Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering

Sarah Louise Ariansen

Clinician Requirements for Genetic Information in the Electronic Health Record

Master's thesis in Healthcare Informatics
Supervisor: Øystein Nytrø
November 2022

https://genomemedicine.biomedcentral.com

NTNU
Norwegian University of
Science and Technology

Sarah Louise Ariansen

Clinician Requirements for Genetic Information in the Electronic Health Record

Master's thesis in Healthcare Informatics
Supervisor: Øystein Nytrø
November 2022

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering

NTNU
Norwegian University of
Science and Technology

Abstract

Introduction
Genetic medicine is one of the fastest growing fields in healthcare and one component of precision medicine that is increasingly used. Using electronic health records (EHRs) to present genetic information for use in clinical care represents a great opportunity, but the information should be well organised and integrated with the patient's clinical information. However, while the use of genetic testing has expanded rapidly, health information systems have not developed accordingly. The additional complexities of genetic data, compounded by lack of standards, are also considerable challenges. Genetic test results are currently delivered to the EHR in PDF reports, resulting in the genetic information being difficult to find and unavailable for clinical decision support.

Purpose
The objective of this thesis is to define the requirements for the organisation and display of genetic information and investigate the possibilities for representation of this information in EHRs to ensure genetic data is readily available to the users and for clinical decision support. This study will contribute to the body of knowledge on display of genetic information in the EHR by identifying user needs in the field of medical genetics and will aid bridge the genetics-health information systems gap within precision medicine.

Methods
A literature review and a user-centred requirement engineering approach was used to answer the research objective. Through qualitative and exploratory methods such as observations, interviews, and paper prototyping, relevant requirements were identified, focusing on end-user involvement.

Results
Mapping of the current genetic analysis workflow revealed a health information system that is complicated, susceptible to errors and inefficient. The present EHR has no functionality for ordering, displaying, or organising genetic information. The developed prototype illustrated how genetic information could be displayed and integrated into the EHR, allowing the users to easily find relevant test results and genetic consultation notes. The use of structured genetic information allowed for clinical decision support in the form of relevant alerts and warnings and included functionalities for electronic test ordering. Twenty-eight user requirements obtained from multiple methods were identified.

Conclusion
The study identified functionalities important to the organisation of genetic information in the EHR and the prototype confirmed that enhanced organisation and display of genetic information using standardised representations of genetic data can result in highly efficient workflows and increased quality of genetic test results. The results also indicate that health information system developments should be driven by user needs identified in clinical settings and involve close collaboration with end users as opposed to technology driven innovation.

Sammendrag

Introduksjon
Genetikk er i dag ett av de raskest voksende feltene innenfor helsevesenet, og bruken av genetisk informasjon øker spesielt innenfor presisjonsmedisin. For å utnytte genetisk informasjon i klinisk sammenheng, er man avhengig av at den presenteres på en organisert måte og er godt integrert med pasientens kliniske informasjon i den elektroniske pasientjournalen (EPJ). Selv om bruken av genetisk testing har økt de siste årene, har ikke helseinformasjonssystemer (HIS) holdt tritt med utviklingen. Dagens situasjon er at genetiske testresultater leveres til EPJ i PDF-rapporter. Dette resulterer ofte i at den genetiske informasjonen er vanskelig å finne og dermed utilgjengelig for klinisk beslutningsstøtte. Den ekstra kompleksiteten i genetisk data, forsterket av mangel på standarder for presentasjon, gir også betydelige utfordringer.

Hensikt
Målet med denne oppgaven er først og fremst å definere kravene til organisering og presentasjon av genetisk informasjon i EPJ. Dette for å sikre at genetisk informasjon er lett tilgjenglig for både brukere og klinisk beslutningsstøtte. Ved å identifisere brukerbehov innen medisinsk genetikk, vil denne studien øke kunnskapsgrunnlaget om organisering av genetisk informasjon i EPJ og bidra til en økt bruk av genetikk i klinisk sammenheng for å bedre realisere presisjonsmedisin.

Metode
For å svare på forskningsspørmålet ble det først utført en litteraturstudie. Deretter ble det gjennom kvalitative og utforskende metoder, som observasjoner, intervjuer og papirprototyping identifisert brukerkrav med en brukersentrert tilnærming.

Resultater
Kartlegging av dagens arbeidsflyt for genetiske analyser avdekket et HIS som er komplisert og ineffektivt. Dagens EPJ støtter ikke funksjonaliteten som behøves for å bestille, presentere eller organisere genetisk informasjon. Den foreslåtte prototypen illustrerer hvordan genetisk informasjon kan presenteres og integreres i EPJ på en god måte, slik at brukerne enkelt kan finne relevante testresultater. Prototypen demonstrerer hvordan genetisk informasjon i større grad kan tas i bruk i klinisk sammenheng ved å implementere klinisk beslutningsstøtte ved hjelp av relevante varsler, samt tilby elektronisk testbestilling med integrasjon av pasientens kliniske informasjon. 28 brukerkrav ble identifisert.

Konklusjon
Studien identifiserte hvilke funksjoner som er viktige for presentasjonen av genetisk informasjon i EPJ. Prototypen bekreftet at innføringen av standardiserte representasjoner av genetisk informasjon førte til mer effektiv arbeidsflyt og økt kvalitet på klinisk bruk av genetiske testresultater. Resultatene indikerer også at utviklingen av HIS bør foregå med en brukersentrert tilnærmingen, og innebære et tett samarbeid med sluttbrukere, i motsetning til teknologidrevet innovasjon.

Acknowledgements

This thesis marks the end of a three-year part-time study towards an experience-based master's in Healthcare Informatics at the Department of Computer Science, Norges Teknisk-Naturvitenskapelige Universitet (NTNU). They have been educational and interesting years, and at times, demanding alongside full-time work and family life.

There are several people I would like to thank for their help and support during this study and I am fortunate to have family, friends, and colleagues who were able to provide encouragement.

I would like to thank my employer, the Department of Medical genetics, OUH for giving me the possibility to complete this master's degree, and to all my good colleagues for continual support throughout the whole study. Thanks also to everyone who has participated and shared their knowledge and experiences, without you there would have been no study! I would like to thank my supervisor, Øystein Nytrø, for all his time and help in guiding me through this process.

Elsa and Åsmund - Thanks for constructive comments, advice and proofreading and giving me the motivation to finish this thesis on time!

