



HL7 Version 3 Domain Analysis Model: **Clinical Sequencing, Release 1**

January 2013

HL7 Comment Only Ballot

Sponsored by:
Clinical Genomics Work Group

Principal Contributors:

Mollie Ullman-Cullere
Kevin Hughes
Daryl Thomas
Grant Wood
Amnon Shabo

Questions or comments regarding this document should be directed to Mollie Ullman-Cullere at mullmancullere@partners.org.

HL7 VERSION 3

DOMAIN ANALYSIS MODEL:

CLINICAL SEQUENCING, RELEASE 1

(1ST BALLOT FOR COMMENT)

ORU^R01

HL7 Version 3

December, 2012

| | |
|--|---|
| 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 | |
| | |
| | |

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 IMPLEMENTORS | 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 |
| 5.6 Scenario 6: Clinical and Research Data Warehouses..... | 13 |

Table of Contents

| | |
|--|-----------|
| 5.6.1 Description of Scenario..... | 13 |
| 6. DATA SET CONSIDERATIONS AND HIT DATA STANDARDS | 14 |
| 6.1 FAMILY HISTORY..... | 14 |
| 6.2 Sequence Variations / Chromosomal change | 14 |
| 6.2.1 Variations within a Gene..... | 14 |
| 6.2.2 Other Sequence Variations - Chromosomal Change | 14 |
| 6.2.3 Future Enhancements..... | 14 |
| 7. GAPS & EXTENSIONS | 14 |
| 7.1 Laboratory Order Entry..... | 14 |
| 7.2 laboratory Genomic Data Standards | 14 |
| 8. OUTSTANDING QUESTIONS | 15 |
| 9. GLOSSARY..... | 15 |
| 10. FUTURE PLANS..... | 15 |
| 10.1 Laboratory Genomic Data Standards | 15 |
| 10.2 Extension of Sequence Variation and Cytogenetic HL7 Models..... | 15 |
| 10.3 Extension to Specimen scenarios..... | 15 |
| 10.3.1 Microbiome analysis of the patient..... | 15 |

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.

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.

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 data system.
- Privacy and security has been implemented at an acceptable level.
- All participants agree to all standards, methodologies, consent, privacy and security.

Chapter 1: Introduction

- 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.

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>)

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. <i>in vitro</i> 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. |

| Internal and External Systems for Stakeholder Groups | | | | |
|--|---|--|--|--|
| | Patient Systems | Healthcare Provider Systems | Laboratory Systems | Geneticist...Systems |
| | <div>PHR (Patient Controlled)</div> <div>Patient Portal (Part of EHR)</div> | <div>EHR: medical record, lab order entry, lab/pathology results, clinical guidelines, e-prescription, image repository, clinical decision support systems, specialized clinical modules, family history analysis tools</div> <div>Genetic Data Repository (containing patient variome or sequence)</div> <div>Genetic Knowledgebase</div> | <div>NCBI Databases: RefSeq, dbSNP, dbVAR, OMIM, PubMed, Genetcs Home Reference, GeneTests.org,</div> <div>Lab Information Management System, BioInformatic Pipeline, results database</div> | <div>NCBI Databases: RefSeq, dbSNP, dbVAR, OMIM, PubMed, Genetcs Home Reference, GeneTests.org,</div> <div>Genetic Databases/ Knowledgebase Gene, Disease, Ethnicity (germline & somatic), family history analysis tools</div> |

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.

