HL7 VERSION 3
DOMAIN ANALYSIS MODEL:
CLINICAL SEQUENCING, RELEASE 1
DRAFT
(1ST BALLOT FOR COMMENT)

ORU^R01

HL7 Version 3

Initial version: January 2013

Next Balloted Version: September 2013

Chapter Chair and Principal Author:	Mollie Ullman-Cullere Dana-Farber Cancer Institute and Partners HealthCare
Chapter Chair and Contributing Author:	Amnon Shabo IBM
Project Chair and Contributing Author:	Grant Wood Intermountain Healthcare
Contributing Author:	Kevin Hughes, MD Massachusetts General Hospital
Contributing Author:	Daryl Thomas Life Technologies
Contributing Author	Larry Babb Partners Healthcare Center for Personalized Genetic Medicine
Contributing Author	Lynn Bry, MD Brigham and Women's Hospital
Seeking Additional Co-Authors/Participants	

This document has been updated to reflect January 2013 HL7 ballot comments. The HL7 Clinical Genomics Work Group is actively seeking comments/feedback from the genetics/genomics community. Please contact Mollie Ullman-Cullere at: mullmancullere@partners.org, if you are interested in participating.

TABLE OF CONTENTS

1. INTRODUCTION **1**

1.1 Purpose 1

1.2 AudIence 1

1.3 Scope 1

1.4 Assumptions 1

1.5 Conventions 2

1.6 IMPLMENTORS 2

2. USE CASE STAKEHOLDERS **3**

3. ISSUES AND OBSTACLES **4**

4. PERSPECTIVE **4**

5. USE CASE SCENARIOS **5**

5.1 Scenario 1: Specimen Identification 5

5.1.1 Germline testing for biomarkers/mutations (usually inherited) *5*

5.1.2 Tumor testing for somatic (tumor specific biomarkers/mutations) *5*

5.1.3 Pediatric testing for biomarkers/mutations causal to rare early childhood conditions *6*

5.1.4 Prenatal testing which may be reported on the maternal medical record (and should be identified as separate from germline testing) *6*

5.1.5 Infectious disease testing, where the biomarker/mutation identified within the disease causing organism is reported into the patient medical record following similar data standards as used for other testing scenarios above. *6*

5.2 Scenario 2: Clinical Sequencing – Germline Testing 7

5.2.1 Description of Scenario (following numbers in the diagram above) *8*

5.2.2 Alternative Flow 1: Chart Review *9*

5.2.3 Alternative Flow 2: New Genetic Knowledge *9*

5.2.4 Alternative Flow 3: New Clinical Indication *9*

5.3 Scenario 3: Cancer Profiling – Somatic Testing 10

5.3.1 Description of Scenario Differences from Germline Workflow *10*

5.4 Scenario 4: Decision Making Tools – Family History and Drug Dosage Calculators 11

5.4.1 Description of Scenario *11*

5.5 Scenario 5: Public Health Reporting 12

5.5.1 Description of Scenario *12*

5.5.2 Cancer Registry workflow *13*

5.5.3 FDA Scenarios in Public Health Reporting *13*

© 2013 Health Level Seven International. All rights reserved.

5.6 National Marrow Donor Program 13
5.7 Scenario 6: Clinical and Research Data Warehouses 13
5.7.1 Description of Scenario 13
6. VOCABULARY CONSTRAINTS 15
6.1 Genetic Tests, Testing Context, Interpretation Code, and Genetic Data 15
6.1.1 LOINC 15
6.2 Associated Disease and/or Drug 16
6.2.1 SNOMED-CT 16
6.2.2 RxNORM 16
6.3 Genes 16
6.3.1 HGNC gene symbols (required) 16
6.4 Sequence Variations 17
6.4.1 HGVS (required) 17
6.4.2 dbSNP (optional, but highly recommended) 17
6.4.3 COSMIC (optional) 18
6.5 Structural variants 18
6.5.1 <Add ISCN> 18
6.5.2 <Add dbVAR> 18
6.6 Copy number change 18
6.6.1 <Need Nomenclature> 18
6.7 Biomarkers --> Is this far enough along to add 18
6.7.1 <Add MedGen/LOINC> 18
6.8 Reference Sequences (required) 18
6.8.1 RefSeq 19
6.8.2 LRG 19
6.9 Publicly Available References (valuable for clinical and translational genomics) 20
6.9.1 OMIM (optional) 20
6.9.2 PubMed (optional) 20
6.9.3 PharmGKB (optional) 21
6.9.4 ClinicalTrials.gov (optional) 21
7. LOGICAL MESSAGE TYPES 22
7.1 INTRODUCTION AND STRATEGY 22
7.2 MESSAGE DEFINITIONS 23
7.3 Message Components 23

 © 2013 Health Level Seven International. All rights reserved.

7.3.1 Test Interpretation *24*
7.3.1.0 Genetic Analysis Master Panel 24
7.3.1.1 Genetic Analysis Summary Panel 24
7.3.2 Test Definition *26*
7.3.2.0 DNA Region Analysis Test Coverage Panel 26
7.3.2.1 DNA Region of Interest Panel 27
7.3.3 Findings *28*
7.3.3.0 Genetic Analysis Discrete Result Panel 28
7.3.3.1 DNA Analysis Discrete Sequence Variation Panel 29
7.4 LOINC Codes 36
7.5 LoINC Answer Lists 38
7.6 SPECIAL SYNTAX and notes on Use Of loinc terms 39
7.6.1 DNA Sequence Variation (see table 7-5) *39*
7.6.2 Genomic Source Class *40*
7.6.3 Genetic Variant Assessment *40*
8. DATA SET CONSIDERATIONS AND HIT DATA STANDARDS 41
8.1 fAMILY HISTORY 41
8.2 Sequence Variations / Chromosomal change 41
8.2.1 Small Genetic Variations *41*
8.2.2 Structural Variations *41*
8.2.3 Copy Number Change *41*
9. HL7 ENCAPSULATION OF GENOMIC DATA FILES 41
10. CLINICAL GRADE-GENOMIC DATA FILE STANDARDS 42
11. GAPS & EXTENSIONS 43
11.1 Laboratory order entry 43
12. OUTSTANDING QUESTIONS 43
13. GLOSSARY 43
14. FUTURE PLANS 43
14.1 Laboratory genomic data standards 43
14.2 Extension of sequence VARIATION AND cytogenetic HL7 models 43
14.3 Extension to Specimen scenarios 43
14.3.1 Microbiome analysis of the patient *43*

© 2013 Health Level Seven International. All rights reserved.

1. Introduction

In March, 2008, the United States Department of Health and Human Services, Office of the National Coordinator for Health IT published the *Personalized Healthcare Detailed Use Case* (Click here to see the use case) in response to a request and specifications from the American Health Information Community. The use case focuses on supporting secure access to electronic genetic laboratory results and interpretations for clinical care, as well as family history and associated risk assessments by authorized parties and is driven by the need for timely electronic access to ordered, referred and historical genetic laboratory results and family history. Ordering clinicians receive genetic laboratory test results as a response to an order by having the genetic test results sent either directly to the clinician's EHR system (local or remote) or to another clinical data system in support of the provisioning of historical results.

1.1 PURPOSE

At the time of writing the 2008 Personalized Healthcare Use Case, single gene tests were the norm and genomic sequencing was not specifically addressed. The *HL7 Version 3 Domain Information Model: Clinical Sequencing, Release 1* extends the Personalized Healthcare Use Case with lessons learned from implementations, as well as technological advancement.

The current version of this document is meant to gather early comments for iteration and extension of future releases.

1.2 AUDIENCE

This guide is designed to be used by analysts and developers who require guidance on incorporation of genomic data in the clinical and clinical research healthcare IT environment. In addition, developers of genomic and healthcare IT data standards may use this guide to extend these standards for support of clinical sequencing. Users of this guide must be familiar with the details of HL7 message construction and processing. This guide is not intended to be a tutorial on that subject.

1.3 SCOPE

This domain information model details a variety of use case scenarios key to personalized genomic medicine and translational research, including more typical scenario for testing of a person's inherited or germline genome, cancer genomics/tumor profiling, early childhood developmental delay, neonatal testing, and newborn screening. In addition, the use case includes two scenarios where test results are manually translated from reports into either a tool for clinical decision making (e.g. family history or drug dosage calculator) or for public health reporting for cancer registries.

1.4 ASSUMPTIONS

Assumptions are summarized as follows:

- Infrastructure is in place to allow accurate information exchange between information systems.
- Providers access laboratory test results through either an EHR or a clinical information system.
- Privacy and security has been implemented at an acceptable level.
- All participants agree to all standards, methodologies, consent, privacy and security.
- Legal and governance issues regarding data access authorizations, data ownership and data use are outside the scope of this document.
- The order, paper or electronic, associated with the laboratory result contains sufficient information for the laboratory to construct the laboratory result message properly.

© 2013 Health Level Seven International. All rights reserved.

1.5 CONVENTIONS

This document is based on conventions used within the 2008 ONC Personalized Healthcare Use Case (click here to view), because to date it has been valuable in articulating a unified vision for which standards have been successfully created and piloted, over a diverse stakeholder group. However, the use case needs to be updated with lessons learned, technological advances, and progress in the field.

1.6 IMPLMENTORS

Since the 2008 publication of the Personalized Healthcare Use Case, several laboratories and providers have piloted the HL7 standards supporting the described functionality.

GENETIC TESTING:
Genetic Testing Laboratories:
Laboratory for Molecular Medicine, Partners HealthCare Center for Personalized Genetic Medicine (formerly Harvard – Partners Center for Genetics and Genomics), Cambridge, MA
ARUP, University of Utah, Salt Lake City UT
Center for Advanced Molecular Diagnostics, Brigham and Women's Hospital, Boston MA
Center for Cancer Genomic Discovery, Dana-Farber Cancer Institute, Boston MA

Receiving Provider Electronic Medical Records:
Partners Healthcare, Boston, MA
Intermountain Healthcare, Salt Lake City, UT

Systems for Discovery Research, including results viewer and research data warehouse:
Dana-Farber Cancer Institute, Boston MA

FAMILY HISTORY / PEDIGREE:
Adaptors of the following open source software, including a significant number of clinical settings and research initiatives.
Hughes Risk Apps – an open source family history, pedigree and risk analysis software product (http://www.hughesriskapps.net/)
My Family Health Portrait – the US Surgeon General's open source family history software (https://familyhistory.hhs.gov/fhh-web/home.action)

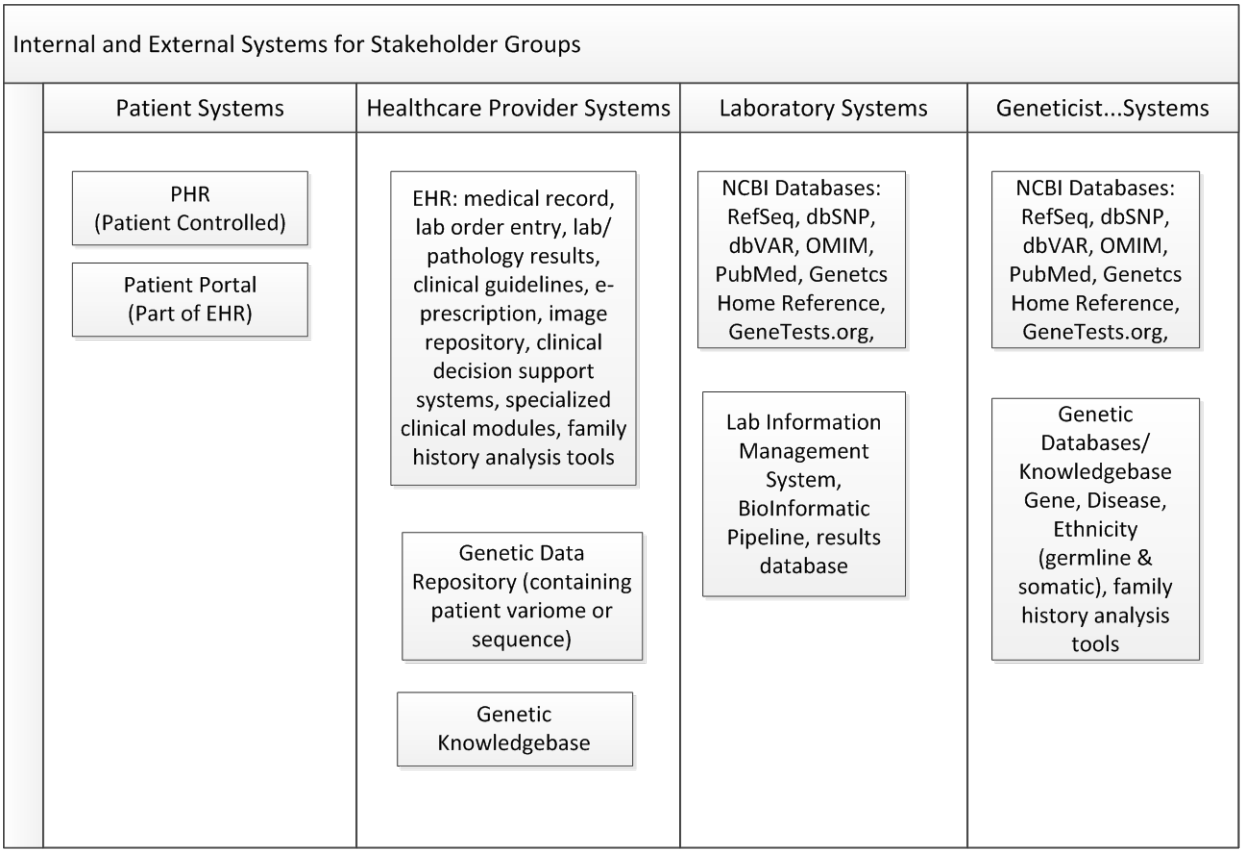
© 2013 Health Level Seven International. All rights reserved.