Finally, special thanks to my husband, who has been incredibly understanding and patient, has put up with me during these last three years and made it possible to combine family, work, and studies.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

API	Application
CDS	Clinical Decision Support
CIM	Clinical Information Model
DIKW	Data Information Knowledge Wisdom
DIPS	Distribuert Informasjons- og Pasientdatasystem i Sykehus
DMG	Department of Medical Genetics
DNA	Deoxyribonucleic acid
EHR	Electronic Health Record
EMR	Electronic Medical Record
FHIR	Fast Health Interoperability Resources
GA4GH	Global Alliance for Genomics and Health
HGVS	Human Gene Variation Society
HIS	Health Information System
HL7	Health Level Seven
HPO	Human Phenotype Ontology
HUGO	Human Genome Organisation
IT	Information Technology
LIMS	Laboratory Information Management System
LOINC	Logical Observation Identifiers Names and Codes
NGS	Next Generation Sequencing
NTNU	Norges teknisk-naturvitenskapelige universitet
OMIM	Online Mendelian Inheritance in Man
OUH	Oslo University Hospital
PDF	Portable Document Format
PGX	Pharmacogenetics
SMART	Substitutable Medical Applications and Reusable Technologies
SNOMED CT	Systematised Nomenclature of Medicine Clinical Terms
SWL	Swisslab
TSD	Service for Sensitive Data
UCRE	User-centered Requirement Engineering
UML	Unified Modeling Language
VA	Variant Annotation
VCF	Variant call format
VRS	Variation Representation Specification
VUS	Variant of Uncertain significance
XML	Extensible Markup Language

1 Introduction

Genetic medicine is one of the fastest growing fields in healthcare and one component of precision medicine that is increasingly used. Although genetics and precision medicine promise many opportunities and benefits to healthcare, clinical implementation has proven challenging. This thesis describes an exploratory approach to study the organisation of genetic information in electronic health records (EHRs) and to specify the functionalities required by users of genetic information that can be used to facilitate further innovation processes within health information systems (HISs).

This chapter will introduce the study by first presenting a broad overview of the topic and contextual factors, followed by a description of established challenges. The research problem and objectives will be then introduced.

1.1 Importance of Genetic Information

Precision medicine can transform healthcare by customising treatments to individual patient needs. Aronson et al (S. J. Aronson & Rehm, 2015) state the goal of precision medicine is to "enable clinicians to quickly, efficiently and accurately predict the most appropriate course of action for a patient". The Norwegian Directorate for Health define precision medicine as an important area for innovation which is set to revolutionise the health service but requires advanced equipment, a multidisciplinary approach and a high level of expertise(Directorate for e-health, 2016).

Genomic medicine considers the functions and interactions of all the genes in the genome and is an emerging multidisciplinary specialty that aims to improve human health. In this thesis, the term genetic information will be used to encompass both genomic (multiple genes) and genetic (single gene) information. Over the last few years, the costs of genetic testing have dramatically decreased, and the speed of generating genetic information has rapidly increased, leading to a revolution in genetic analysis in clinical medicine. However, genetics can only improve healthcare if clinicians can identify when genetic information may be useful.

The EHR system is an essential component of any clinician's workflow. Consequently, utilising the EHR to present genetic test results for use in clinical care represents a great opportunity but the information should be organised and displayed in a manner that integrates both the patient's clinical information and the clinician's workflow. In this study I have used the term EHR to describe an EHR system, complete with the user interface, as opposed to an isolated electronic health record.

Up to ten percent of all cancer is hereditary, caused by genetic variants (Cancer Research UK). Integrating genetic information into the EHR can assist clinical diagnosis and the treatment of e.g., hereditary cancer and allows for the potential to provide clinical decision support (CDS). If a patient has genetic testing, variations that cause cancer can be identified early, personalised patient treatment ordered, surgeries can be recommended to reduce the health risks, healthcare providers can be alerted of potentially harmful prescription drug interactions based on the patients' genes, and family members can be notified of hereditary risks.

Most rare disorders are genetic in origin, with young children being the largest patient group, and they are difficult to diagnose based on clinical features alone (Bick, Jones, Taylor, Taft, & Belmont, 2019). Comprehensive genetic testing of these patients

increases the potential for a clinical diagnosis, leading to better care and personalised treatment.

1.2 Current Challenges

The gap between what we can do to improve healthcare with genetic testing and precision medicine and what is actually done continues to widen, and clinical implementation has proven challenging (Hooker, 2021). While using the EHR to present genetic information is a great opportunity, numerous barriers prevent the effective use of the EHR for this purpose.

Current interoperability standards cannot appropriately represent the discrete details of genetic data and the genetic community have yet to reach a consensus on which standards will be most functional for managing genetic information (Carter et al., 2022). The EHR obtains genetic results from laboratory systems, which requires interfaces that span multiple organisations and systems from different vendors but due to this complexity, results are often transmitted as PDF documents, making it difficult to organise results in the EHR. PDFs are designed to be human readable but future use of genetic results would then require a clinician to know a genetic test was performed, know where to find the PDF, and be able to use significant time in searching for this information. Genetic analysis workflows have a poor fit with the perception of "flow", and as genetic data is transferred manually from paper test orders to the laboratory systems and then to the EHR, the lack of interoperability can result in lost information and errors.

The absence of electronic test order systems that would provide functionalities for the inclusion of clinical information and eliminate paper test orders also plays a role in preventing optimal use of clinical genetic information. Electronic test ordering in the EHR is complicated, as the list of available genetic tests is constantly growing and changing. There is a need for a sustainable infrastructure and nomenclature to describe genetic tests and enable comparison between tests intended for the same purpose (Hooker, 2021).

Linking genetic and clinical information from individual patients in the EHR is a critical component of precision medicine as genetic variants must be interpretated in relation to a specific condition. However, EHRs are not ready to send accurate coded clinical and family information to the laboratory together with an electronic test order(Carter et al., 2022).

Patients are also beginning to expect their healthcare providers to understand and use genetics, but referring clinicians, often with limited experience in genetics, struggle to interpret complex genetic results. The scope of testing varies greatly, and genetic test results differ in structure and content between laboratories. Inconsistencies in how genetic data is shared and delivered to the EHR make it difficult for clinicians to locate and interpret genetic information and leads to variations in use. There is currently no consensus on how to represent genetic data in the patient's journal and the consequences of a poor display and presentation can potentially lead to missed diagnoses and reduced quality of care (Paxton, 2022).

1.3 Research Problem

The fast-evolving technology, quantity, and complexity of genetic testing has outpaced HIS and EHR systems ability to be interoperable with genetic workflows. Genetic information has become increasingly complex but it's clinical use is hindered by lack of

integration in the EHR. To realise the full potential of precision medicine, the EHR must also evolve to keep up with developing genetic technology.