5. Use Case Scenarios

5.1 SCENARIO 1: SPECIMEN IDENTIFICATION

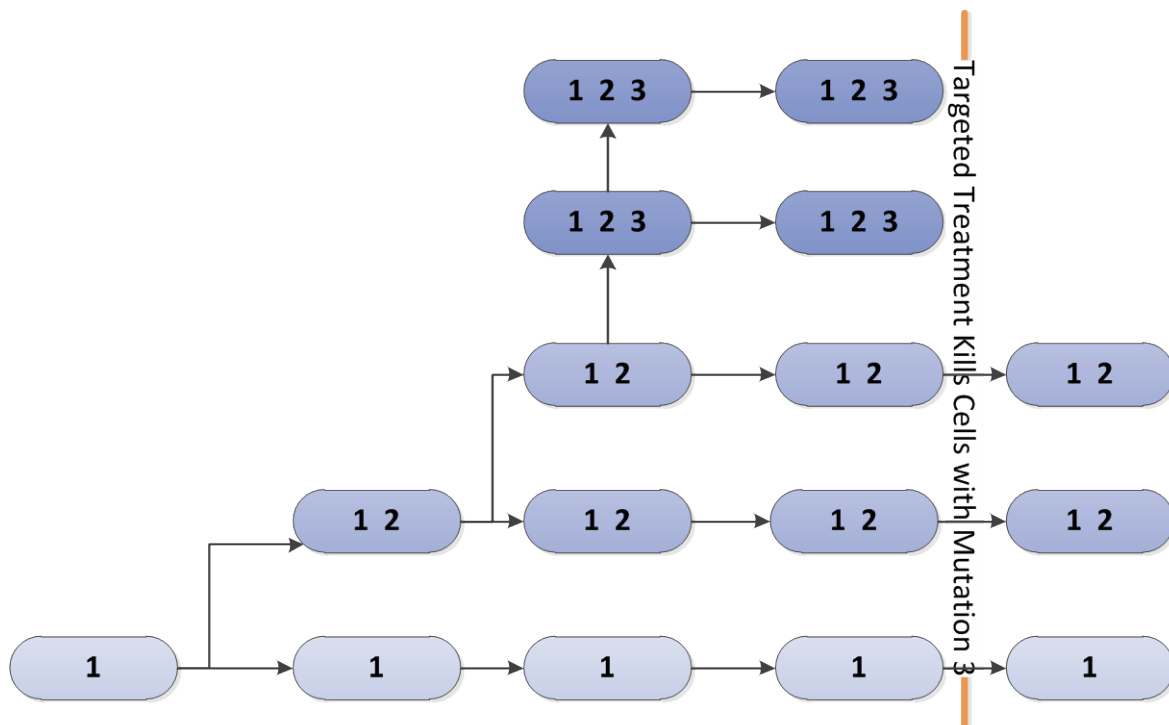
Use Cases for sequencing require explicate identification of 1 or more specimens to be used in laboratory analysis. This likely requires the identification of specimen groups (i.e. separate specimens and associated derivatives) originating from the same patient/subject or related patients/subjects.

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. In order to accurately 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.