2. Use Case Stakeholders

Stakeholder	Contextual Description
Anatomic & Surgical Pathology	For cancer profiling (i.e. genetic testing of cancer specimens), the pathologic diagnosis will play a key role in testing and interpretation of the findings.
Geneticist / Medical Geneticist / Molecular Pathologist	Professionals interpreting the clinical implications of a patient's genetic data. These professionals may work within the laboratory setting or outside the laboratory.
Healthcare Entities	Organizations delivering healthcare.
Healthcare Payors	Healthcare Insurers and Centers for Medicare & Medicaid Services
Information Technology Vendors	Vendors supplying information technology solutions and support.
Laboratories - Reference	Testing laboratories outside the hospital environment either as a separate corporate entity or separate unit of the same organization.
Laboratories - Hospital	Testing laboratory which is part of the hospital entity and hospital laboratories.
Manufacturers/Distributors	Entities involved in the development, production, and distribution of products used in healthcare (e.g. *in vitro* diagnostic tests)
Patients	Members of the public that use healthcare services.
Public Health Agencies	Agencies which help to protect and improve health and healthcare of the public.
Registries	Systems for the collection, analysis, and distribution of data for the improvement of public health.

<need to expand>

© 2013 Health Level Seven International. All rights reserved.



3.Issues and Obstacles

Numerous challenges exist in the area of policy, patient and clinician education, and reimbursement, which are beyond the scope of this document, unless requiring consideration within the information technology solutions (e.g. clinical decision support). Key challenges for information technology include: data security, adoption of electronic health records and laboratory information management systems, and interoperability, and structuring of useful data. This document informs information technology vendors of key functionality for clinical sequencing, and outlines considerations for healthcare providers and laboratories investing in information technology.

4.Perspective

This document includes perspectives of stakeholder groups outlined in section 2. Integration of molecular diagnostics into the clinical workflow is key for safe, efficient and effective adoption. For instance, the potential for medical error during drug order entry is reduced with clinical decision support which alerts the clinician, if ordering a drug which is contraindicated. Developing systems which are capable of consideration of genetic markers associated with drug metabolism, efficacy, and toxicity during the order entry process will reduce medial error, as our knowledge increases.

© 2013 Health Level Seven International. All rights reserved.

5.Use Case Scenarios

5.1 SCENARIO 1: SPECIMEN IDENTIFICATION

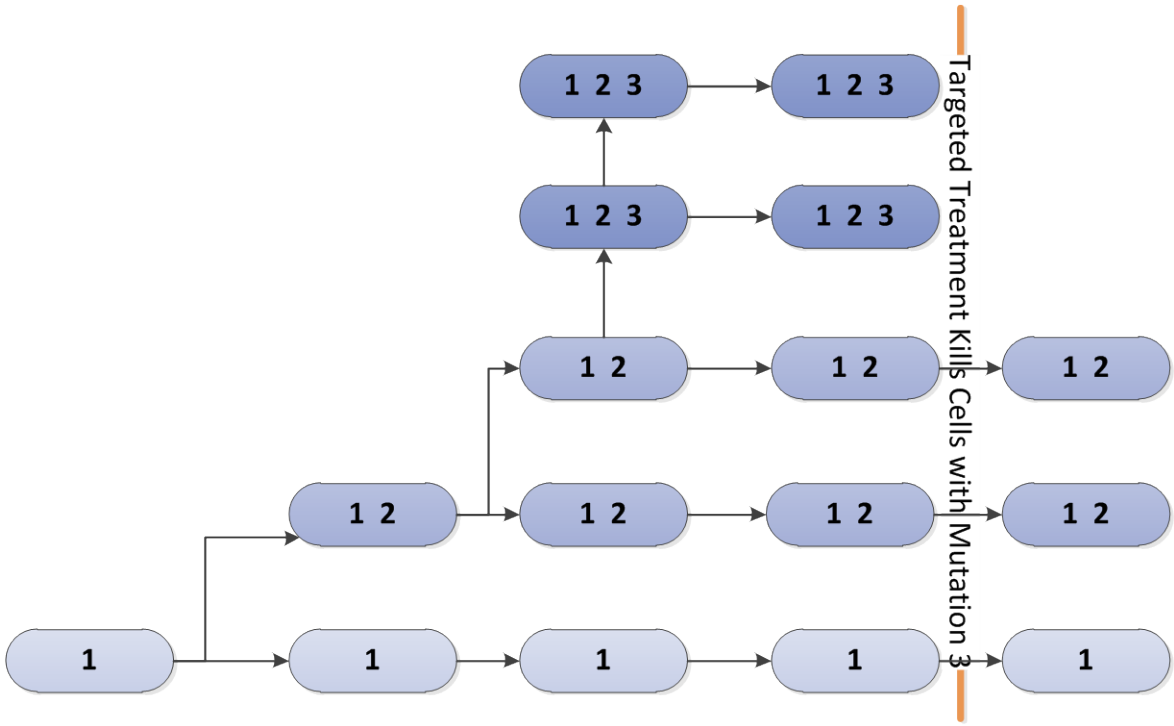
Use Cases for sequencing require explicate identification of 1 or more specimens to be used in laboratory analysis. This likely requires the identification of specimen groups (i.e. separate specimens and associated derivatives) originating from the same patient/subject or related patients/subjects.

5.1.1 Germline testing for biomarkers/mutations (usually inherited)

In terms of specimen identification, this is the most straightforward scenario. Typically a blood sample or cheek swab will be taken from the patient and DNA extracted. Except for low level heterogeneity, the genome/variome/mutations identified in this specimen are ubiquitously present throughout every cell in the patient and are inherited from their mother and father (except in the case of spontaneous mutations). This specimen is not limited in quantity, like a tumor specimen, because the laboratory may request an additional sample.

5.1.2 Tumor testing for somatic (tumor specific biomarkers/mutations)

To identify somatic (i.e. acquired) mutations within a cancer specimen, in general a laboratory will analyze both a germline specimen and somatic specimen. The somatic/cancer specimen contains both germline sequence and mutations as well as the somatic mutations present in cancer. To definitively classify a mutation as somatic the laboratory compares the two sequences and to identify mutations unique to the cancer. Note this can be a complicated process, because cancer cells acquire mutations throughout their lifespan and pass them on to daughter cells.



Simplified representation of cancer cells acquiring mutations or sequence variants, represented as numbers 1 2 and 3, in dividing cancer cells. Note targeted therapy can kill a specific population of cancer cells.

© 2013 Health Level Seven International. All rights reserved.

Changes in the population of cells with particular mutations will change overtime as well as in conjunction with events such as therapy. For instance, targeted chemotherapy may kill a specific population of cancer cells with specific mutations and other cancer cell populations may survive and continue to divide. Therefore, clearly annotating these specimens as somatic and capturing annotations related to a time relevant to a treatment timeline may be critical for analysis.

In some scenarios, a laboratory may focus sequence analysis on well studied genes/mutations identified only in cancer. Commonly these mutations are only found in cancer, because they cause extreme behavioral changes at the cellular level (e.g. uncontrolled cell division), which would result in embryonic death if present in the embryo. Specimens, sequence, and identified variants/mutations from these studies should be clearly annotated as somatic.

Summary

a. Matched specimens for germline and somatic analysis, where comparison will result in the identification of tumor specific mutations/biomarkers
b. Tumor specimen without a matched germline specimen, where mutations/biomarkers are believed to be specific to tumors.

5.1.3 Pediatric testing for biomarkers/mutations causal to rare early childhood conditions

a. Matched specimens of patient and maternal and paternal specimens, where comparison aids in identification of original biomarkers/mutations within the patient

5.1.4 Prenatal testing which may be reported on the maternal medical record (and should be identified as separate from germline testing)

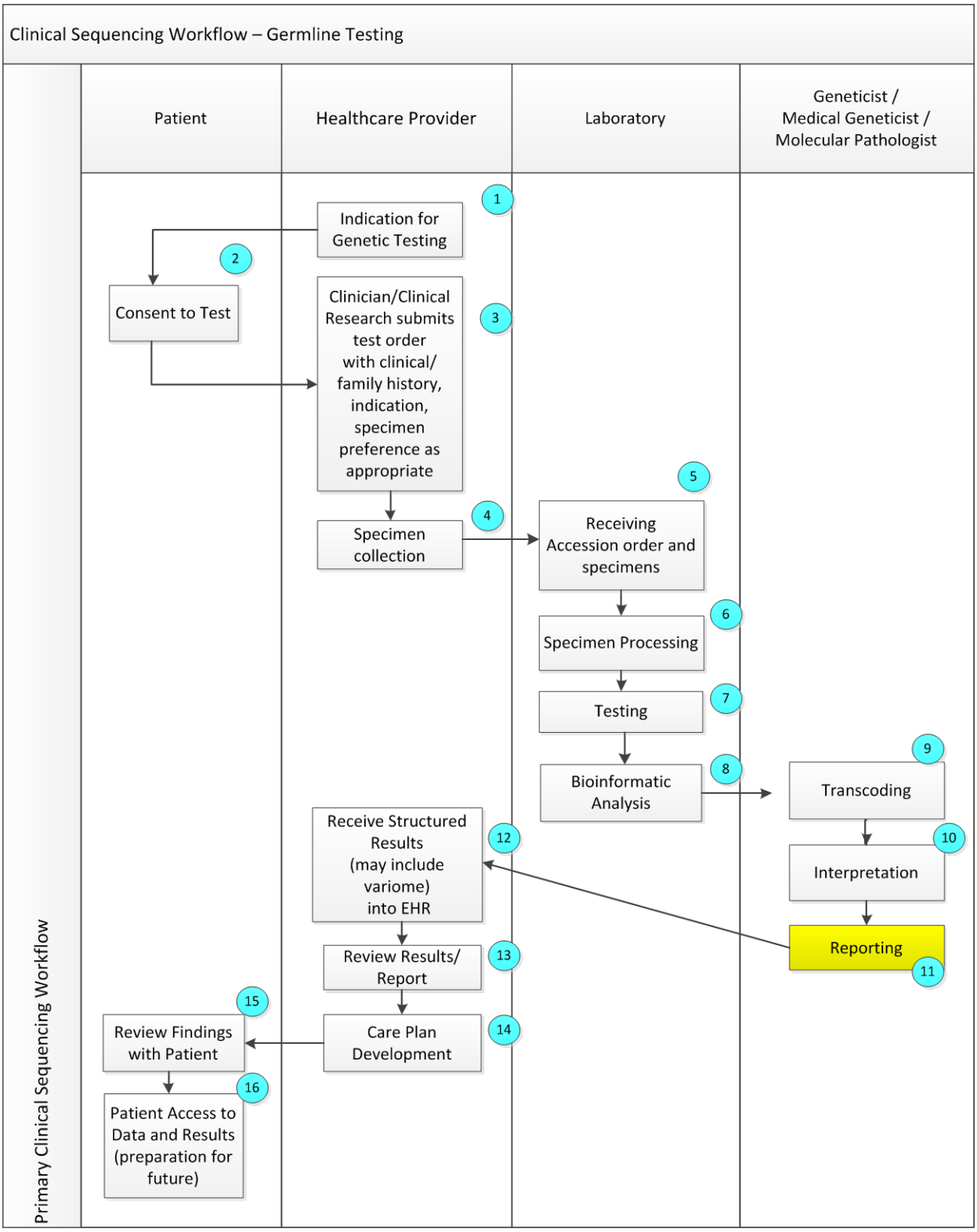
a. Often have matched prenatal/fetal and maternal specimens for analysis

5.1.5 Infectious disease testing, where the biomarker/mutation identified within the disease causing organism is reported into the patient medical record following similar data standards as used for other testing scenarios above.

Derivatives which may be analyzed from the above testing scenarios include: DNA, RNA, and Protein

© 2013 Health Level Seven International. All rights reserved.

