Array CGH
- Genetic Test Report *(alt. format for transmission of codified findings)*

Under Development

- Gene Variants for Tumor Profiling
- Expression Profiling

Piloting Organizations

- Healthcare Providers: Partners Healthcare, Dana-Farber/Brigham and Women's Cancer Center, and Intermountain Healthcare
- Laboratories: Laboratory for Molecular Medicine at Partners Healthcare, Dana-Farber/Brigham and Women's Cancer Center at Harvard Medical School, and ARUP Laboratories at University of Utah

What's Needed for Standards Development?

Roadmap, Pilot Projects, Tools, Collaborators, and Community Involvement

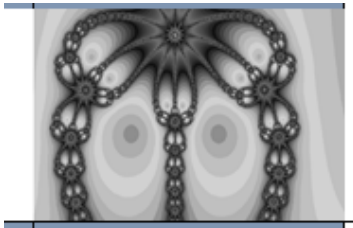
Healthcare IT (HIT) Standards Development Best Practices

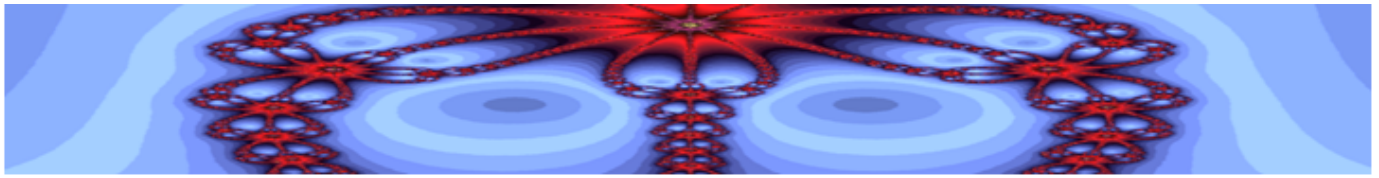
1. Participate in the healthcare standards communities (HIT and Genetics), as well as national initiatives
   - Participate in community review and publication (which is a separate track from journal publication).
   - Use tools for generation/translation/validation into standard representation e.g., HGVS's Mutalyzer tool: http://www.mutalyzer.nl/2.0/

2. Align standards development and real-world implementation projects
   - Does it enable useful functionality while supporting professional, legal and policy requirements?

3. Collaborate with key stakeholders
   - **Vocabulary/Message Standards:** NCBI, NLM, HGVS, HL7 and LOINC
   - **Practitioners:** Geneticists, clinicians, pathologists, and molecular diagnostic laboratories
   - **IT Professionals:** EHR, Clinical Research, Bioinformatics, and LIMS Developers

Mutalyzer – HGVS Nomenclature Name Generator/Checker/dbSNP Converter





- Home
- Name Checker
- Syntax Checker
- Position Converter
- SNP Converter
- Name Generator
- Batch Jobs
  - Name Checker
  - Syntax Checker
  - Position Converter
  - SNP Converter
- GenBank Uploader
- Webservices