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 and

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

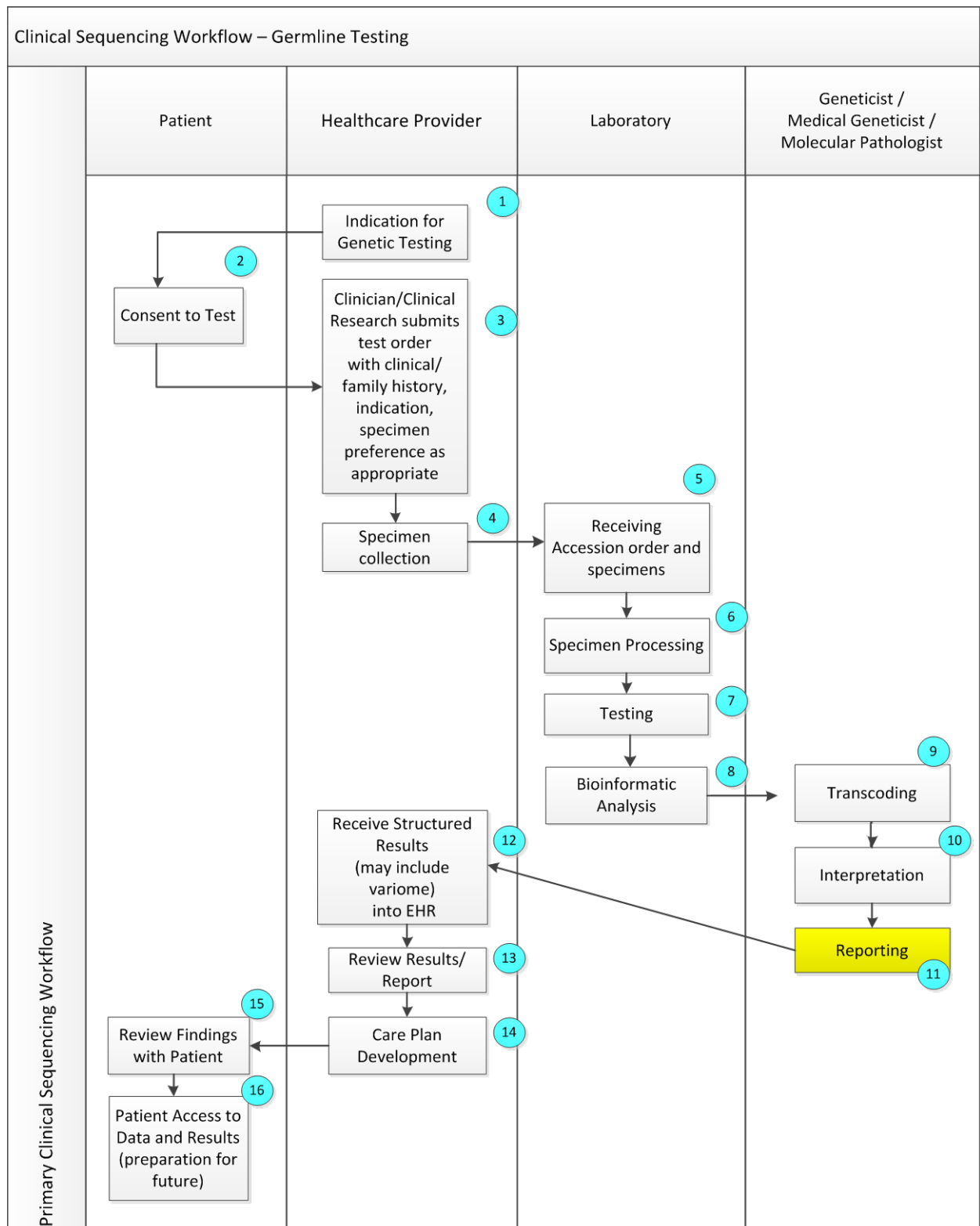
5.1.4 Prenatal testing which may be reported on the maternal medical record (and should be identified as separate from germline testing)