5.2 SCENARIO 2: CLINICAL SEQUENCING – GERMLINE TESTING

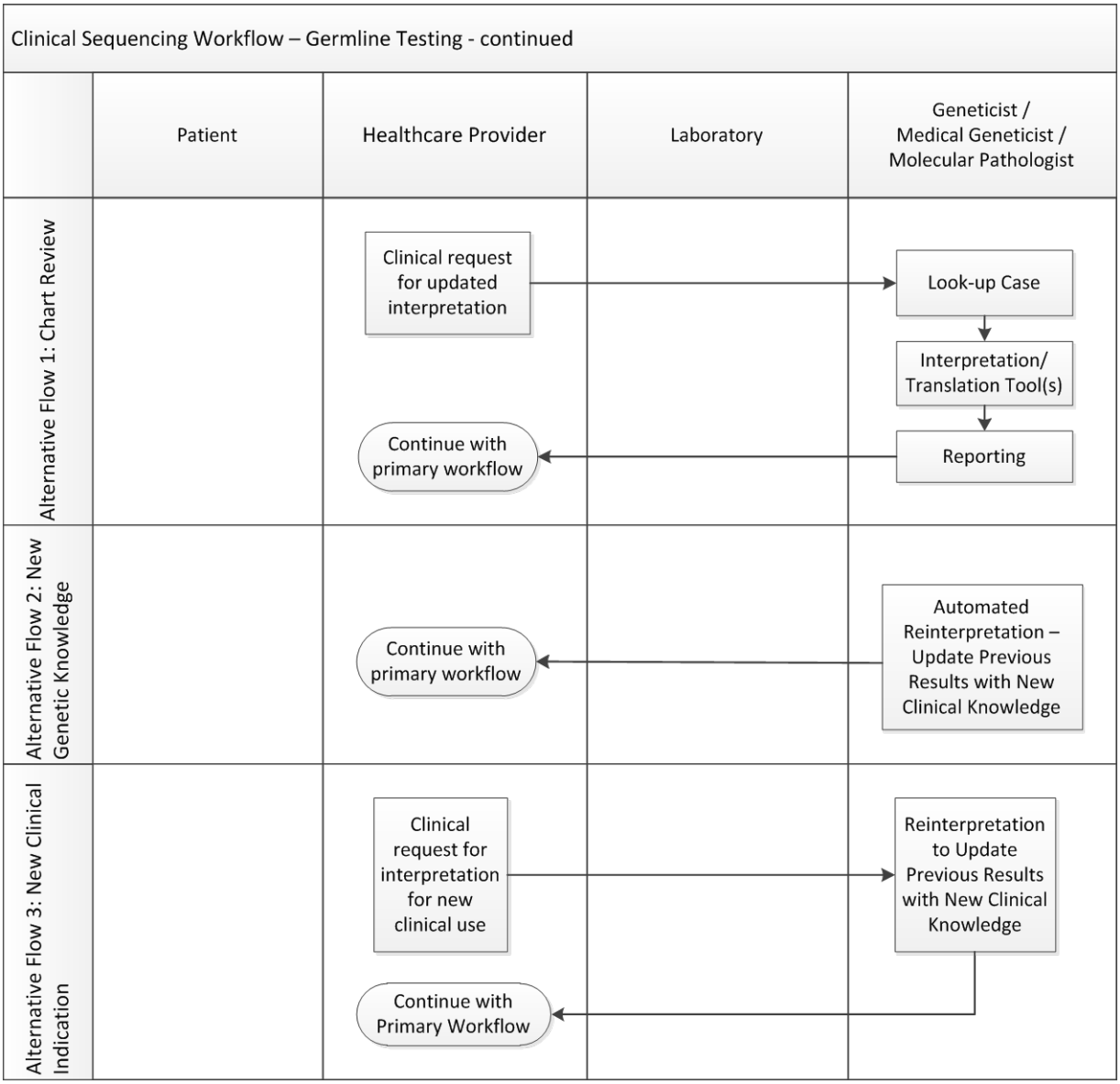


© 2013 Health Level Seven International. All rights reserved.

5.2.1 Description of Scenario (following numbers in the diagram above)

1. Clinician determines that a genetic test is needed to inform patient care decisions. Often this includes family history based risk assessment.
2. Clinician obtains patient consent for testing.
3. Order entry for genetic testing, including relevant data to aid in evaluation and interpretation of findings: indication for testing, family history, and relevant clinical data for the patient.
4. Blood is drawn or cheek swabbed for cells containing DNA
5. Laboratory receives the order and specimen(s) for testing
6. Specimens are processed (e.g. DNA extracted) and prepared to be loaded on the sequencing instrument.
7. Specimens are sequenced.
8. Data from the instrument passes through a bioinformatics pipeline for data processing: alignment and identification of sequence variants, as well as quality assurance
9. During the **'Transcoding'** process, genetic data is transformed from bioinformatic format into healthcare IT data standards.
10. Genetic results are interpreted for clinical implications
11. Genetic report is created, including narrative findings and interpretation as well as the equivalent information structured in machine readable formats using interoperable healthcare IT data standards.
12. Genetic report and structured results are received in the Electronic Health Record system (EHR)
13. Clinician reviews the results/report
14. Clinician develops (or modifies) a care plan taking into consideration the genetic findings
15. Clinician reviews the genetic findings and care plan with the patient
16. Genetic results are made available to the patient in the web-based patient portal

© 2013 Health Level Seven International. All rights reserved.



5.2.2 Alternative Flow 1: Chart Review

If a sequence variant (i.e. mutation) of 'Unknown Significance' were identified in a patient or the clinical implications of an identified variant are suspected of change, then the clinician may contact the testing laboratory prior to a follow-up patient appointment (e.g. annual exam).

5.2.3 Alternative Flow 2: New Genetic Knowledge

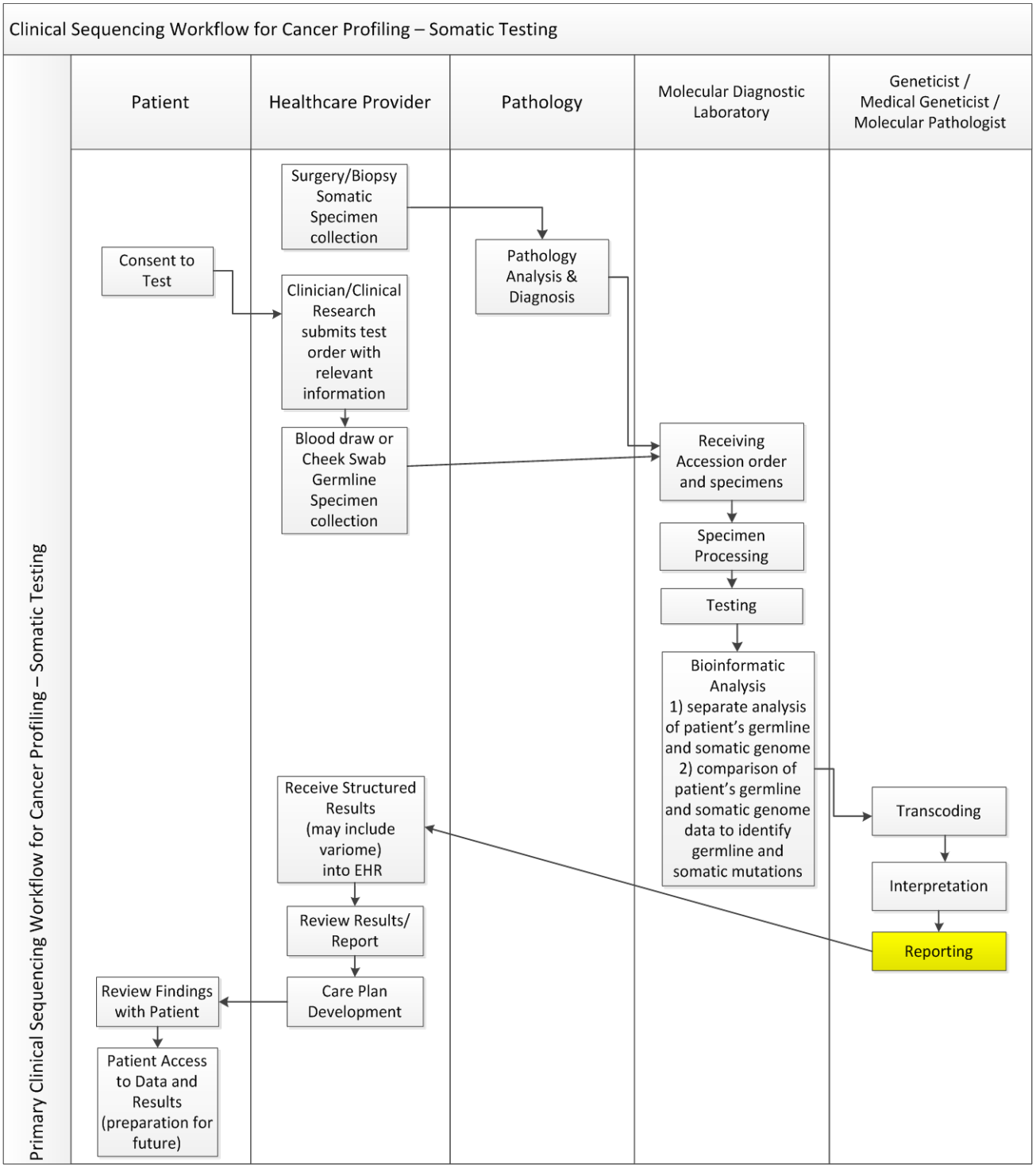
A testing laboratory may contact the ordering clinician, if the clinical implications of a sequence variant (i.e. mutation), previously identified in the patient, have changed.

5.2.4 Alternative Flow 3: New Clinical Indication

If genetic data from previous testing may inform a new clinical decision, the clinician may contact the laboratory for a new interpretation of existing data. As confidence in data quality increases and size of data sets increases, alternative flow may become more common.

© 2013 Health Level Seven International. All rights reserved.

5.3 SCENARIO 3: CANCER PROFILING – SOMATIC TESTING



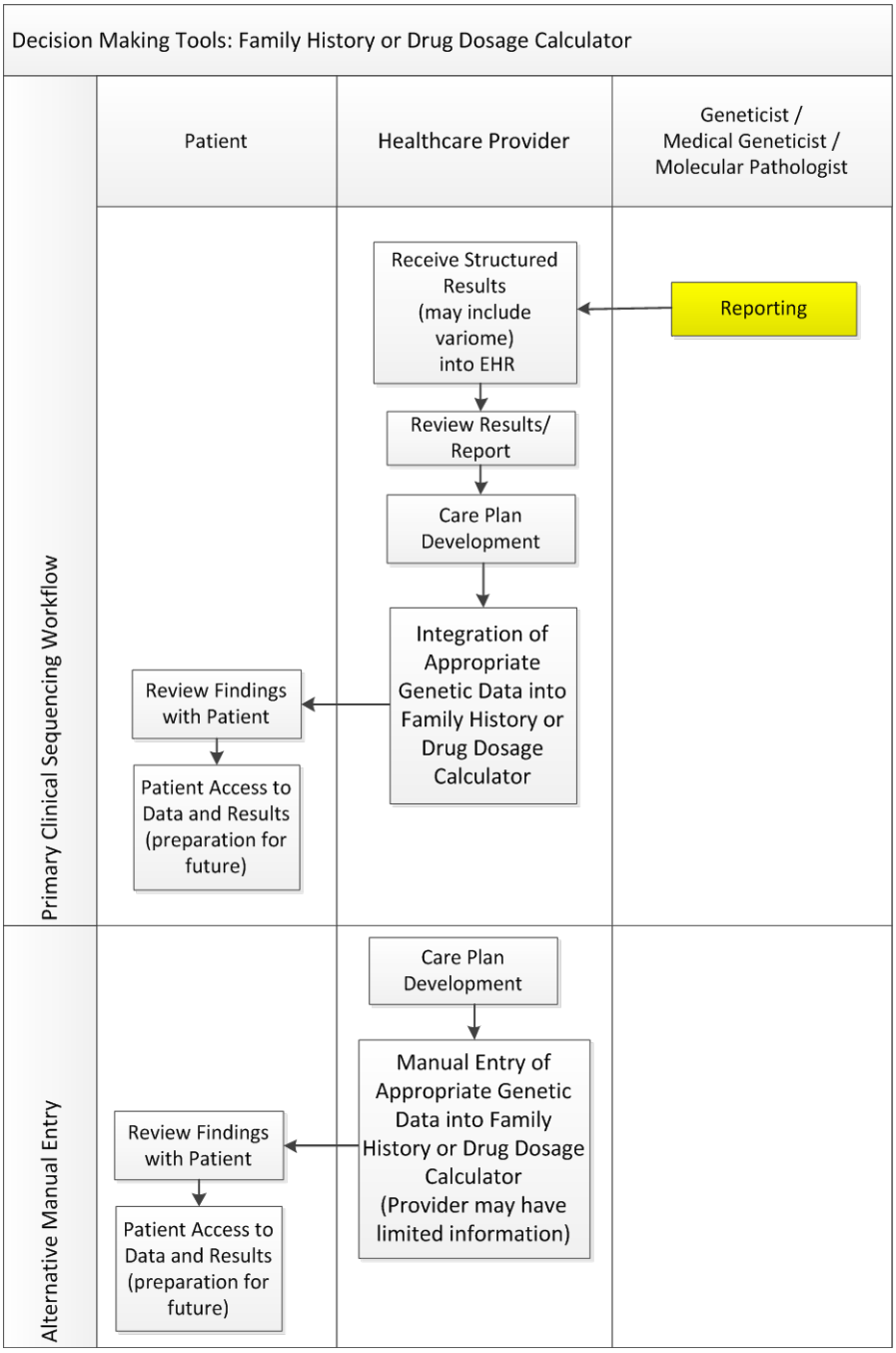
5.3.1 Description of Scenario Differences from Germline Workflow

In cancer profiling, pathology plays a key role. For instance, the same mutation identified in different cancers has different clinical implications. In addition, ideally clinical sequencing will include analysis of

© 2013 Health Level Seven International. All rights reserved.

both a germline specimen and a cancer specimen, so that cancer specific mutations can be identified with more certainty. For more information on specimens within this workflow, see section 5.1.2.