Clinicians need well designed and functional HISs to help interpret and apply this new data. Successful development of HISs can increase efficiency and production while reducing medical errors (Andre Kushniruk, Monkman, Borycki, & Kannry, 2015). Many HIS developments take a techno-centric approach, limiting the human and social components in the system which can explain numerous failures (Heeks, 2006). User-centred design, involving the user's knowledge and workflow to design an interface, results in a product that meets the user's needs and leads to improved interactions between technology and people (Teixeira, Ferreira, & Santos, 2012). Genetic experts must be involved in EHR innovation to ensure the needed functionalities are defined and improved.

This master thesis focuses on the integration and display of genetic information in the EHR at Oslo University Hospital.

1.3.1 Research Objective

Designing useful and usable HISs is a major challenge but end-user involvement with design and evaluation has been recognised as being a critical success factor. Using a framework from software requirement engineering to specify user needs, my aim is to elucidate functional requirements in a manner that is understandable for all involved in the development process and which can be used to further facilitate an innovation process.

The objective is to define the requirements for the organisation and display of genetic information and to investigate the possibilities for representation of this information in the EHR, to ensure genetic data is readily available to the users and for clinical decision support.

The following research questions provide a basis for answering the main research objective:

1. What requirements do clinicians have for optimal use of genetic information in the EHR?
2. Which solutions are available to display and integrate genetic information with other patient information in the EHR?
3. What are the challenges of integrating genetic information with other patient information in the EHR?

This study will contribute to the body of knowledge on display of genetic information in the EHR by identifying user needs in the field of medical genetics in which technology and utilisation are rapidly evolving. This will aid bridge the genetics-HIS gap within precision medicine.

1.4 Motivation

I have many years' experiences from the genetic laboratory. I have observed the development of genetic testing from single gene tests that were complicated and time-consuming to perform, to the current high-throughput and effective genomic sequencing tests that we use today. The laboratory and HISs have, in the same period, also changed and developed, but not at the same pace as the sequencing technology, resulting in workarounds and compromises at multiple steps in the analysis. I have, however, little knowledge of how a genetic test report generated in the laboratory, is presented to the

clinicians in the EHR. When I became aware that the clinicians using genetic information, experience severe challenges in their everyday work, I saw it as an opportunity to be able to study a real problem. As an end-user myself of many HISs, I have also reaped the benefits of user-centered design and it was especially motivating to include user centered methods in this study.

1.5 Outline

Chapter 1 provides background information on the role of genetics in precision medicine, and introduces the problems posed in integrating genetic information in the EHR. This chapter identifies the research questions and motivation for this study.

Chapter 2 gives a brief description of the technology associated with a genetic analysis workflow and provides an overview of HISs and genetic interoperability standards.

Chapter 3 is a literature review relevant to the research question. It presents a state-of the-art review of the integration and organisation of genetic information in the EHR. It addresses challenges and solutions in this domain and identifies gaps in the literature.

Chapter 4 outlines the methodology and research design used in the study. This consisted of user-centred requirement engineering, with explorative data-gathering methods including observations, interviews, and paper prototyping.

Chapter 5 presents the results from the qualitative research and includes a list of identified functional user requirements necessary to achieve optimal use of genetic information in the EHR.

Chapter 6 contains a discussion of the results and their relevance to the research questions and existing knowledge. It also outlines the quality of the study and limitations of the work.

Chapter 7 consists of a summary and future work.

2 Background: Genetic Health Informatics

This chapter will briefly describe a typical genetic analysis workflow with relevant genetic terminology and the HISs that are critical for the successful clinical use of genetic information. Emerging interoperability standards are also explained to provide necessary background information for the reader.

2.1 Genetic Analysis Workflow

Delivering genetic test results to clinicians is a complex process. Genetic testing is comprised of multiple steps and involves many actors, requiring all of these to work together to create an optimal workflow. There are many critical and time-sensitive steps, which must be delivered to the referring clinician within an expected turn-around time. (Cutting et al., 2015).

Briefly summarised, most genetic tests start with identification of a patient who will benefit from genetic testing, the referring clinician must then identify a relevant test and place an order. This is often in paper form, with limited clinical information included. The laboratory performs the genetic test, interpretates the findings and return the results to the referring clinician, again as a paper document or as a PDF document in the EHR.

Earlier, genetic testing has typically focused on a single gene analysis, but due to improvements in the technology performance, it is now increasingly common to analyse many genes or all the genes at the same time by next-generation sequencing (NGS). Methods that determine the nucleotide sequence of DNA molecules are called sequencing and today most labs use a high-throughput sequencing technology that enables billions of nucleotides to be sequenced in parallel. The results of the sequencing are returned as short sequencing reads, stored in in text files in FASTQ format. Metadata, the information associated with the produced data, such as the type of sequencing assay, sequencing instruments, quality values or sample type are also collected. (Bernasconi, Canakoglu, Masseroli, & Ceri, 2021).

During data analysis, bioinformatic pipelines are necessary to process the raw data into data values that can be evaluated and interpreted, and to ensure the raw data fulfils the quality requirements. Genetic sequences are computationally reconstructed by using overlaps between the short sequencing reads, and aligned to a reference genome, which produces binary alignment map (BAM) files. The differences between the sequenced DNA and the reference can be identified by performing variant calling, which produces variant call format (VCF) files.

Post-processing activities include marking specific features in the DNA sequence with descriptive information about structure or function. These annotations can include the addition of positional information for each DNA variant, gene descriptions and descriptions of variant types. Annotations can also include external identifiers that point to databases containing genetic resources. This metadata is organised in a structured format but can vary from laboratory to laboratory in different levels of granularity and entity descriptions (Bernasconi et al., 2021).

Variant interpretation is a complex process where many pieces of information, often inconclusive and sometimes contradictory, must be pieced together to reach an answer. The laboratory studies the variations found in a patient's DNA to differentiate between normal variation and disease-causing variants. Interpreting a DNA variant involves many thoughts of reasoning, e.g. the frequency of the variant (we assume that frequent variants are not disease causing), assessment of the chemical properties and expected biological impact of the changes produced by the variant, and accessing information about reported associations between variants and disease from research literature ("IT6103," 2019). The interpretation of DNA variants is often the major bottleneck for processing of the patients' genetic test results.

An informative report for the referring clinician, including information on the assessment and potential impact of the genetic variants and their association with disease interpretation is generated. The laboratory should also inform referring clinicians if new, relevant information emerges on a previously reported variant (S. Aronson et al., 2016).