- a. Often have matched 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.

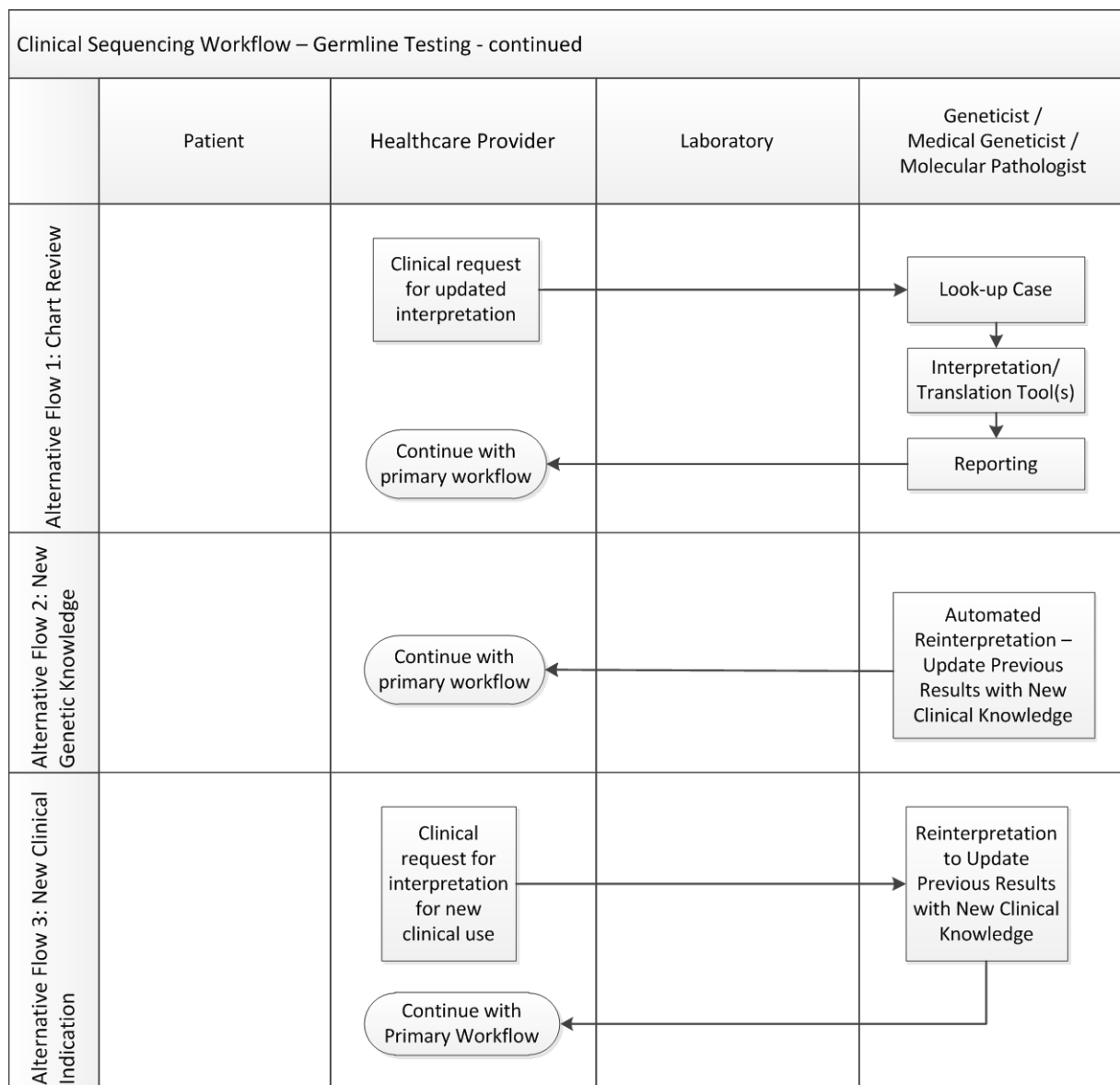
Derivatives which may be analyzed from the above testing scenarios include: DNA, RNA, and Protein