5.4 SCENARIO 4: DECISION MAKING TOOLS – FAMILY HISTORY AND DRUG DOSAGE CALCULATORS

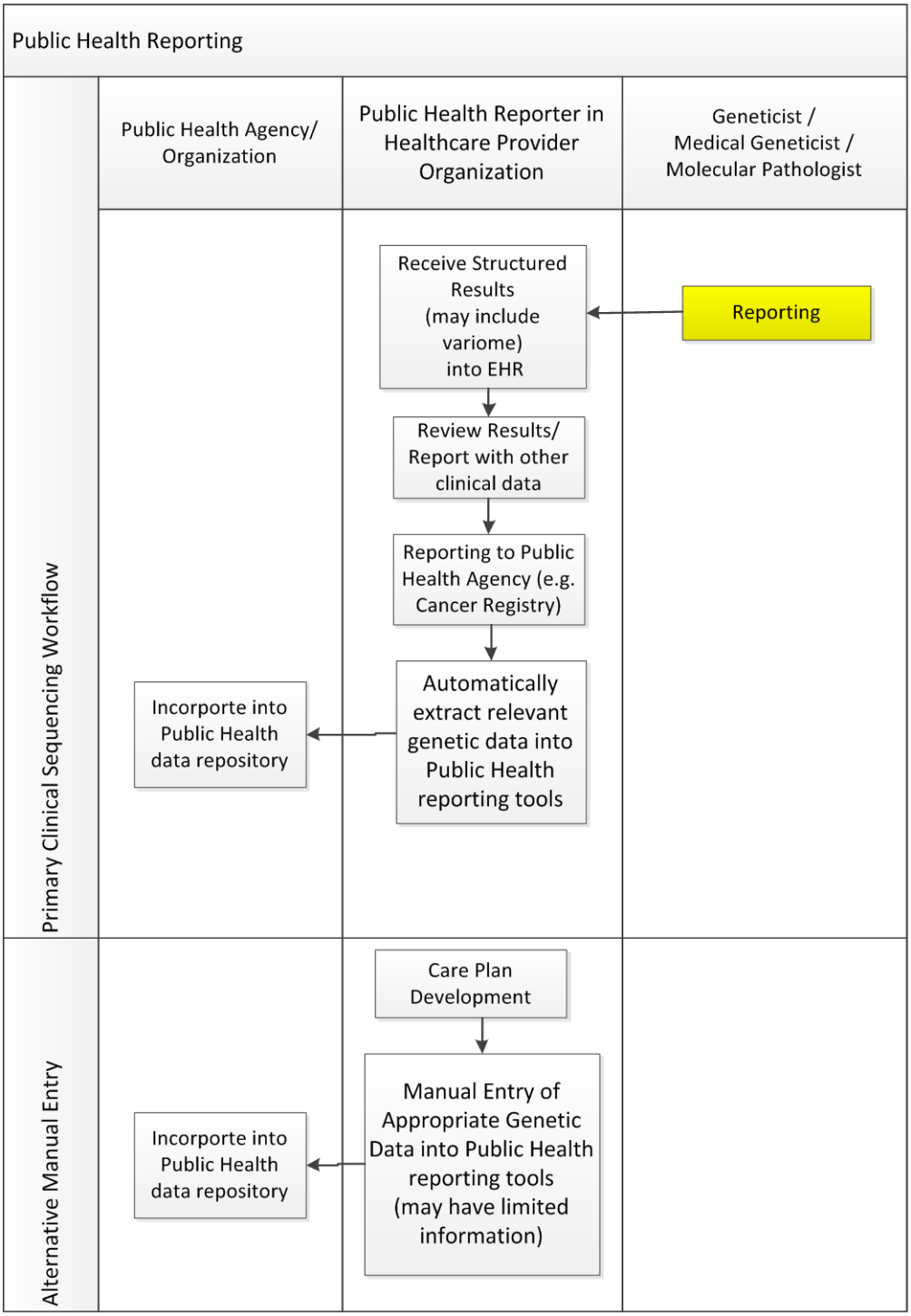


5.4.1 Description of Scenario

Today clinicians translate (i.e. manually reenter) genetic data into tools for decision making. This includes family history tools and drug dosage calculators. In the future, this data will automatically be incorporated into clinical decision making tools.

© 2013 Health Level Seven International. All rights reserved.

5.5 SCENARIO 5: PUBLIC HEALTH REPORTING



5.5.1 Description of Scenario

Today Registrars manually translate clinical data into public health reporting systems. This data is used to monitor and improve public health (e.g. surveillance and clinical research). In the future, this data will be extracted from the EHR in an automated (or semi-automated) manner.

© 2013 Health Level Seven International. All rights reserved.

5.5.2 Cancer Registry workflow

Add workflow for next version

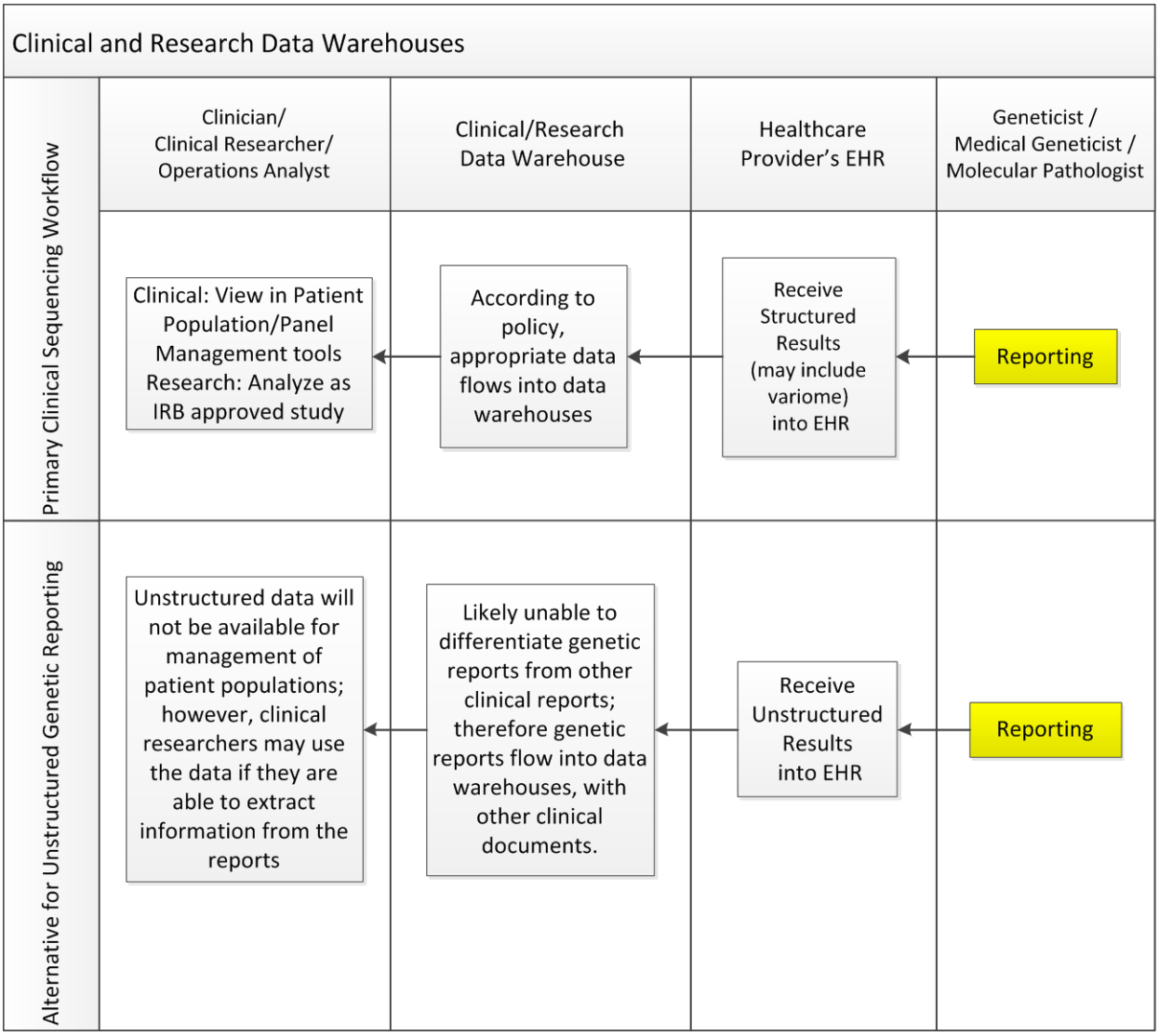
5.5.3 FDA Scenarios in Public Health Reporting

Add workflow for next version.

5.6 NATIONAL MARROW DONOR PROGRAM

Add workflow for next version.

5.7 SCENARIO 6: CLINICAL AND RESEARCH DATA WAREHOUSES



5.7.1 Description of Scenario

Electronic health records (EHRs) are optimized for transactional data and working with one patient record at a time. To enable clinicians to view populations of similar patients (e.g. a primary care provider may

© 2013 Health Level Seven International. All rights reserved.

want to see last mammography dates for all their patients with increased risk of breast cancer), clinical data is incorporated into clinical data warehouses. Similar data warehouses support use of clinical data, for clinical research, according to Institutional Review Board polices. If genetic data is not structured, it doesn't support these activities.

© 2013 Health Level Seven International. All rights reserved.

Nomenclatures, Code Systems and Value Sets

6.Vocabulary Constraints

6.1 GENETIC TESTS, TESTING CONTEXT, INTERPRETATION CODE, AND GENETIC DATA

6.1.1 LOINC

TABLE 6-1 – LABORATORY LOINC	
Code sets, vocabularies, terminologies and nomenclatures that need to be constrained	All LOINC laboratory result codes
Minimum attributes of the component	HL7 value sets not established. Considered value sets should include: HEDIS (Health plan Employer Data and Information Set) reported tests accounting for 95% of routine laboratory orders. Proposed value sets for micro and cytology codes per HITSP/C80. Category A, B, & C bioterrorism agents/diseases. Public Health jurisdiction and Federal reportable disease conditions. The LOINC codes for reporting clinical genomics variation are listed in this implementation guide explicitly – and were developed and tested as described below.
Other Comments	LOINC - Vocabularies and code sets, useful in the reporting of genetic test result data into the EHR, in formats that can be leveraged by clinical decision support, have been defined as a result of the 2 year clinical pilot of the HL7 Version 3 Genetic Variation model. These vocabularies and code sets have be submitted to LOINC, and through ongoing collaborations between the National Library of Medicine's Lister Hill Center for Biomedical Communication, Partners HealthCare Center for Personalized Genetic Medicine (formerly the Harvard – Partners Center for Genetics and Genomics), Partners Healthcare, and Intermountain Healthcare, these vocabularies and codes will be piloted more broadly. In addition, the above collaborators have detailed these vocabularies and code sets in the HL7 implementation guide, balloted in Fall 2009, entitled: HL7 Version 2 Implementation Guide: Clinical Genomics; Fully LOINC-Qualified Genetic Variation Model, Release 1. The full LOINC data base can be downloaded from the Regenstrief Institute, Indianapolis, Indiana at http://loinc.org/.

© 2013 Health Level Seven International. All rights reserved.

6.2 ASSOCIATED DISEASE AND/OR DRUG

6.2.1 SNOMED-CT

TABLE 6-2 – SNOMED-CT	
Code sets, vocabularies, terminologies and nomenclatures that need to be constrained	SNOMED-CT
Minimum attributes of the component	SNOMED-CT FDA SPL Problem List Subset
Other Comments	FDA SPL Problem List Subset available at http://www.fda.gov/ForIndustry/DataStandards/StructuredProductLabeling/ucm163377.htm . The SNOMED terminology is used in the coding of disease associated with sequence variants or genes. Utilization of SNOMED provides linkage of genetic data with other clinical data stored in clinical applications.