To integrate genetic data into the clinical record requires an Information Technology (IT) infrastructure which includes a Laboratory Information System to support test orders, data collection and reporting, an interpretation tool to assist in translating results into clinical implications and an EHR to accept and utilise genetic data (Ullman-Cullere, Clark, & Aronson, 2018), as illustrated in figure 2.1.

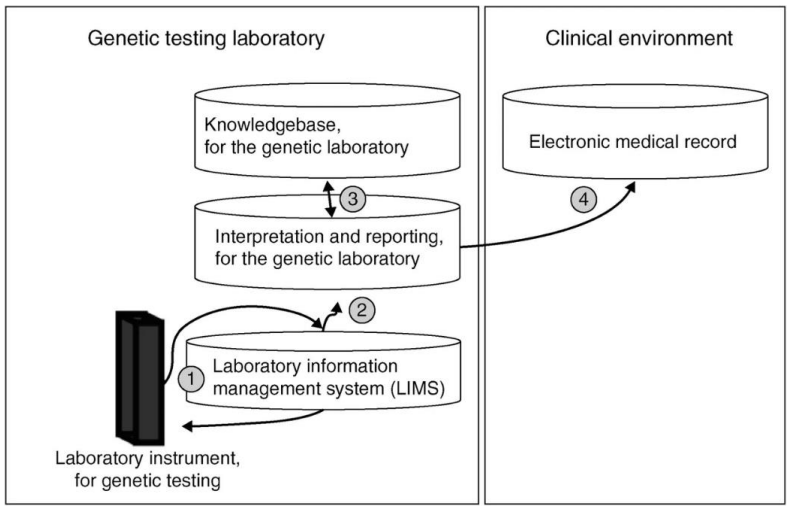


Figure 2.1 IT infrastructure for reporting of genetic test results (Ullman-Cullere et al., 2018)

2.2 Health Information Systems

A health information system (HIS) describes a system designed to manage healthcare data e.g., EHRs, laboratory information management systems (LIMS), CDSs or patient portals. HISs have existed for many years as a tool to improve quality and effectivity in health care and their role extends significantly beyond the simple storage and retrieval of patient data.

Healthcare depends increasingly on HISs, whose introduction is often characterised by limited implementation or rejection. It has become apparent that a social-technical approach when implementing HISs may lead to a higher success rate (Greenhalgh, Stones, & Swinglehurst, 2014). This requires looking at the whole picture, observing the users involved, mapping their needs, identifying what the technology can assist with and

planning what changes will be necessary. Although HISs are composed of machines, devices and hard physical technology, they require substantial social, organisational and intellectual investments to make them perform optimally.

2.2.1 Laboratory Information System

A laboratory information system (LIMS) is a software system that records, manages, and stores data for clinical laboratories. The LIMS provides the framework for the laboratory to exert control over the testing process and is often based on relational databases (McCudden, Henderson, & Jackson, 2020). Typically, information from each laboratory testing process will be transferred to the LIMS including preanalytical steps such as ordering and labelling, analytical steps with results, and finally, report generation and billing.

Khalifa (A. Khalifa et al., 2021b) describe a generic model for genetic laboratory systems where the lab uses a central LIMS that receives the lab order, the raw test results are obtained, then the test results are sent to another specialised system for interpretation. The final report is uploaded to LIMS and transmitted to the EHR.

2.2.2 Electronic Health Record

The EHR is defined by the International Organisation for Standardization (Standardization, 2019) as: " a data repository regarding the health and healthcare of a subject of care where all information is stored on electronic media. Its primary purpose is the support of continuing, efficient, and quality integrated health care.". An EHR is patient-centred, longitudinal (from birth to death), comprehensive (includes a record from all institutions) and prospective (includes future plans/orders).

An electronic medical record (EMR) is considered an internal organisational system whereas an EHR is inter-organisational (Heart, Ben-Assuli, & Shabtai, 2017). In this study, the term EHR is used to encompass both types of medical records.

The EHR represents the sum of information and communication systems available for clinical care and is the single point of deposition and access for nearly all elements of clinical data(Coiera, 2015). EHRs are suitably positioned to be pivotal in genetic information-technology support, as clinicians have access to all of a patient's data, including genetic information. However, this is dependent on the information being organised and displayed in a way that integrates with the clinicians workflow and facilitates diagnostic and treatment decisions (S. J. Aronson & Rehm, 2015). Currently most results in genetic reports and information are transmitted from the laboratory as a PDF file or a text block, with no opportunity for reliably extracting structured data (Shirts et al., 2015).

2.2.2.1 DIPS

After several years of development within Nordland hospital, the journal system,DIPS (Distribuert Informasjons og Pasientdatasystem i Sykehus) Classic was implemented in 1992. DIPS AS have now an agreement with three of the four health regions in Norway, and is the medical record system for 4.3 million patients and 150,000 health employees (AS) in Norway.

DIPS Arena, a new version of the journal system, is built on an open technology platform and international standards. The new journal system will support advanced and complicated patient routes through improved interaction and data-sharing. DIPS Arena is

still in development and is only available in certain hospital departments. The version in use at OUH, Dept. of medical genetics is based on DIPS Classic.

2.2.3 Clinical Decision Support

A clinical decision support (CDS) system can be defined as a health information technology system, designed to enhance decision making by providing health professionals with relevant information and data to assist with clinical decision-making (Greenes, 2014). The focus is specifically computer-based support in aiding decisions, which infers that the system is dependent on a medical professional in the process. They are primarily used at the point-of-care, for the clinician to combine their own knowledge with information or suggestions provided by the CDS and help providers answer certain questions such as what tests are most appropriate or which treatment would be the best ("IT6122," 2020).

A CDS system, at its simplest, is composed of a knowledge-base where both available data and rules/logic expressions and constraints are programmed, an inference engine which takes the programmed rules and data structures and applies them to a patient's data to generate an output or action, which is presented to the clinician through a communication interface. See Figure 2.2

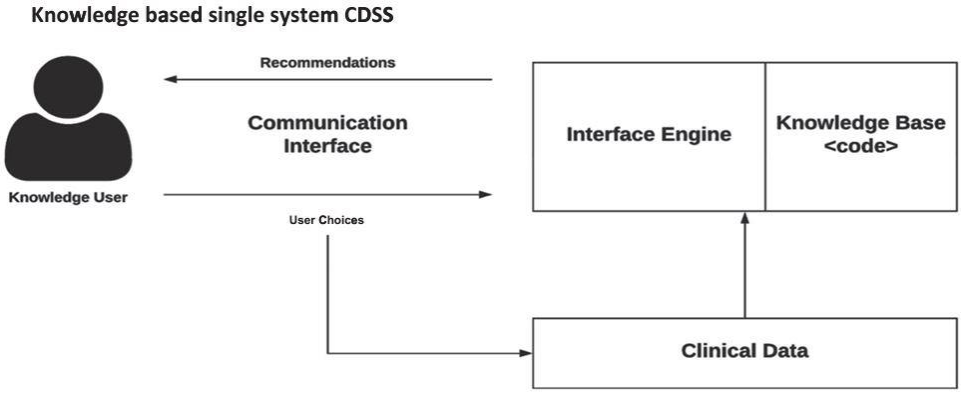


Figure 2.2 Key interactions in a CDS (Sutton et al., 2020)