5.2 SCENARIO 2: CLINICAL SEQUENCING – GERMLINE TESTING



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 (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



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).

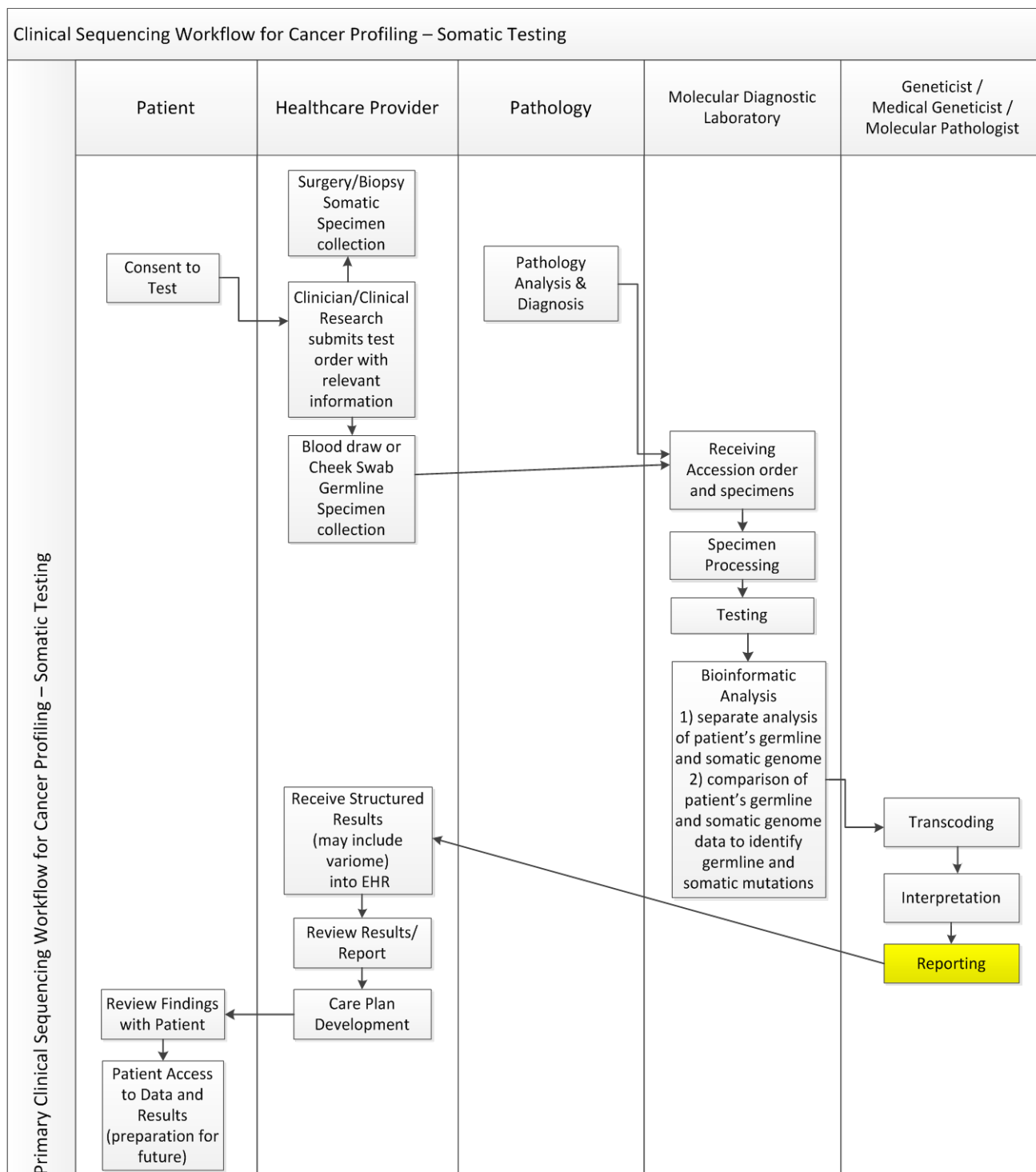
5.2.3 Alternative Flow 2: New Genetic Knowledge