6.2.2 RxNORM

TABLE 6-3 – RXNORM	
Code sets, vocabularies, terminologies and nomenclatures that need to be constrained	RxNORM
Minimum attributes of the component	Medication List Subset
Other Comments	Use RxNORM ingredient codes to identify drugs that are the target of pharmacogenomics studies. Utilization of RxNORM provides linkage of genetic data to other clinical data stored in clinical applications. An overview of RXNorm can be obtained from http://www.nlm.nih.gov/research/umls/rxnorm/overview.html

6.3 GENES

6.3.1 HGNC gene symbols (required)

TABLE 6-4 – HGNC	
Code sets, vocabularies, terminologies and nomenclatures that need to be constrained	HGNC
OID	2.16.840.1.113883.6.281
Minimum attributes of the component	Gene symbol
Other Comments	Human Gene Nomenclature Committee (HGNC maintains a database of gene names and symbols. They are a non-profit body which is jointly funded by the US National Human Genome Research Institute (NHGRI) and the

© 2013 Health Level Seven International. All rights reserved.

	Wellcome Trust (UK). They operate under the auspices of Human Genome Organization. The database can be found at http://www.genenames.org/.

6.4 SEQUENCE VARIATIONS

6.4.1 HGVS (required)

TABLE 6-5 – HGVS	
Code sets, vocabularies, terminologies and nomenclatures that need to be constrained	HGVS
OID	2.16.840.1.113883.6.282
Minimum attributes of the component	Sequence variation
Other Comments	Human Genome Variation Society (HGVS) Nomenclature standards for the description of sequence variations are maintained at http://www.hgvs.org/mutnomen/. This standard is well accepted by the clinical genetic community and is extended on an ongoing basis to support genetic findings. HGVS maintains a tool for the generation/translation/verification of HGVS nomenclature. This tool can be found at: http://www.mutalyzer.nl/2.0/index

6.4.2 dbSNP (optional, but highly recommended)

TABLE 6-6 – DBSNP	
Code sets, vocabularies, terminologies and nomenclatures that need to be constrained	dbSNP
OID	2.16.840.1.113883.6.284
Minimum attributes of the component	Rs number and nucleotide change
Other Comments	The Single Nucleotide Polymorphism database (dbSNP). Is maintained by National Center for Biotechnology Information. Available at: http://www.ncbi.nlm.nih.gov/projects/SNP/ Databases and knowledgebases defining sequence variants will be increasingly important. Although sequencing based tests which can result in the identification of novel variants require HGVS nomenclature standards for complete results reporting, genotyping tests which probe for the existence of known variants can additionally report results using an 'RS number' (i.e. identifier in dbSNP) and the associated nucleotide change. (Within the clinical environment results reporting using HGVS nomenclature is required with an option to additionally specify the RS number.)

© 2013 Health Level Seven International. All rights reserved.

6.4.3 COSMIC (optional)

Variants/Mutations can also be reported with a COSMIC mutation identifier associating the findings with internationally compiled cancer mutation data.

TABLE 6-10 – COSMIC	
Code sets, vocabularies, terminologies and nomenclatures that need to be constrained	COSMIC (Catalogue Of Somatic Mutations In Cancer)
Responsible Body	Sanger Institute
OID	2.16.840.1.113883.3.912
Minimum attributes of the component	COSMIC ID
Other Comments	Catalogue Of Somatic Mutations In Cancer (COSMIC) serves as a repository for somatic mutations identified in specific cancer specimens. These mutations are recorded associated with structured description of the specimen. Available at: http://www.sanger.ac.uk/genetics/CGP/cosmic/

6.5 STRUCTURAL VARIANTS

6.5.1 <Add ISCN>

6.5.2 <Add dbVAR>

6.6 COPY NUMBER CHANGE

6.6.1 <Need Nomenclature>

6.7 BIOMARKERS --> IS THIS FAR ENOUGH ALONG TO ADD

6.7.1 <Add MedGen/LOINC>

6.8 REFERENCE SEQUENCES (REQUIRED)

Reference sequences are the baseline from which variation is reported. For example, sequence variants are identified in a patient by comparing the patient's DNA sequence to a reference sequence standard, used in the laboratory. Typically, differences between the patient and reference sequence are called sequence variation and are cataloged, interpreted and reported. Documentation of the reference sequence used is becoming increasingly important for normalization of results between laboratories. To meet this need NCBI is cataloging reference sequences used in clinical testing in the Core Nucleotide Database and can be referred to through the RefSeq identifiers. In collaboration with NCBI, the European BioInformatics Institute (EBI) is also developing a database of reference sequences called Locus Reference Genomic Sequences (LRG). The standard is still in draft status. Importantly, NCBI's RefSeq and EBI's LRG will contain the same reference sequences, annotations and cross references to each other.

© 2013 Health Level Seven International. All rights reserved.

6.8.1 RefSeq

TABLE 6-7 – REFSEQ	
Code sets, vocabularies, terminologies and nomenclatures that need to be constrained	RefSeq
OID	2.16.840.1.113883.6.280
Minimum attributes of the component	RefSeq ID
Other Comments	National Center for Biotechnology Information (NCBI) Reference Sequences contained in Core Nucleotide database. (Note version numbers are required to uniquely identify the reference.) Available at: http://www.ncbi.nlm.nih.gov/nuccore?db=nuccore

6.8.2 LRG

TABLE 6-8 – LRG	
Code sets, vocabularies, terminologies and nomenclatures that need to be constrained	LRG
OID	2.16.840.1.113883.6.283
Minimum attributes of the component	LRG ID
Other Comments	Locus Reference Genomic Sequences an emerging standard led by the European Bioinformatics Institute. Available at: http://www.ebi.ac.uk/ebisearch/search.ebi?db=lrg&t=gene And http://www.lrg-sequence.org/page.php

© 2013 Health Level Seven International. All rights reserved.

6.9 PUBLICLY AVAILABLE REFERENCES (VALUABLE FOR CLINICAL AND TRANSLATIONAL GENOMICS)

6.9.1 OMIM (optional)

Clinical genetic/genomic results can be reported with an OMIM id for association to relevant information in the OMIM knowledgebase, which contains a compendium of information on genetic based disease, genes and mutations.

TABLE 6-9 – OMIM	
Code sets, vocabularies, terminologies and nomenclatures that need to be constrained	OMIM (Online Mendelian Inheritance in Man)
Responsible Body	Johns Hopkins
OID	2.16.840.1.113883.6.174
Minimum attributes of the component	OMIM ID
Other Comments	Knowledgebase for genes, variants/mutations and genetic based phenotypes. Note this information includes somatic or acquired variants/mutations and phenotypes and is not limited to inherited variants/mutations and phenotypes. Available at: http://www.omim.org/ and through NCBI at http://www.ncbi.nlm.nih.gov/omim Additionally, dbSNP contains links to variants in OMIM.

6.9.2 PubMed (optional)

Coding of references may include PubMed ids to peer reviewed literature (e.g. publications within a medical journal).

TABLE 6-10 – PUBMED	
Code sets, vocabularies, terminologies and nomenclatures that need to be constrained	PubMed
Responsible Body	United States National Library of Medicine
OID	2.16.840.1.113883.13.191
Minimum attributes of the component	PubMed ID
Other Comments	"PubMed comprises more than 20 million citations for biomedical literature from MEDLINE, life science journals, and online books. Citations may include links to full-text content from PubMed Central and publisher web sites." Available at: http://www.ncbi.nlm.nih.gov/pubmed/

© 2013 Health Level Seven International. All rights reserved.

6.9.3 PharmGKB (optional)

PharmGKB ids to community curated information on emerging pharmacogenomic associations.

TABLE 6-10 – PHARMGKB	
Code sets, vocabularies, terminologies and nomenclatures that need to be constrained	PharmGKB (Pharmacogenomic Knowledge Base)
Responsible Body	Stanford University, Department of Genetics
OID	2.16.840.1.113883.3.913
Minimum attributes of the component	PharmGKB ID
Other Comments	The mission of PharmGKB is "to collect, encode, and disseminate knowledge about the impact of human genetic variations on drug response. We curate primary genotype and phenotype data, annotate gene variants and gene-drug-disease relationships via literature review, and summarize important PGx genes and drug pathways." Available at: http://www.pharmgkb.org/

6.9.4 ClinicalTrials.gov (optional)

ClinicalTrials.gov id maybe transmitted as part of the interpretation indicating which clinical trials the patient may qualify.

TABLE 6-11 – CLINICALTRIALS.GOV	
Code sets, vocabularies, terminologies and nomenclatures that need to be constrained:	ClinicalTrials.gov
Responsible Body:	U.S. National Institutes of Health and Lister Hill National Center for Biomedical Communications
OID	2.16.840.1.113883.3.1077
Minimum attributes of the component:	ClinicalTrials.gov Identifier
Other Comments:	"ClinicalTrials.gov is a registry of federally and privately supported clinical trials conducted in the United States and around the world. ClinicalTrials.gov gives you information about a trial's purpose, who may participate, locations, and phone numbers for more details. This information should be used in conjunction with advice from health care professionals." http://clinicaltrials.gov/ct2/home

© 2013 Health Level Seven International. All rights reserved.

7.Logical Message Types

7.1 INTRODUCTION AND STRATEGY

The Genetic Test Result Reporting message is defined by a set of four nested LOINC panels, which serve as templates for the messages. In general, LOINC panel definitions include one LOINC code to identify the whole panel and a set of LOINC codes for each child element of that panel. A child element can also be a LOINC panel, and such panels can repeat, to provide a structure that can accommodate many reporting patterns. For each such child element, the panel definition also includes its data type, units of measure, optionality and answer list, as applicable. The definitional information for the four panels used to report Genetics Test Result Reports is included in this guide. It can also be obtained in electronic form from the LOINC web site.

In a message, each new panel of observations begins with an OBR segment that carries the LOINC ID for that panel, and is followed by a series of OBX's each of which carry caries the LOINC ID (OBX-3), and the value (OBX-5) of a particular observation. In a message, the first panel is the master panel for the reporting of genetic analysis. The first child panel delivers an overall summary of the study results and includes options for reporting the traditional narrative report, the overall study impression, and a few other items. Depending on the study being reported, the summary panel may contain variables required to summarize a pharmacogenomics study, or those required to summarize the genetic findings associated with a disease or the risk of a disease (see table 7-2). Next comes the discrete results panel (table 7-5), which contains the detailed results pay load in a series of one or more "*DNA sequence analysis discrete sequence variation panels"* (see table 7-6). This last panel repeats as many times as needed to report all of the variations of interest.

The pattern of panels is shown in the following diagram:

© 2013 Health Level Seven International. All rights reserved.

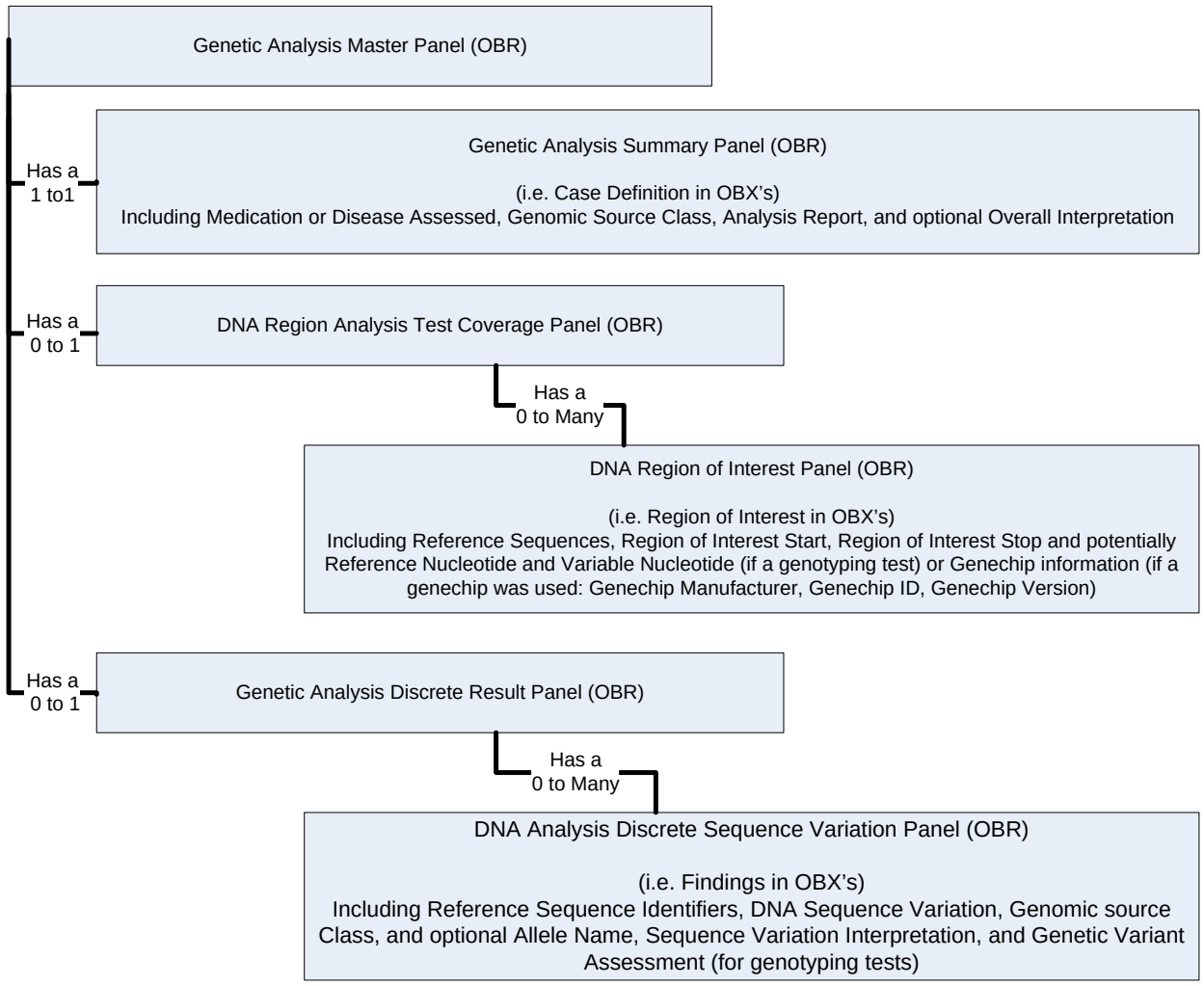


Figure 1. Object model of elements contained within the genetic results message. The master OBR (*Genetic Analysis Master Panel*) contains a child OBR, the *Genetic Analysis Summary Panel*. If information detailing the region of DNA examined is to be included in the HL7 message, then the *DNA Analysis Test Coverage Panel* will be added as a child OBR of the *Genetic Analysis Master Panel*. The *DNA Analysis Test Coverage Panel* itself has one or more child OBR's, the *DNA Region of Interest Panel*. If DNA sequence variations are identified, then the *Genetic Analysis Master Panel* will have another child OBR, the *Genetic Analysis Discrete Result Panel*. This child OBR (the *Genetic Analysis Discrete Result Panel*) itself has one or more child OBR's, the *DNA Analysis Discrete Sequence Variation Panel,* which are used to report genetic findings (sequence variations and gene alleles).