A well-functioning CDS system can lead to improved patient safety, improved quality of care and the reduction of prescribing and dosing errors. Studies have shown that CDS systems can increase adherence to clinical guidelines and alert clinicians to reach out to patients who have not followed management plans. (Sutton et al., 2020). It is impossible for a clinician to have an overview over all the available knowledge that can be relevant in any given situation – CDS is a method to spread such knowledge, in a manner that is more advanced than distribution of pure text.

A CDS system can have a wide variety of functions, including alerts and reminders, continual surveillance of laboratory results, computerised guidelines, order sets, documentation templates, treatment planning, prescription control, information searches and image analysis (Sutton et al., 2020). An optimal CDS for genetic data should consist of both active and passive components (Sen, Al Kawam, & Datta, 2019), where the active components could provide alerts. Other tools such as Infobuttons can provide context-sensitive links embedded within the EHR, which act as a form of passive CDS to deliver targeted and relevant information resources to a clinician such as current guidelines and test interpretations.

Masys et al (Masys et al., 2012) present seven key principles in their technical desiderata for storing genetic data in the EHR, developed further by Welch et al (Welch, Eilbeck, Fiol, Meyer, & Kawamoto, 2014) to also include the support for CDS. These 14 principles are summarised in Table 2.1.

Table 2.1 Desiderata for integration of genetic information into EHR (Welch et al., 2014)

1	Maintain separation of primary molecular observations from the clinical interpretations of those data
2	Support lossless data compression from primary molecular observations to clinically manageable subsets
3	Maintain linkage of molecular observations to the laboratory methods used to generate them
4	Support compact representation of clinically actionable subsets for optimal performance
5	Simultaneously support human-viewable formats and machine-readable formats in order to facilitate implementation of decision support rule
6	Anticipate fundamental changes in the understanding of human molecular variation
7	Support both individual clinical care and discovery science
8	CDS knowledge must have the potential to incorporate multiple genes and clinical information
9	Keep CDS knowledge separate from variant classification
10	CDS knowledge must have the capacity to support multiple EHR platforms with various data representations with minimal modification
11	Support a large number of gene variants while simplifying the CDS knowledge to the extent possible
12	Leverage current and developing CDS and genetics standards
13	Support a CDS knowledge base deployed at and developed by multiple independent organisations
14	Access and transmit only the genetic information necessary for CDS

The combination of the above requirements is essential for the integration of genetic information within EHRs using CDS.

2.3 Knowledge Representation

The data-information-knowledge-wisdom (DIKW) hierarchy, also known as the knowledge pyramid or wisdom hierarchy, is a widely recognised model in information and knowledge studies (Rowley, 2007). The hierarchy is used to contextualise data, information, knowledge, and wisdom with respect to each other and to describe the processes involved in the transformation from level to another. It has gained importance in public health with the increasing use of large amounts of data and highlights that data cannot be used by itself but needs to be transformed into information and knowledge and eventually wisdom.

The use of genetic information can also be aligned with the DIKW model (Timothy M. Herr et al., 2015), resulting in a genetic funnel, see figure 2.3.

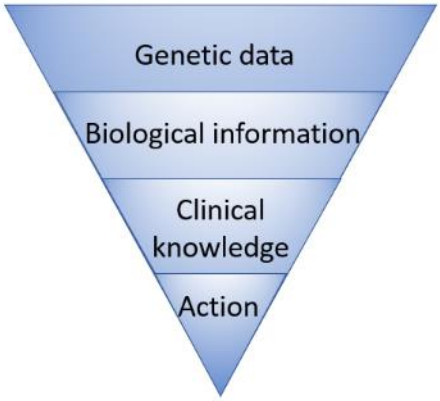


Figure 2.3 The genetic funnel (Timothy M. Herr et al., 2015)

The Genetic Data layer is represented by the millions of raw sequence reads from high throughput sequencing platforms which must be filtered and processed before it can be acted upon.

The Biological Information layer contains information about an individual's genome and how it varies from a reference sequence with a description of e.g., single-nucleotide variants or copy number variants. Many of these variations have no validated significance and are rarely helpful in the clinical setting, so the information again must be filtered for it to be useful. The transition from data to biological information requires published research into the nature of genes and proteins.

The Clinical Knowledge layer contains information that has clinically relevant associations. The number of clinically significant variants is currently small but growing (Campbell et al., 2019). Our understanding is also changing overtime and clinical recommendations need to be continually updated, resulting in provider organisations having to maintain a current and accurate knowledge of the field. This demand leads to the development of central repositories for evidence-based guidelines e-g. the Clinical Pharmacogenetics Implementation Consortium[1] for clinically actionable gene-drug interactions and Clingen[2] for clinically relevant genetic variants. Pharmacogenetics (PGx) is the study of how genes affect a person's response to medicines. This knowledge can be represented in a variety of formats, either a text report or preferably as discrete data elements in the EHR through a recognised genetic standard. Inferring clinical actionability represents a severe bottleneck in the process, with extensive manual curation.

The Action layer represents methods by which clinical knowledge is translated to the point of care and applied to change a clinician's choices, often in combination with other clinical factors to personalise care e.g., a patients genetic information together with weight and smoking status to individualise dose level prescriptions. However, this clinical information is typically only found in text notes in the EHR, unavailable to CDS.

Current EHRs often only show a snapshot of the state of the individual patient, representing the data layer. Longitudinal data on individual patients and interpretation of this, forms information. By interpreting data over time, future EHRs could provide support for correlating clinical features with genetic testing over time. Developing insights that add to knowledge about similar patients, will lead to improved clinical

[1] https://cpicpgx.org/guidelines/
[2] https://clinicalgenome.org/

guidelines, and ultimately new care processes in response to the gained knowledge and to the final layer of the pyramid, action.

Clinicians have historically depended on their own experiences and knowledge, together with knowledge and experiences from relevant health institutions and other specialists, to make clinical decisions, often without the necessary documentation of health effects. This could result in patients with similar symptoms, being treated quite differently.