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. (Unfortunately, reimbursement models may need to change to make this feasible.)

5.3 SCENARIO 3: CANCER PROFILING – SOMATIC TESTING

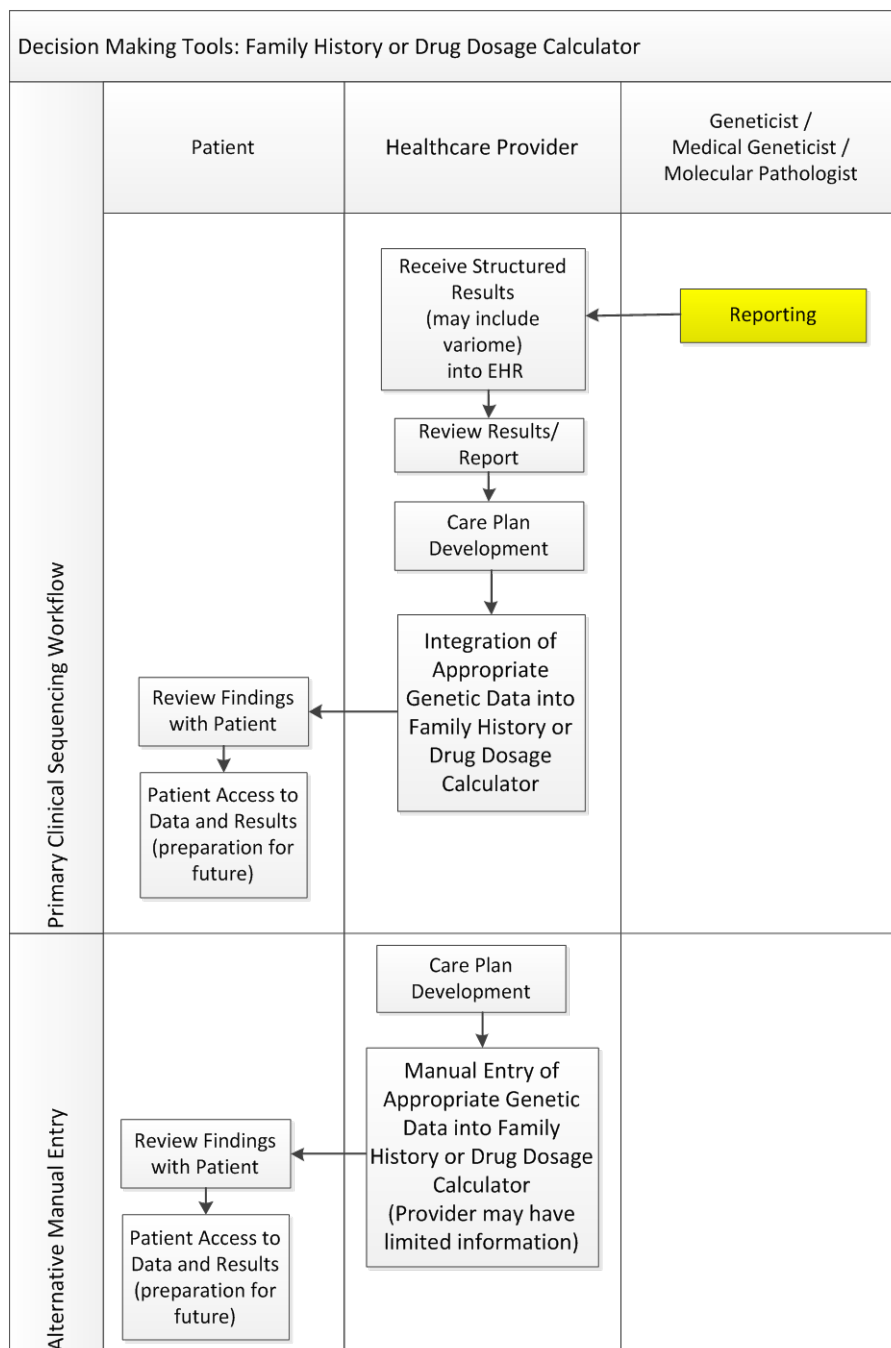


5.3.1 Description of Scenario Differences from Germline Workflow