7.2 MESSAGE DEFINITIONS

The complexity of genetic data requires additional coding of the message components using LOINC. These codes are listed in tables 7.1 to 7.8. As already stated in the Introduction, LOINC coding has several advantages including more robust representation of the data when persisted in a database, increased accuracy when supporting multiple HL7 message formats, and consistency for clinical decision support, and applicability to many complex reporting requirements.

7.3 MESSAGE COMPONENTS

PLEASE NOTE:

© 2013 Health Level Seven International. All rights reserved.

The following tables 7-1 through 7-6 are NOT segment definitions. They specify the content of LOINC panels.

7.3.1 Test Interpretation

You can view the list of the individual test components included in each panel. The elements will be accompanied by a usage flag that will denote the expected appearance of the panel element in the panel when resulted. A usage flag is always one of three states:

- **R**equired. The panel element is always expected to be reported when the panel is resulted.

- **O**ptional. Report if available. The panel element may not be reported with a panel result depending upon institutional policy or capabilities of the reporting lab.

- **C**onditional. The panel element is a key finding in the panel report and should be assumed to be negative, absent or not present if the panel result does not include data for this element.

7.3.1.0 Genetic Analysis Master Panel

The *Genetic Analysis Master Panel* is an OBR which will contain a child OBR with summary information of the genetic analysis (i.e. *Genetic Analysis Summary Panel*). In addition, if genetic biomarkers were identified (e.g. a DNA Sequence Variant), then the *Genetic Analysis Master Panel* will have a second child OBR, the *Genetic Analysis Discrete Result Panel* (which in itself will have child OBR's for the Sequence Variants).

TABLE 7-1 – GENETIC ANALYSIS MASTER PANEL

OBR/ OBX	DT	Usage	Card-inality	Value Set	LOINC Code	LOINC Element Name	Description/Comments
OBR		R	1..n		55233-1	Genetic Analysis Master Panel	This is the parent OBR for the panel holding the summary of genetic analysis (i.e. Genetic Analysis Summary Panel).

7.3.1.1 Genetic Analysis Summary Panel

The *Genetic Analysis Summary Panel* is used to report the summary of the genetic analysis. This will fall into several categories: including disease risk/diagnosis, carrier testing, drug metabolism, and drug efficacy. It includes the genetic report, appropriate overall interpretation answer list, disease or drug assessed, and genomic source class.

TABLE 7-2 –GENETIC ANALYSIS SUMMARY PANEL

OBR/ OBX	OBX-2 Value Type	Usage*	Card-inality	Value Set	LOINC Code	LOINC Element Name	Description/Comments
OBR		R	1..n		55232-3	Genetic Analysis Summary Panel	The summary panel for a genetic analysis for one or more laboratory tests (e.g. analysis for disease risk, diagnosis or pharmacogenetics) on a

© 2013 Health Level Seven International. All rights reserved.

							single accession.
OBX	CWE	C1		SNOMED	51967-8	Genetic disease assessed	A coded disease which is associated with the region of DNA covered by the genetic test (recommend SNOMED).
OBX	CWE	C1		RxNORM	51963-7	Medication Assessed	A coded medication accessed in a pharmacogenic test (recommend RxNorm).
OBX	CWE	R			48002-0	Genomic Source Class	The genomic class of the specimen being analyzed: Germline for inherited genome, somatic for cancer genome (e.g. DNA from tumor cells), and prenatal for fetal genome. LOINC Answer List values can be seen in Table 7.8.
If the study is intended to assess disease risk or diagnosis based on genetic findings, then the *Genetic Disease Analysis Overall Interpretation* is used (see below).							
OBX	CWE	C2			51968-6	Genetic Disease Analysis Overall Interpre-tation	Interpretation of all identified DNA Sequence Variations along with any known clinical information for the benefit of aiding clinicians in understanding the results overall in either the context of diagnosis or increased risk of disease. LOINC Answer List values can be seen in Table 7.8.
If reporting a genetic test specifically performed for carrier testing, then the *Genetic Disease Analysis Overall Carrier Interpretation* (below) should replace the *Genetic Disease Analysis Overall Interpretation.*							
OBX	CWE	C2			53039-4	Genetic Disease Analysis Overall Carrier Interpre-tation	Carrier Identification interpretation of all identified DNA Sequence Variations along with any known clinical information for the benefit of aiding clinicians in understanding the results overall. LOINC Answer List values can be seen in Table 7.8.
If the study is intended to assess drug efficacy, include the following LOINC term in the report.							
OBX	CWE	C2			51964-5	Drug Efficacy Analysis Overall Interpre-tation	Overall predicted phenotype for drug efficacy for all DNA Sequence Variations identified in a single case. LOINC Answer List values can be seen in Table 7.8.

© 2013 Health Level Seven International. All rights reserved.

If the study is intended to assess the effect on metabolism, include the following LOINC term in the report.							
OBX	CWE	C2			51971-0	Drug metabolism analysis overall interpret-tation	Overall predicted phenotype for drug metabolism for all DNA Sequence Variations identified in a single case. LOINC Answer List values can be seen in Table 7.8.
OBX	FT	O	0..1		51969-4	Genetic analysis summary report	Narrative report in disease diagnostic-based format, which is used for pharmacogenomic reporting as well and disease risk or diagnosis. These reports currently follow the same formatting recommendations.
OBX	FT	O			53577-3	Reason for Study Additional Note	Descriptive text to further annotate the coded *Reason for Study* associated with an ordered test.

* R = Required; O = Optional; C1 = Conditional, either of the C1 values is required according to the following rules. C1 Rules – If the test is performed for diagnostic confirmation or identification of pre-clinical disease risk, then 'Genetic Disease Assessed', or if the genetic test is performed for pharmacogenomic usage, then 'Medication Assessed' is used. C2 = Conditional, only one of the values can be used according to the following rules: C2 Rules – If the test is performed for diagnostic confirmation or identification of pre-clinical disease risk, then 'Genetic Disease Analysis Overall Interpretation' is used. If the test is performed for identification of carrier status, then 'Genetic Disease Analysis Overall Carrier Interpretation' is used. If a pharmacogenomic test is performed for indication of drug efficacy, then 'Drug Efficacy Analysis Overall Interpretation' is used. If a pharmacogenomic test is performed for indication of drug metabolism, then 'Drug metabolism analysis overall interpretation' is used.

7.3.2 Test Definition

7.3.2.0 DNA Region Analysis Test Coverage Panel

The *DNA Region Analysis Test Coverage Panel* is a child OBR of the *Genetic Analysis Master Panel*, and will contain child OBR's defining the discrete regions of interest. Currently, these are gene (or partial gene) sequencing or genotyping in the *DNA Region of Interest Panel*.

Prior to the widespread adoption of whole genome sequencing, knowing which region of the genome was tested is often important for clinical interpretation of the results and determining need for follow-up testing (specifically for negative or inconclusive results). For instance, if only genetic regions most commonly found to have mutations associated with a specific disease were tested, the clinician may want to follow-up with a test which examines the remaining genetic regions currently known to be associated with the disease. Additionally, as new regions are discovered (to be associated with the specific disease), the clinician may want to order these newly available genetic tests. For genotyping tests, the reference and variable nucleotides would also be reported, if known. If the test is performed using a genechip, then information on the manufacturer, genechip and version should also be listed. This information should be captured within the *DNA Region of Interest Panel.*

Important note: For sequencing tests, receiving systems should not assume that they can derive the patient sequence by comparing the reference sequence and the reported mutations, because polymorphisms and other differences from the reference (which are thought to be non-clinically important

© 2013 Health Level Seven International. All rights reserved.

at the time of reporting) may not have been reported. Derivation of the patient sequence will only be possible if ALL differences from the reference sequence are reported in the findings. In this scenario, using the 'Sequence Variation Interpretation' to prioritize variants displayed within the electronic medical record would provide a mechanism to not overwhelm the clinician with data on benign variants (not currently deemed important for clinical care).

TABLE 7-3 – DNA REGION ANALYSIS TEST COVERAGE PANEL							
OBR/ OBX	**DT**	**Usage**	**Card-inality**	**Value Set**	**LOINC Code**	**LOINC Element Name**	**Description/Comments**
OBR		R	1..n	N/A	51956-1	*7.3.2.0.1 DNA Region Analysis Test Coverage Panel*	This is the parent OBR for genetic panels for regions of interest (e.g. *DNA Region of Interest Panel).*

7.3.2.1 DNA Region of Interest Panel

The *DNA Region of Interest Panel* describes the regions examined for variation or confirmation of matching to the reference sequence (e.g. wild-type or 'normal'). For sequencing based tests, one or more regions will be sequenced and the region of interest will fall within the sequenced region.

TABLE 7-4 – DNA REGION OF INTEREST PANEL							
OBR/ OBX	**OBX-2 Value Type**	**Usage***	**Card-inality**	**Value Set**	**LOINC Code**	**LOINC Element Name**	**Description/Comments**
OBR		R	1..n	N/A	53041-0	*7.3.2.1.1 DNA Region of Interest Panel*	A set of observations which comprise the definition of a contiguous sequence of nucleotides in the genome. This is intended for use in genetic test coverage panels in genetic test reporting.
OBX	CWE	R	0…1		48013-7	*7.3.2.1.2 Genomic reference sequence identifier*	This field carries the ID for the genomic reference sequence. The genomic reference sequence is a contiguous stretch of chromosome DNA that spans all of the exons of the gene and includes transcribed and non transcribed stretches. For this ID, use either the NCBI genomic nucleotide RefSeq IDs with their version number (see: NCBI.NLM.NIH.Gov/RefSeq), or use the LRG identifiers (see: http://www.lrg-sequence.org/)

© 2013 Health Level Seven International. All rights reserved.

OBX	CWE	O	0…1		69545-2	*7.3.2.1.3 Region of Interest Start*	This is the nucleotide number starting the region which was examined for variants (or wild-type confirmation) during testing, and is to be used in conjunction with *Region of Interest Stop*.
OBX	CWE	O	0…1		69546-0	*7.3.2.1.4 Region of Interest Stop*	This is the nucleotide number ending the region which was examined for variants (or wild-type confirmation) during testing, and is to be used in conjunction with *Region of Interest Start*.
OBX	CWE	O	0…1		69547-8	*7.3.2.1.5 Reference Nucleotide*	For genotype tests, this is the nucleotide sequence within the region of interest on the reference sequence, and is to be used in conjunction with *Variable Nucleotide*.
OBX	CWE	O	0…1		69551-0	*7.3.2.1.6 Variable Nucleotide*	For genotype tests, this is the nucleotide sequence for the variation (or wild-type confirmation) tested for, and is to be used in conjunction with *Reference Nucleotide*.
OBX	CWE	O	0…1		48010-3	*7.3.2.1.7 Genechip ID*	For Genechip based tests, this contains the Genechip identifier, and is to be used in conjunction with Genechip *Manufacturer ID* and *Genechip Version*.
OBX	CWE	O	0…1		48009-5	*7.3.2.1.8 Genechip Manufacturer ID*	For Genechip based tests, this contains the Genechip manufacturer, and is to be used in conjunction with Genechip *ID* and *Genechip Version*.
OBX	CWE	O	0…1		48011-1	*7.3.2.1.9 Genechip Version*	For Genechip based tests, this contains the Genechip version number, and is to be used in conjunction with Genechip *Manufacturer ID* and *Genechip ID*.