However, a seamless integration between the EHR and for example, CDS is necessary to achieve evidence-based medicine. In addition to the diversity that exists between health information systems, medical system interoperability has been recently documented as a significant bottleneck for the sharing of medical information and data, and procurement strategy templates targeting this need have now been proposed in USA (Pronovost, 2018). A carefully designed CDS is a critical component that will make it possible for a clinician to understand and use the otherwise overwhelming amounts of genetic information, but to make the transition from clinical knowledge to action requires data sharing and medical system interoperability.

2.4 Interoperability

Interoperability is critical for achieving the full potential of digitalisation in healthcare and medicine but awareness of the topic is relatively low among healthcare professionals and progress is slow (Lehne, Sass, Essenwanger, Schepers, & Thun, 2019). Within health care, interoperability has been defined by the Healthcare Information and Management Systems Society, an international organisation with a mission to reform global health through the power of information and technology, as "the ability of different information systems, devices and applications to access, exchange, integrate and cooperatively use data in a coordinated manner, within and across organisational, regional and national boundaries, to provide timely and seamless portability of information and optimise the health of individuals and populations globally" (HIMSS, 2022).

Most definitions distinguish between several levels of interoperability, and although they can differ slightly between definitions, they generally follow a framework ranging from low-level technical components to higher-level organisational components:

Foundational/technical: Ensures basic data exchange between systems, e.g. moving data from a memory stick to a computer. Technical interoperability establishes the inter-connectivity requirements needed for one system or application to securely communicate data to and receive data from another but to process the data and extract useful information, the systems would require syntactic and semantic interoperability.

Use-case/Workflow: Constraining of the sequence of messages exchanged, which coordinates the work processes across healthcare systems between laboratory instruments and laboratory information systems.

Structural/Syntactic: Syntactic interoperability specifies the format, syntax and structure of data exchange e.g., an extensible markup language (XML) document, including at the data field level for interpretation. XML is a simple text-based markup language that defines a set of rules for representing and describing documents in a format that is both human and computer readable.

Semantic: Full semantic interoperability is achieved using medical terminologies, nomenclatures, and ontologies. Semantic interoperability provides for common

underlying models and ensures that systems understand data in the same way, and unambiguous use and interpretation of the data. This level of interoperability brings, in addition to the standardisation and formatting of health data, the possibility of inferring based on data.

Organisational: Includes governance, legal and organisational considerations to facilitate the secure and seamless communication and use of data both within and between organisations. These components enable shared consent and integrated end-user workflows.

The definitions of these specific levels of interoperability are techno-centric because they are primarily defined by how information is sent, received, and stored by the information technology systems. However, interoperability can also be characterised based on a user-centric framework. Guarrear eta l (Guarrera et al., 2014) describe levels of interoperability based on the methods by which people, together with technology, accomplish work activities. Using observation data from clinical staff, workflows involving different technologies were characterized, resulting in 7 levels of user-centric interoperability. See Table 2.2.

Table 2.2 User-centric interoperability (Guarrera et al., 2014)

Level	Level definition	Examples
Level 0	No form of HIS is used to perform the task and no additional information is accessed or recorded via HIS.	Paper charts, Manual whiteboard
Level 1	Additional information related to the task is available through a form of HIS, but only with view/read-only access; no data entry or modification of the data is permitted.	Lab results entered into own laboratory system
Level 2	The information in a database (HIS) is entered by an office but can only be read within the office.	Stand-alone EHR system with view-only access of additional information, Online reports printed, then scanned into local EHR
Level 3	Information may be exchanged electronically between offices (e.g., email), but a given HIS does not have the capability to exchange information with another HIS .	Information is e-faxed or emailed but as attachments and must be manually inputted by a person into the EHR
Level 4	Information is forwarded to the correct HIS, but goes into the system as ˜uncategorised" the system has no provisions to correctly file the information.	Outside lab results are forwarded to an EHR, but go into a "miscellaneous" tab instead of the "lab result" tab
Level 5	Information is forwarded to the correct location within the HIS, but the meaning of specific data is not defined.	Weight (pounds) and weight (kilograms) are both displayed in the correct location in the EHR
Level 6	Information is forwarded to the correct HIS and within the correct location of the HIS, and the information content is defined.	Abnormal laboratory results are flagged

This framework shifts the focus from technology capability to performance and impact on the human-technology system. The different levels of interoperability can be compared to levels of automation e.g., from no automation (paper only) to automatic transfer and categorising of data, and finally automatic interpretation of information. However higher levels of interoperability aren't necessarily the most desirable for the end-user and particular tasks maybe designed where a clinician is required to review information

before integration into a patient's health record. Higher levels of interoperability may also result in disconnected tasks which do not fit into the workflow, e.g., re-typing of information from one system to another, were perhaps working at a lower level in the framework until technology supports direct transfer is more desirable.

2.5 Healthcare Terminologies and Classification Systems

Healthcare terminology standards address a fundamental requirement for effective communication and provide the ability to represent concepts in a precise manner (HIMSS, 2022). They offer a common language for representing and communicating medical information and are vital to facilitate interoperability among healthcare organisations. Standardisation and terminologies can provide the ability to share and reuse knowledge once it is created (Kawamoto & Greenes, 2014).

To achieve a high degree of semantic interoperability, health data must be standardised, and systems that can be easily expanded and adapted to changing user needs be in place. The health domain is in constant development, and it is necessary to be able to expand information systems as to keep in line with the rapid increase of huge amounts of data. Healthcare interoperability relies on health information standards that allow the different concepts of interoperability to be put in practice (Gansel, Mary, & van Belkum, 2019).

The need for standards in e-health is great and is driven by, especially within genetics, of the increased amounts of data generated. International standards facilitate faster and more efficient development of coherent health and care services. In Norway, the use of international standards will be a requirement for future interaction and the Directorate of e-health has prepared a plan for use of international standards which describes measures and recommendations on the use in priority areas (Directorate for e-health, 2021).

The Directorate for e-health has also established the Common Health Language (Felles Språk) with the goal to simplify and improve the interoperability of Norwegian data in healthcare. The Directorate define a common health language as an ecosystem of health-related terminology, classifications and variables in context which will be used in health information systems(Directorate for e-health, 2019). The common language consists of interconnected terminologies, administrative code lists, medical classifications, and health registry variables, with the goal to ensure that information utilised during patient treatment can be reused after being initially registered.

Using standards to represent and transfer the content of a genetic test report should be relatively straightforward as many computational tools are used in the analysis and interpretation of the results but can also be challenging with regard to the requirement of tailored reports. Genetic results often have unclear or changing interpretations, which requires the communication of uncertainties associated with the result (A. Khalifa et al., 2021a). Genetic data is also generated by several different laboratory methods, as shown in figure 2.4, with data often stored in noncompatible data structures, with genetic variants described with different nomenclatures (Alterovitz et al., 2020).