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

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.

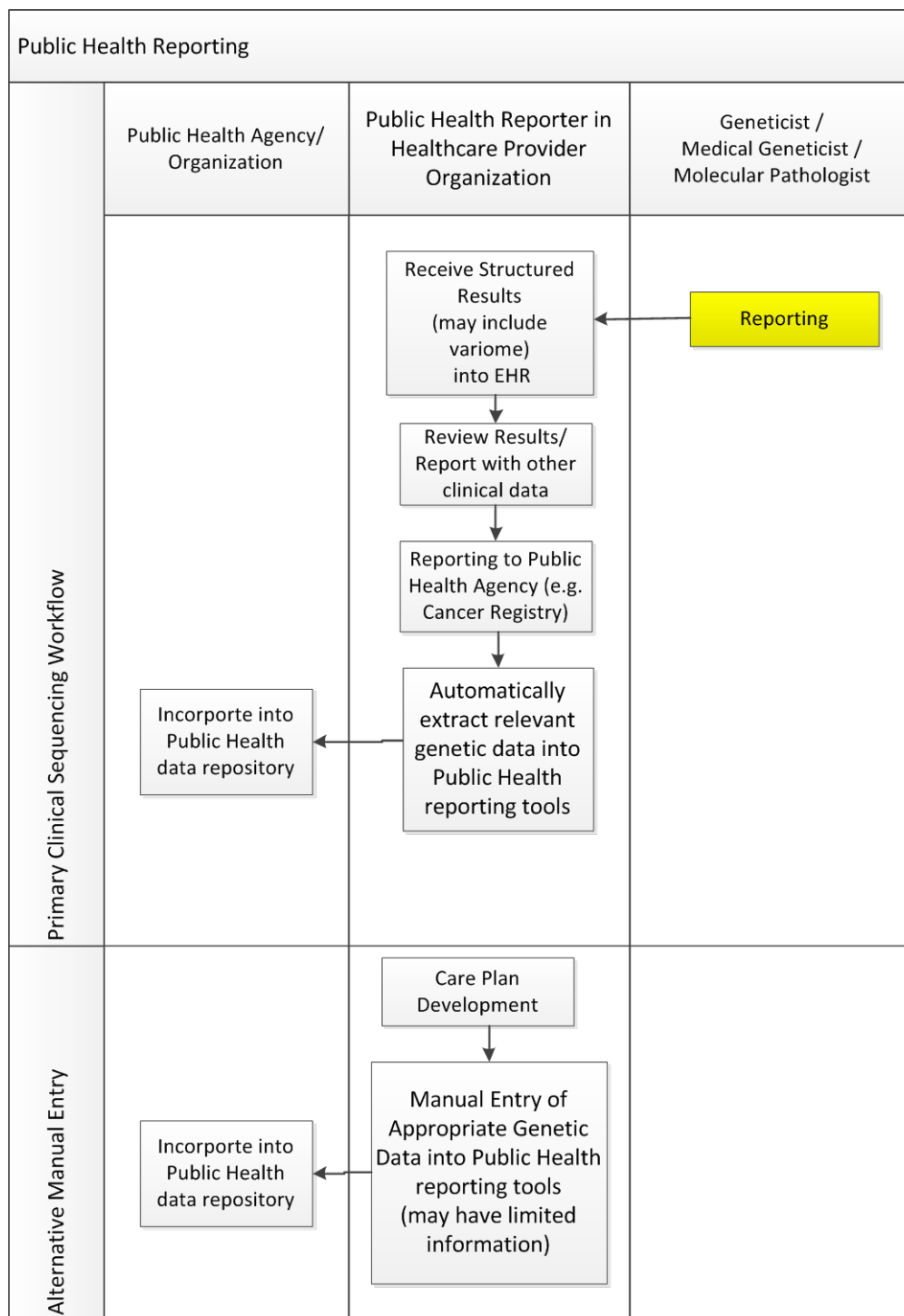
5.4 SCENARIO 4: DECISION MAKING TOOLS – FAMILY HISTORY AND DRUG DOSAGE CALCULATORS