7.3.3 Findings

7.3.3.0 Genetic Analysis Discrete Result Panel

The *Genetic Analysis Discrete Result Panel* is a child OBR of the Genetic Analysis Master Panel, and will contain child OBR's defining the discrete findings. Currently, these are DNA sequence variants (in *DNA Analysis Discrete Sequence Variant Panel*), but will expand in future releases of the guide to include other types of genetic biomarkers, as mentioned in Chapter 9.

© 2013 Health Level Seven International. All rights reserved.

TABLE 7-5 – GENETIC ANALYSIS DISCRETE RESULT PANEL							
OBR/ OBX	**DT**	**Usage**	**Card-inality**	**Value Set**	**LOINC Code**	**LOINC Element Name**	**Description/Comments**
OBR		R	1..n	N/A	55207-5	*7.3.3.0.1 Genetic Analysis Discrete Result Panel*	This is the parent OBR for genetic panels of genetic findings (e.g. *DNA Analysis Discrete Sequence Variant Panel).*

7.3.3.1 DNA Analysis Discrete Sequence Variation Panel

The *DNA Analysis Discrete Sequence Variant Panel* corresponds to the Genetic Variation model SequenceVariation class. It describes the characteristics of an identified SequenceVariation - either of clinical relevance, or as a benign difference from the reference sequence, and reported for completeness.

TABLE 7-6 – DNA ANALYSIS DISCRETE SEQUENCE VARIATION PANEL							
OBR/ OBX	**OBX-2 Value Type**	**Usage***	**Card-inality**	**Value Set**	**LOINC Code**	**LOINC Element Name**	**Description/Comments**
OBR		R	1..n	N/A	55208-3	*7.3.3.1.1 DNA Analysis Discrete Sequence Variant Panel*	The set of observations representing all identified DNA Sequence Variations (variants and wild type) for all of the genetic assays performed for a single accession.
OBX	CWE	O		HGNC	48018-6	Gene Identifier	HGNC gene Identifier set by the Human Genome Organization Nomenclature Committee.

© 2013 Health Level Seven International. All rights reserved.

OBX	CWE	C3		NCBI	48013-7	Genomic Reference Sequence Identifier	This field carries the ID for the genomic reference sequence. The genomic reference sequence is a contiguous stretch of chromosome DNA that spans all of the exons of the gene and includes transcribed and non transcribed stretches. For this ID use either the NCBI genomic nucleotide RefSeq IDs with their version number (see: NCBI.NLM.NIH.Gov/RefSeq), or use the LRG identifiers without transcript (t or p) extensions when they become available. (See-Report sponsored by GEN2PHEN at the European Bioinformatics Institute at Hinxton UK April 24-25, 2008). C3 = third set of conditionals. Either 'Genomic Reference Sequence Identifier' or 'Transcript Reference Sequence Identifier' is required. Both may be used.

© 2013 Health Level Seven International. All rights reserved.

OBX	CWE	C3		NCBI	51958-7	Transcript Reference Sequence Identifier	This field carries the ID for the transcribed reference sequence that part of the genetic reference sequence that is converted to Messenger RNA. For this ID use either the NCBI nucleotide RefSeq IDs for transcribed DNA, plus the version number (NCBI.NLM.NIH.Gov/RefSeq), or use the LRG identifiers with transcript (t or p) extensions when they become available. (Report sponsored by GEN2PHEN at the European Bioinformatics Institute at Hinxton UK April 24-25, 2008). The NCI RefSeq transcripts IDs have a prefix of “NM” for genes from the nuclear chromosomes. NCBI does not currently provide a transcript RefSeq for mitochondrial genes. The LRG transcripts Identifiers have a prefix of “LRG_” plus a t extension. Mitochondrial genes are out of scope of LRG. C3 = third set of conditionals. Either ‘Genomic Reference Sequence Identifier’ or ‘Transcript Reference Sequence Identifier’ is required (with version number). Both may be used.
OBX	CWE	O			48008-7	Allele Name	The published, commonly used name for a gene allele is recommended if it exists.
OBX	CWE	O		NCBI	48003-8	DNA Sequence Variation Identifier	A DNA Sequence Variation identifier conveys a universal or standard repository identifier for definitive characteristics of a DNA Sequence Variation. (If available, recommend using NCBI dbSNP ids - RS#)

© 2013 Health Level Seven International. All rights reserved.

OBX	CWE	C4		HGVS	48004-6	DNA Sequence Variation	Human Genome Variation Society (HGVS) nomenclature for a single or set of DNA Sequence Variation(s) identified in testing. The use of the nomenclature is also used to describe non-variations (aka. wild types). Either the DNA Sequence Variation is required or the Amino Acid Change. NOTE: If NCBI's dbSNP IDs (RS#) is used, then the DNA Sequence Variation is required to uniquely define the Variant, as the number is unique to the nucleotide location and requires the details of nucleotide change. C4 = fourth set of conditionals. Either 'DNA Sequence Variation' or 'Amino Acid Change' is required. Both may be used.
OBX	CWE	O		HGVS	48019-4	DNA Sequence Variation Type	Codified type for associated DNA Sequence Variation. DNA Sequence Variations use the HGVS notation which implies the DNA Sequence Variation Type, but the concurrent use of this code will allow a standard and explicit type for technical and display convenience. LOINC Answer List values can be seen in Table 7.8.

© 2013 Health Level Seven International. All rights reserved.

OBX	CWE	C4		HGVS	48005-3	Amino Acid Change	Human Genome Variation Society (HGVS) nomenclature for an amino acid change. This value is derivable from the DNA Sequence Variation value if available. It is provided for convenience. The use of the nomenclature is also used to describe non-variations (aka. wild types). Either the DNA Sequence Variation is required or the Amino Acid Change. C4 = fourth set of conditionals. Either 'DNA Sequence Variation' or 'Amino Acid Change' is required. Both may be used.
OBX	CWE	O		HGVS	48006-1	Amino Acid Change Type	Codified type for associated Amino Acid Change. Amino Acid Change's use the HGVS notation which implies the Amino Acid Change Type, but the concurrent use of this code will allow a standard and explicit type for technical and display convenience. LOINC Answer List values can be seen in Table 7.8.
OBX	CWE	O			47999-8	DNA Region Name	A human readable name for the region of interest. Typically Exon #, Intron # or other. NOTE: This is not standardized and is mainly for convenience and display purposes.

© 2013 Health Level Seven International. All rights reserved.

OBX	CWE	O			53034-5	Allelic State	The level of occurrence of a single DNA Sequence Variation within a set of chromosomes. Heterozygous indicates the DNA Sequence Variation is only present in one of the two genes contained in homologous chromosomes. Homozygous indicates the DNA Sequence Variation is present in both genes contained in homologous chromosomes. Hemizygous indicates the DNA Sequence Variation exists in the only single copy of a gene in a non-homologous chromosome (the male X and Y chromosome are non-homologous). Hemiplasmic indicates that the DNA Sequence Variation is present in some but not all of the copies of mitochondrial DNA. Homoplasmic indicates that the DNA Sequence Variation is present in all of the copies of mitochondrial DNA. LOINC Answer List values can be seen in Table 7.8.
OBX	CWE	O			48002-0	Genomic Source Class	The genomic class of the specimen being analyzed: germline for inherited genome, somatic for cancer genome, and prenatal for fetal genome. LOINC Answer List values can be seen in Table 7.8.
OBX	ST	O			47998-0	DNA Sequence Variation Display Name	Thumbnail "textual display" convention of a DNA Sequence Variation and its interpretation. For example: c.2235_2249del (Glu746_Ala750del), Exon 19, EGFR, Responsive

© 2013 Health Level Seven International. All rights reserved.

If reporting a genetic test specifically performed for identification of DNA Sequence Variations associated with disease, then the *Genetic Disease Sequence Variation Interpretation* should be used for coding interpretations at the DNA Sequence Variation level.							
OBX	CWE	C5			53037-8	Genetic Disease Sequence Variation Interpre-tation	Interpretation of the pathogenicity of the DNA Sequence Variation in the context of the assessed genetic disease. LOINC Answer List values can be seen in Table 7.8.
If reporting a genetic test specifically performed for identification of DNA Sequence Variations associated with drug metabolism, then the *Drug Metabolism Sequence Variation Interpretation* should be used for coding interpretations at the DNA Sequence Variation level.							
OBX	CWE	C5			53040-2	Drug Metabolism Sequence Variation Interpre-tation	Predicted phenotype for drug efficacy. A sequence variation interpretation value known to allow (responsive) or prevent (resistant) the drug to perform. LOINC Answer List values can be seen in Table 7.8.
If reporting a genetic test specifically performed for identification of DNA Sequence Variations associated with drug efficacy, then the *Drug Efficacy Sequence Variation Interpretation* should be used for coding interpretations at the DNA Sequence Variation level.							
OBX	CWE	C5			51961-1	Drug Efficacy Sequence Variation Interpre-tation	Predicted phenotype for ability of drug to bind to intended site in order to deliver intended effect. A Sequence Variation interpretation value known to allow (responsive) or prevent (resistant) the drug to perform. LOINC Answer List values can be seen in Table 7.8.
If reporting a genetic test specifically performed for determining the presence or absence of a known set of DNA Sequence Variations (e.g. molecular profile of the tumor) without an association to phenotype, then the *Genetic Variant Assessment* should be used for coding interpretations at the DNA Sequence Variation level.							
OBX	CWE	C5			69548-6	Genetic Variant Assessmen t	Denoting the genotyping result (e.g. present, absent, or no call) for a known set of mutations within a genetic profiling test. LOINC Answer List values can be seen in Table 7.8.

*R = Required; O = Optional; C3 = Conditional, one of the C3 values is required. Both can be provided. C4 = Conditional, one of the C4 values is required. Both can be provided. C5 = Conditional. Only one of the values can be used with the following rules: C5 rules – If the test is performed for diagnostic confirmation or identification of pre-clinical disease risk, then *Genetic Disease Analysis Overall Interpretation* is used in the *Genetic Analysis Summary Panel* and *Genetic Disease Sequence Variation Interpretation* is used in *DNA Analysis Discrete Sequence Variant Panel*. If the test is performed for identification of carrier status, then *Genetic Disease Analysis Overall Carrier Interpretation* is used in the *Genetic Analysis Summary Panel* and *Genetic Disease Sequence Variation Interpretation* is used in *DNA Analysis Discrete Sequence Variant Panel*. If a pharmacogenomic test is performed for indication of

© 2013 Health Level Seven International. All rights reserved.

drug efficacy, then *Drug Efficacy Analysis Overall Interpretation* is used in the *Genetic Analysis Summary Panel* and *Drug Efficacy Sequence Variation Interpretation* is used in the *DNA Analysis Discrete Sequence Variant Panel*. If a pharmacogenomic test is performed for indication of drug metabolism, then *Drug Metabolism Analysis Overall Interpretation* is used in the *Genetic Analysis Summary Panel* and *Drug Metabolism Sequence Variation Interpretation* is used in the *DNA Analysis Discrete Sequence Variant Panel*.

7.4 LOINC CODES

TABLE 7-7 – LOINC CODES

LOINC #	Component	Property	Time	System	Scale	Method
47998-0	DNA Sequence Variation display name	Txt	Pt	Bld/Tiss	Nar	Molgen
47999-8	DNA region name	ID	Pt	Bld/Tiss	Nom	Molgen
48002-0	Genomic source class (LOINC Answer List values can be seen in table 7.8)	Type	Pt	Bld/Tiss	Nom	Molgen
48003-8	DNA Sequence Variation identifier	ID	Pt	Bld/Tiss	Nom	Molgen
48004-6	DNA Sequence Variation	Find	Pt	Bld/Tiss	Nom	Molgen
48005-3	Amino acid change	Find	Pt	Bld/Tiss	Nom	Molgen
48006-1	Amino acid change type (LOINC Answer List values can be seen in table 7.8)	Type	Pt	Bld/Tiss	Nom	Molgen
48008-7	Allele name	ID	Pt	Bld/Tiss	Nom	Molgen
48013-7	Genomic reference sequence identifier	ID	Pt	Bld/Tiss	Nom	Molgen
48018-6	Gene identifier	ID	Pt	Bld/Tiss	Nom	Molgen
48019-4	DNA Sequence Variation type (LOINC Answer List values can be seen in table 7.8)	Type	Pt	Bld/Tiss	Nom	Molgen
51958-7	Transcript reference sequence identifier	ID	Pt	Bld/Tiss	Nom	Molgen
51959-5	DNA region of interest	Prod	Pt	Bld/Tiss	Nom	Molgen
51961-1	Drug efficacy sequence variation interpretation (LOINC Answer List values can be seen in table 7.8)	Imp	Pt	Bld/Tiss	Ord	Molgen
51963-7	Medication assessed	Prid	Pt	Bld/Tiss	Nom	Molgen
51964-5	Drug efficacy analysis overall interpretation (LOINC Answer List values can be seen in table 7.8)	Imp	Pt	Bld/Tiss	Nom	Molgen
51967-8	Genetic disease assessed	ID	Pt	Bld/Tiss	Nom	Molgen
51968-6	Genetic Disease Analysis Overall Interpretation	Imp	Pt	Bld/Tiss	Nom	Molgen

© 2013 Health Level Seven International. All rights reserved.