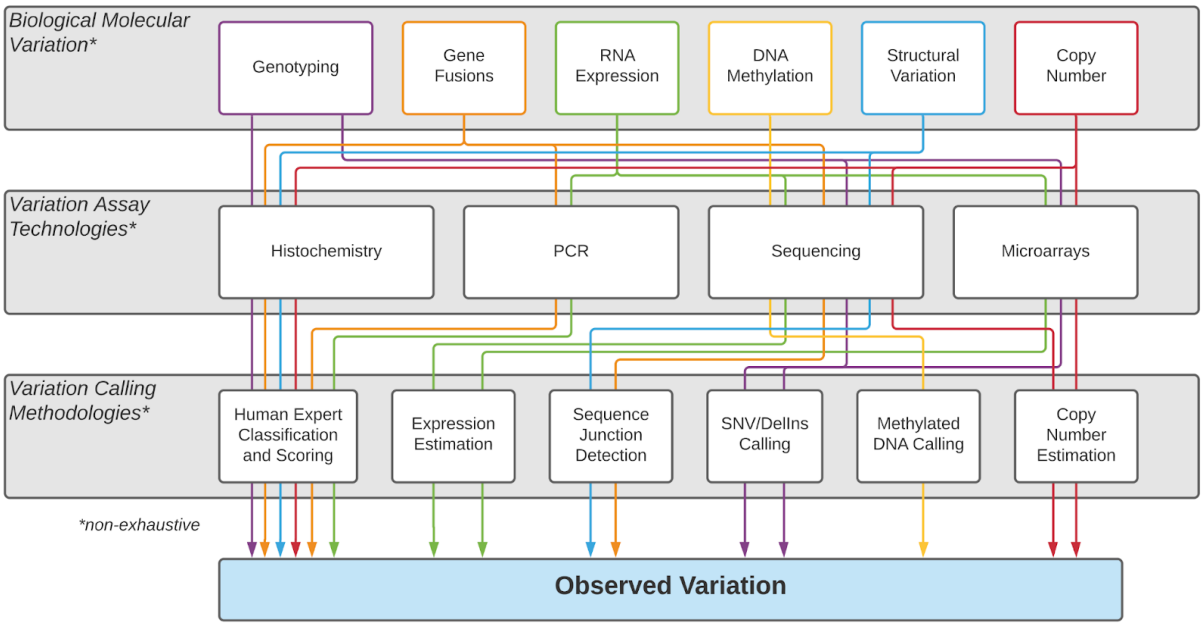


Figure 2.4 Diversity of genetic information sources (GA4GH, 2021)

There are well-established nomenclatures for the representation of genetic variants e.g. Human Genome Organisation Gene Nomenclature for gene names and Human Genome Variation Society (HGVS) for the description of small DNA variants, but implementation of these standards in the clinical domain is not complete (Warner, Jain, & Levy, 2016). One genetic variant can be described in a single database with multiple names due to lack of a universal standard for genetic variation descriptions, see figure 2.5. Although the distinctions may seem minimal, the downstream implications for interoperability are significant.

NM_000016.6(ACADM):c.985A>G (p.Lys329Glu)
NC_000001.11:g.75761161A>G
NM_000001.10:g.76226846A>G
K304E
985A>G
985A>G (K304E)
c.985A>G (p.K304E)
c.985A>G (p.Lys304Glu)
p. Lys304Glu
LYS304GLU

Figure 2.5 Nomenclature for one identical genetic variant in Clinvar

Pharmacogenetic variants, however, use again a different nomenclature system, star alleles, which raise many challenges. A "star allele" is an identified unique gene sequence and unique variant change, e.g., CYP2C19*27, quite different from a HGVS nomenclature description shown in Figure 2.5.

In addition, there are multiple tools, databases and other resources that are used to capture clinically relevant evidence or knowledge about a specific genetic variation which presently require human resources to extract relevant information. Therefore, flexible and effective standards are required to integrate genetic results from various formats with other clinical data in the EHR and external resources and to provide mechanisms for scalability.

The Directorate for e-health(Directorate for e-health, 2019) use the following definitions for several key concepts within interoperability standards:

- Terminology/Codework - a collection of concepts with associated codes, and that the concepts are put in a context with each other.
- Classification - a collection of unique concepts and associated codes in meaningful hierarchies
- Ontologies -a formal representation of a set of concepts within a field of knowledge.

Ontologies include classes (terms or concepts), instances (particular things), relationships among these things, properties of the things, functions and processes, and constraints and rules involving the things. They are expressed in a logic-base language e.g., OWL[3], so that accurate, consistent, and meaningful distinctions can be made. Ontologies can be used to build computer-interpretable semantic representations of domain knowledge and by using formal logic-based specifications, domain concepts and their attributes and relationships can be expressed. This allows computers to make inferences on assertions in a domain model or a knowledgebase. Ontologies speed up the implementation of interoperability due to the availability of robust tools and frameworks that promote reuse and they provide a flexible approach to integrating data and sharing meaning (Liyanage, Krause, & De Lusignan, 2015). Figure 2.6 shows that classifications and ontologies have fundamentally different properties but are also complementary to each other.

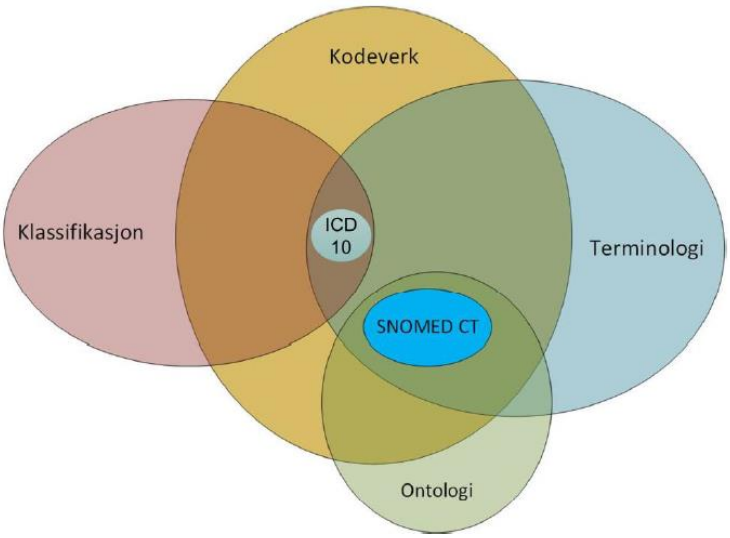


Figure 2.6 Differences and similarities between codeworks, classification, terminologies and ontologies (Directorate for e-health, 2019)

[3] https://www.w3.org/OWL/

SNOMED CT (Systematised Nomenclature of Medicine - Clinical Terms) is an example of a general-purpose terminology for advancing semantic interoperability, consisting of more than 350,000 medical concepts, and often combined with other domain-specific terminologies eg. ICD-10, Logical Observation Identifiers Names and Codes (LOINC) or Human Phenotype Ontology (HPO) (SNOMED International), shown in figure 2.7 and Table 2.3. SNOMED CT will be central in the Norwegian Common language and will connect all existing terminologies that are in use side by side in present health care settings to enable improved information transfer.