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.

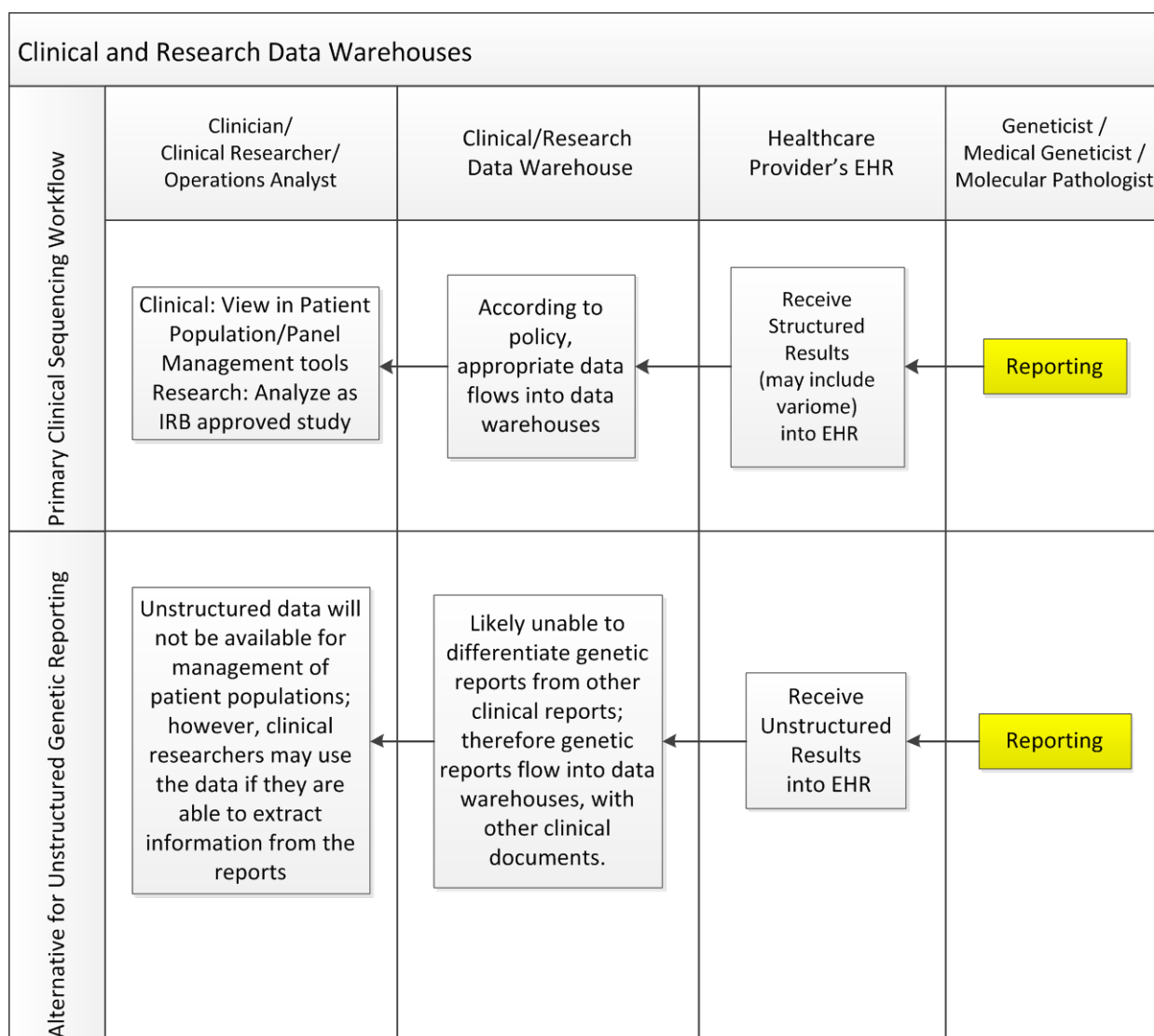
5.5 SCENARIO 5: PUBLIC HEALTH REPORTING



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.

5.6 SCENARIO 6: CLINICAL AND RESEARCH DATA WAREHOUSES



5.6.1 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 PCP may 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 policies. If genetic data is not structured, it doesn't support these activities.

6.Data Set Considerations and HIT Data Standards

6.1 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)

6.2 SEQUENCE VARIATIONS / CHROMOSOMAL CHANGE

6.2.1 Variations within a Gene

HL7 Clinical Genomics Workgroup has a published standard which extends the HL7 2.5.1 Laboratory Reporting standard entitled: HL7 Version 2 Implementation Guide: Clinical Genomics; Fully LOINC-Qualified Genetic Variation Model. This standard structures genetic change (i.e. mutations or sequence variants) within a gene (including gene alleles) and optionally supports reporting of interpretations associated with drug metabolism, efficacy, and confirmatory diagnosis or identification of risk of an inherited condition. This standard uses established clinical genetic standards.

Additionally, CDA has been modeled for support of genetic test reporting which uses the LOINC codes available in the 2.5.1 laboratory report. This standard has passed DSTU ballot and will be published in early 2013.

6.2.2 Other Sequence Variations – Chromosomal Change

HL7 Clinical Genomics Workgroup has a published 2.5.1 Implementation Guide for the reporting of cytogenetic test results using ISCN nomenclature: HL7 2.5.1 Implementation Guide: Clinical Genomics; fully LOINC-Qualified Cytogenetic Model, Release 1. The CDA-based Genetic Test Report also enables inclusion of cytogenetic results, as modeled in the 2.5.1 cytogenetic reporting guide.

6.2.3 Future Enhancements

As part of the Clinical Sequencing project, the standards listed above will be extended to support established bioinformatic representation of DNA and protein changes, as well as additional requirements identified through this project.

7.Gaps & Extensions

7.1 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.

7.2 LABORATORY GENOMIC DATA STANDARDS

High-throughput sequencing has transformed laboratory processes, reducing human analysis of sequencing data and significantly increasing computer based analysis. As such, HL7 Clinical Genomics

must work with these new data standards, to ensure required annotations are either supported within the laboratory systems or can be added after bioinformatic analysis during the transcoding, interpretation or reporting process steps.

8.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?
2. What will future reimbursement models look like? Currently laboratories are reimbursed at a significantly higher rate for actual testing of specimens and reimbursement levels for interpretation of findings is very low, although this can be the most costly step.

9.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.

10.Future Plans

10.1 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).

10.2 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.

10.3 EXTENSION TO SPECIMEN SCENARIOS

10.3.1 Microbiome analysis of the patient

- a. Includes analysis of microorganisms living in the patient's gastrointestinal tract or Genitourinary system and may aid in diagnosis

This page intentionally left blank.