TABLE 7-7 – LOINC CODES						
LOINC #	**Component**	**Property**	**Time**	**System**	**Scale**	**Method**
	(LOINC Answer List values can be seen in table 7.8)					
51969-4	Genetic analysis summary report	Find	Pt	Bld/Tiss	Doc	Molgen
51971-0	Drug metabolism analysis overall interpretation (LOINC Answer List values can be seen in table 7.8)	Imp	Pt	Bld/Tiss	Nom	MolGen
53034-5	Allelic state	Find	Pt	Bld/Tiss	Nom	Molgen
53037-8	Genetic disease sequence variation interpretation (LOINC Answer List values can be seen in table 7.8)	Imp	Pt	Bld/Tiss	Nom	Molgen
53039-4	Genetic disease analysis overall carrier interpretation (LOINC Answer List values can be seen in table 7.8)	Type	Pt	Bld/Tiss	Nom	Molgen
53040-2	Drug metabolism sequence variation interpretation (LOINC Answer List values can be seen in table 7.8)	Imp	Pt	Bld/Tiss	Nom	Molgen
53577-3	Reason for study additional note	Txt	Pt	Bld/Tiss	Nar	Molgen
55208-3	DNA Analysis Discrete Sequence Variation Panel	-	Pt	Bld/Tiss	-	Molgen
55207-5	Genetic Analysis Discrete Result Panel	-	Pt	Bld/Tiss	-	Molgen
55233-1	Genetic analysis master panel	-	Pt	Bld/Tiss	-	Molgen
55232-3	Genetic analysis summary panel	-	Pt	Bld/Tiss	-	Molgen
51956-1	DNA Region Analysis Test Coverage Panel	-	Pt	Bld/Tiss	-	Molgen
53041-0	DNA Region of Interest Panel	-	Pt	Bld/Tiss	-	Molgen
69545-2	Region of Interest Start	Prid	Pt	Bld/Tiss	Nom	Molgen
69546-0	Region of Interest Stop	Prid	Pt	Bld/Tiss	Nom	Molgen
69547-8	Reference Nucleotide	Prid	Pt	Bld/Tiss	Nom	Molgen
69551-0	Variable Nucleotide	Prid	Pt	Bld/Tiss	Nom	Molgen
48010-3	Genechip ID	ID	Pt	Bld/Tiss	Nom	Molgen
48009-5	Genechip Manufacturer ID	ID	Pt	Bld/Tiss	Nom	Molgen
48011-1	Genechip Version	ID	Pt	Bld/Tiss	Nom	Molgen
69548-6	Genetic Variant Assessment (LOINC Answer List values can be seen in table 7.8)	Imp	Pt	Bld/Tiss	Nom	Molgen

© 2013 Health Level Seven International. All rights reserved.

TABLE 7-7 – LOINC CODES						
LOINC #	**Component**	**Property**	**Time**	**System**	**Scale**	**Method**
69549-4	Genetic Knowledge Reference	Prid	Pt	Bld/Tiss	Nom	Molgen
69550-2	Genetic Knowledge Reference Type (LOINC Answer List values can be seen in table 7.8)	Type	Pt	Bld/Tiss	Nom	Molgen

7.5 LOINC ANSWER LISTS

TABLE 7-8 – LOINC ANSWER LISTS				
LOINC code	**LOINC component**	**Sequence**	**Answer text**	**LOINC answer code**
53034-5	Allelic state	1	Heteroplasmic	LA6703-8
		2	Homoplasmic	LA6704-6
		3	Homozygous	LA6705-3
		4	Heterozygous	LA6706-1
		5	Hemizygous	LA6707-9
48006-1	Amino acid change type	1	Wild type	LA9658-1
		2	Deletion	LA6692-3
		3	Duplication	LA6686-5
		4	Frameshift	LA6694-9
		5	Initiating Methionine	LA6695-6
		6	Insertion	LA6687-3
		7	Insertion and Deletion	LA9659-9
		8	Missense	LA6698-0
		9	Nonsense	LA6699-8
		10	Silent	LA6700-4
		11	Stop Codon Mutation	LA6701-2
48019-4	DNA sequence variation type	1	Wild type	LA9658-1
		2	Deletion	LA6692-3
		3	Duplication	LA6686-5
		4	Insertion	LA6687-3
		5	Insertion/Deletion	LA6688-1
		6	Inversion	LA6689-9
		7	Substitution	LA6690-7
51964-5	Drug efficacy analysis overall interpretation	1	Responsive	LA6677-4
		2	Resistant	LA6676-6
		3	Negative	LA6577-6
		4	Inconclusive	LA9663-1
		5	Failure	LA9664-9
51961-1	Drug efficacy sequence variation interpretation	1	Resistant	LA6676-6
		2	Responsive	LA6677-4
		3	Presumed resistant	LA9660-7
		4	Presumed responsive	LA9661-5
		5	Unknown Significance	LA6682-4
		6	Benign	LA6675-8
		7	Presumed Benign	LA6674-1
		8	Presumed non-responsive	LA9662-3
51971-0	Drug metabolism analysis overall interpretation	1	Ultrarapid metabolizer	LA10315-2
		2	Extensive metabolizer	LA10316-0

© 2013 Health Level Seven International. All rights reserved.

TABLE 7-8 – LOINC ANSWER LISTS				
LOINC code	**LOINC component**	**Sequence**	**Answer text**	**LOINC answer code**
		3	Intermediate metabolizer	LA10317-8
		4	Poor metabolizer	LA9657-3
		5	Inconclusive	LA9663-1
53040-2	Drug metabolism sequence variation interpretation	1	Ultrarapid metabolizer	LA10315-2
		2	Extensive metabolizer	LA10316-0
		3	Intermediate metabolizer	LA10317-8
		4	Poor metabolizer	LA9657-3
		5	Unknown Significance	LA6682-4
53039-4	Genetic disease analysis overall carrier interpretation	1	Carrier	LA10314-5
		2	Negative	LA6577-6
		3	Inconclusive	LA9663-1
		4	Failure	LA9664-9
51968-6	Genetic disease analysis overall interpretation	1	Positive	LA6576-8
		2	Negative	LA6577-6
		3	Inconclusive	LA9663-1
		4	Failure	LA9664-9
53037-8	Genetic disease sequence variation interpretation	1	Pathogenic	LA6668-3
		2	Presumed pathogenic	LA6669-1
		3	Unknown significance	LA6682-4
		4	Benign	LA6675-8
		5	Presumed benign	LA6674-1
48002-0	Genomic source class	1	Germline	LA6683-2
		2	Somatic	LA6684-0
		3	Prenatal	LA6685-7
		4	Likely Germline	LA18194-3
		5	Likely Somatic	LA18195-0
		6	Likely Prenatal	LA18196-8
		7	Unknown Genomic Origin	LA18197-6
69548-6	Genetic Variant Assessment	1	Present	LA9633-4
		2	Absent	LA9634-2
		3	No Call	LA18198-4
69550-2	Genetic Knowledge Reference Type (LOINC Property = Type)	1	Disease	LA18199-2
		2	Drug	LA18200-8
		3	Gene	LA18201-6
		4	DNA Sequence Variant	LA18202-4
		5	Amino Acid Change	LA18203-2
		6	Test	LA7556-9

7.6 SPECIAL SYNTAX AND NOTES ON USE OF LOINC TERMS

7.6.1 DNA Sequence Variation (see table 7-5)

The expression or nomenclature used to describe SNP locations for sequence variations is as follows:

c.-32_129+10^^HGVS

Discussions regarding the uniform and unequivocal description of sequence variants in DNA and protein sequences (mutations, polymorphisms) were initiated by two papers published in 1993; Beaudet AL & Tsui LC and Beutler E. The original suggestions presented were widely discussed, modified, extended

© 2013 Health Level Seven International. All rights reserved.

and ultimately resulted in nomenclature recommendations that have been largely accepted and are applied world-wide.

The most important rule is that all variants should be described at the most basic level, i.e. the DNA level. Descriptions should always be in relation to a **reference sequence** (see section 6.1.5), either a **genomic** or a **coding DNA** reference sequence. Discussions on which type of reference sequence to prefer, genomic or coding DNA, have been very lively. Although theoretically a genomic reference sequence seems best, in practice a coding DNA reference sequence is preferred (*see Reference Sequence discussions*). Finally, version numbers are required to uniquely identify the reference.

http://www.hgvs.org/mutnomen/refseq.html

7.6.2 Genomic Source Class

Where testing of germline reference material may not have been performed (due to workflow or unavailability of the specimen), the Genomic Source Class list as been extended to support 'Likely Somatic', 'Likely Germline', 'Likely Prenatal' and 'Unknown Genomic Origin'.

It is strongly recommended to use the full list of *Genomic Source Class* values at the mutation level; however, top level genomics source class (around the testing context) should only use the original three (somatic, germline or prenatal) as it codifies the testing context (i.e. tumor testing, inherited genetic testing, or prenatal testing).

7.6.3 Genetic Variant Assessment

Where testing scenarios are intended to assess the presence or absence of a known set of DNA variants (e.g. tumor profiling using genotyping technology), then the Genetic Variant Assessment answer list supports structured communication of these findings. Of note, 'No Call' is different from 'Absent', because 'No Call' did not result in the determination of the marker's presence or absence. This may be due to test failure or specimen specific context which renders the test ineffective. If no result is produced, it can be for one of two reasons: 1) the assay failed or 2) the region of the chromosome/gene containing the sequence variation being genotyped is deleted. For instance, if a portion of the PTEN gene is deleted, then all assays for mutations within the deleted region would be 'no call'. In this scenario, it should not be assumed that this region is deleted, because the assays covering this region of interest may have simply failed. Sequencing remains the gold standard for identification of these variants.

© 2013 Health Level Seven International. All rights reserved.

8.Data Set Considerations and HIT Data Standards

8.1 FAMILY HISTORY

A minimal core data set for family history can be found at in the ONC/HHS family history data requirements as developed by the multi-stakeholder workgroup (available at: http://healthit.hhs.gov/portal/server.pt/community/use_cases_and_requirements_documents/1202/personalized_healthcare/15671)

8.2 SEQUENCE VARIATIONS / CHROMOSOMAL CHANGE

8.2.1 Small Genetic Variations

8.2.2 Structural Variations

8.2.3 Copy Number Change

9.HL7 Encapsulation of Genomic Data Files

<Amnon is developing this content>

© 2013 Health Level Seven International. All rights reserved.

10.Clinical Grade-Genomic Data File Standards

<Content derived from federally-mediated workgroup for development of clinical grade-GVF/VCF files>

© 2013 Health Level Seven International. All rights reserved.

11.Gaps & Extensions

11.1 LABORATORY ORDER ENTRY

One significant gap is the need to develop a laboratory order implementation guide for clinical sequencing/molecular diagnostics, which is capable of including relevant clinical history and a fully structured family history with familial mutations and risk assessment. Currently, laboratory orders are paper or pdf based, which has fulfilled the need while volumes remain low. However, as genetic analysis becomes a standards part of clinical care, paper-based order entry will not scale.

12.Outstanding Questions

1. Will electronic health records (EHRs) incorporate a genomic repository housing a patient's genome/variome for access on demand, in much the same way images are stored in PACS (picture archiving and communication system)? Or will EHRs contain a pointer to a centralized repository? Or will the laboratory continue to sequence a patient's DNA each time a test is ordered?

13.Glossary

Genome: Entirety of a patient's inherited genetic information, unless specified as the cancer genome.

Sequence Variation: Variation from a common DNA reference sequence and synonymous with mutation.

Transcoding: Process of converting genetic data from a bioinformatic representation into a clinical representation, following healthcare IT data standards.

Variome: Variation from a reference sequence. That is a patient's DNA sequence can either be stored as a true sequence of nucleotide, or can be stored as a series of variations from a common reference sequence.

14.Future Plans

14.1 LABORATORY GENOMIC DATA STANDARDS

Identify and collaborate with stakeholders for laboratory genomic data standards, to ensure support for required annotations key to clinical processing and reporting (e.g. germline vs. somatic variants).

14.2 EXTENSION OF SEQUENCE VARIATION AND CYTOGENETIC HL7 MODELS

Current HL7 standards for sequence variation and cytogenetic findings use established clinical standards. These will be extended to support inclusion of established bioinformatic representation, to support linking to research and clinical information systems.

14.3 EXTENSION TO SPECIMEN SCENARIOS

14.3.1 Microbiome analysis of the patient

a. Includes analysis of microorganisms living in the patients gastrointestinal tract or Genitourinary system and may aid in diagnosis

© 2013 Health Level Seven International. All rights reserved.

This page intentionally left blank.

© 2013 Health Level Seven International. All rights reserved.