- Help
- FAQ
- Exercise
- Disclaimer
- Feedback

Mutalyzer 2.0 β-8

released on 31 Jan 2011

HGVS nomenclature version 2.0

Welcome to the Mutalyzer web site

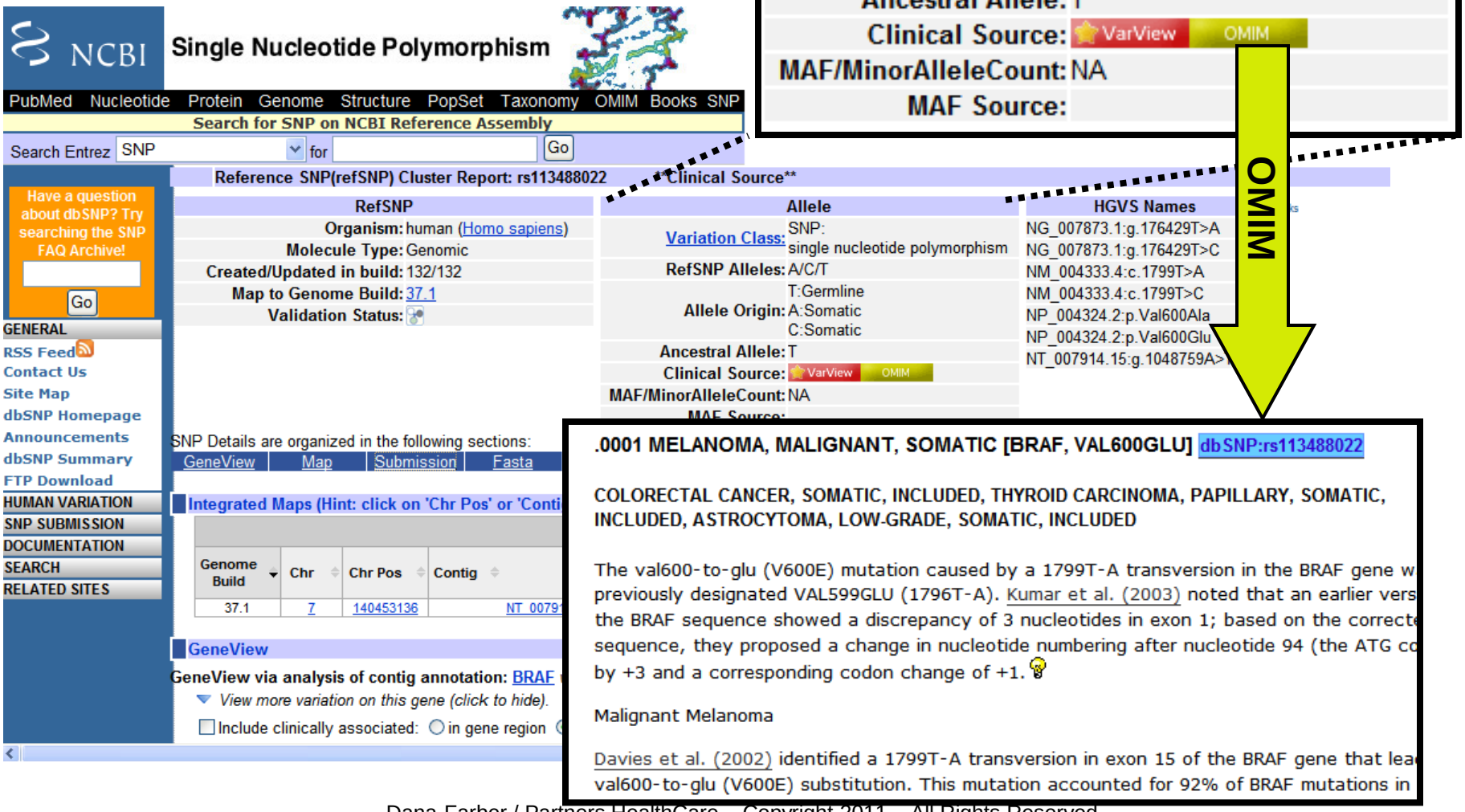
The aim of this program suite is to support checks of sequence variant nomenclature according to the guidelines of the Human Genome Variation Society.

Different interfaces are provided to collect the information necessary for the checks:

- The Name Checker takes the complete sequence variant description as input and checks whether it is correct.
- The Syntax Checker takes the complete sequence variant description as input and checks whether the syntax is correct.
- The Position Converter can convert chromosomal positions to transcript orientated positions and vice versa.
- The GenBank Uploader allows you to upload and use your own reference sequence.
- The SNP converter allows you to convert a dbSNP rsId to HGVS notation.
- The Webservices page provides instructions for the webservices.
- The Batch Checkers are interfaces that accept a list of inputs. These interfaces can be used for large quantities of checks.

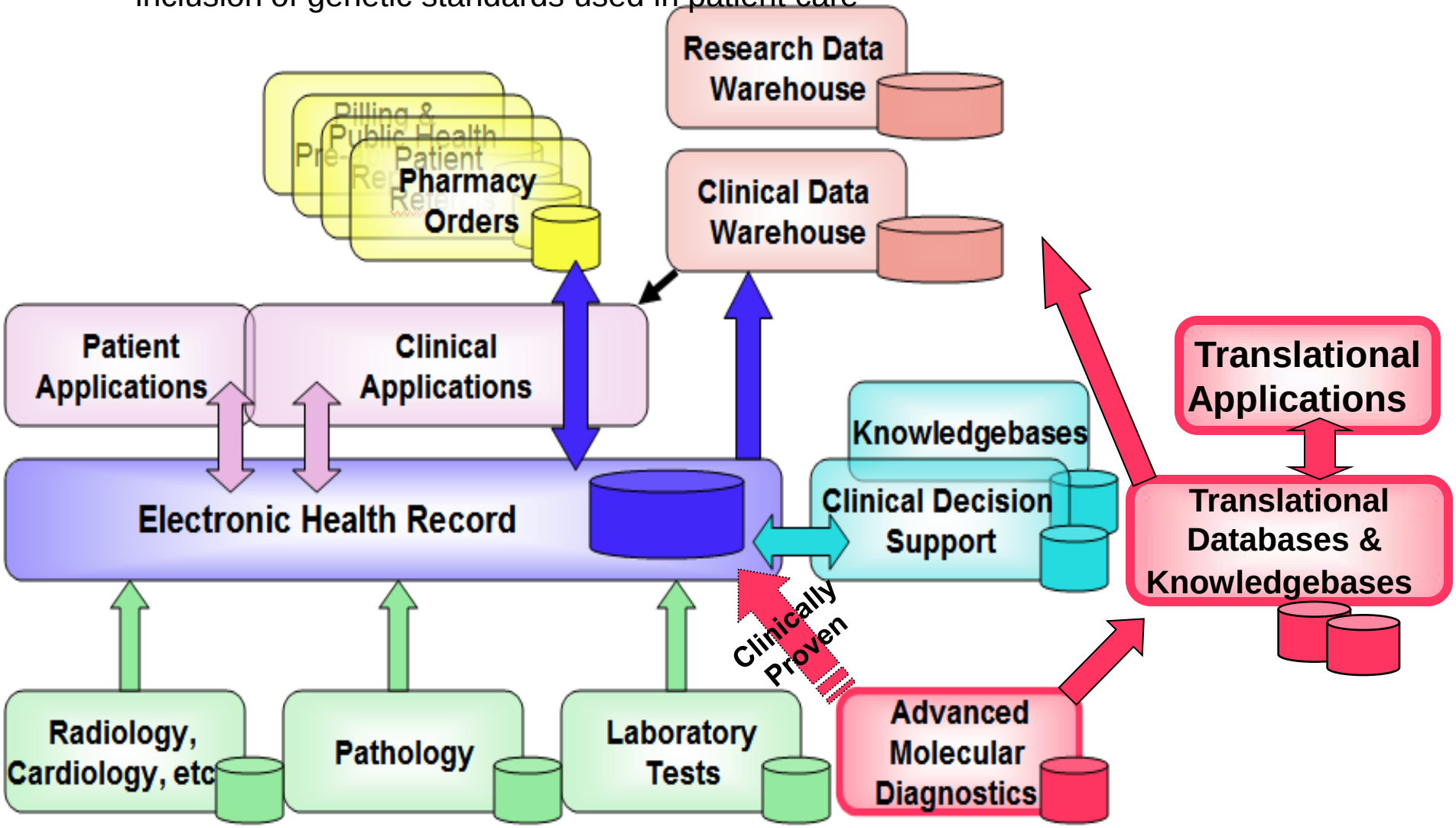
GenBank sequences are retrieved from the NCBI (Copyright and Disclaimers).
This project is sponsored by SUN Microsystems with server hardware within the scope of the Academic Excellence Grant (AEG) program (award EDUD-7832-080223-CNE).

NCBI's dbSNP – Enhanced for the Clinical Perspective



Translational Frameworks

- Adheres to HIT standards development best practices
- Focus on structured/codified data and terminologies extending HIT data standards for inclusion of genetic standards used in patient care



Recommendations to Extend Healthcare IT (HIT) Standards for Personalized Medicine

1. Define roadmap for parallel development of Electronic Health Records and Personalized Genomic Medicine

For example:

- Standard(s) for coding genetic based disease
- Standard(s) for representation of genetic data (for human and computer consumption)
- Minimal core data sets
- Standard(s) for representation of clinical associations
- Standard(s) for representation of clinical decision support rules

2. Fund tool development generating/translating/validating standard representation of data for Personalized Genomic Medicine (e.g. HGVS's Mutalyzer)

3. Make HIT resources easier to find by listing published standards and implementation guides within PubMed

4. Incorporate HIT standards into grants for genomic medicine