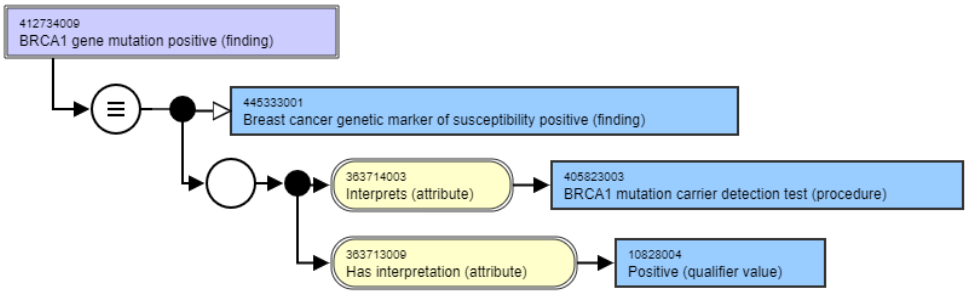


Figure 2.7 Example of a SNOMED CT coding for a BRCA1 variant (SNOMED International)

Table 2.3 Terminology standards relevant to genetic testing

Name	Description
LOINC[4]	Universal code system for identifying laboratory tests, health measurements, observations, and documents. https://loinc.org/
HGVS[5]	DNA sequence variant nomenclature, authorised from HUGOhttps://varnomen.hgvs.org/
HPO[6]	Standard vocabulary for phenotypic abnormalities in human disease
HGNC[7]	Unique symbols and descriptive names for human genes,
Norsk laboratoriekodeverk[8]	Based on the international NPU terminologies, codework for Norwegian laboratory analyses
ICD-10[9]	Medical classification system from World Health Organisation, containing codes for diseases, symptoms and causes of diseases, global reference base for morbidity and mortality statistics,

2.6 Clinical Information Models

Terminology standards are used to structure clinical information, but Clinical Information Models (CIM's) are required to maintain the consistency of clinical information structures

[4] https://loinc.org/
[5] https://varnomen.hgvs.org/
[6] https://hpo.jax.org/app/
[7] https://www.genenames.org/
[8] https://www.ehelse.no/kodeverk/laboratoriekodeverket
[9] https://icd.who.int/browse10/2019/en#/

within a health information system and enable semantic interoperability across different systems and organisations. CIM's are structural and semantic artefacts that facilitate organising, storing, querying, and displaying clinical data; exchanging that data between different information systems and performing data analytics (Moreno-Conde et al., 2015). A CIM is defined by an underlying reference model that provides the attributes needed to represent data instances or clinical terminologies. Complete semantic interoperability is only achieved by using both the reference model and terminologies to describe the semantics of the information structures. See figure 2.8.

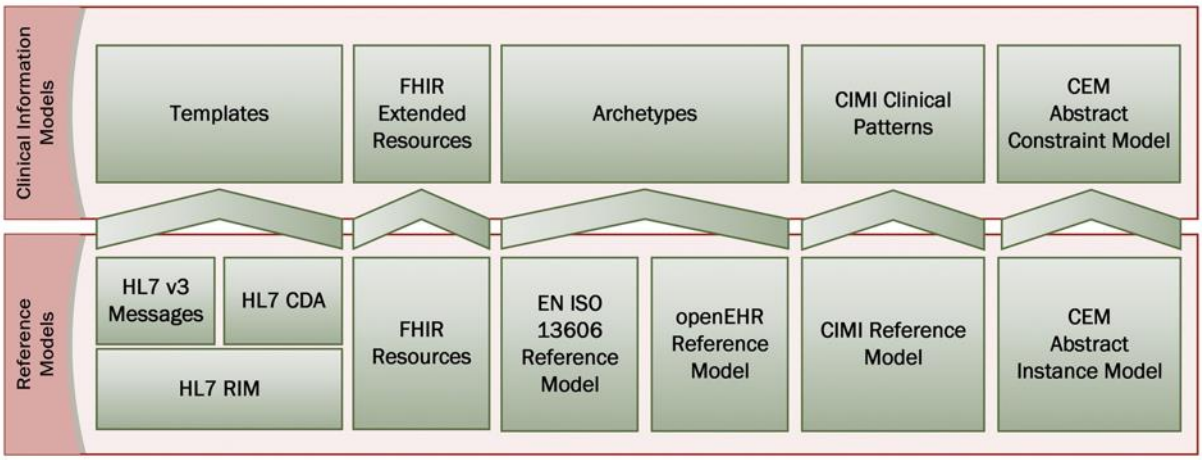


Figure 2.8 Reference models and CIM artefacts (Moreno-Conde et al., 2015)

Earlier terminologies and standards have been approached in a top-down manner, with a minimum of user influence (Ulriksen, Pedersen, & Ellingsen, 2017). This has resulted in complex systems that do not support a flexible workflow and led to challenges with implementation due to the heterogeneity of HISs, with users having to frequently make substantial changes to their workflows and the introduction of workarounds. As an example, in the United Kingdom, the National Health Service Commissioning Board spent more than £12bn on a one-size fits all EHR system, which was eventually discarded and replaced by a new system driven by local decision making(NHS, 2015). Already in 2009, Coiera questioned the feasibility for large complicated top-down implementations and suggests an approach that accommodates the stakeholders to have different starting points, goals and resources (Coiera, 2009). To achieve more flexibility in the standardisation process and to provide standards that accommodate existing practices, more bottom-up processes are being introduced, where the users are given extensive control earlier in the process.

The Clinical Information Modelling process (figure 2.9) requires that technical and clinical experts cooperate to define clinical needs that satisfy levels of a single organisation, a health region or a country, and once defined are then governed to ensure correct managements and development. The modelling process includes a domain analysis, which is essential to ensure a comparable quality and homogenous design of CIMs created by different organisations (Moreno-Conde et al., 2015).

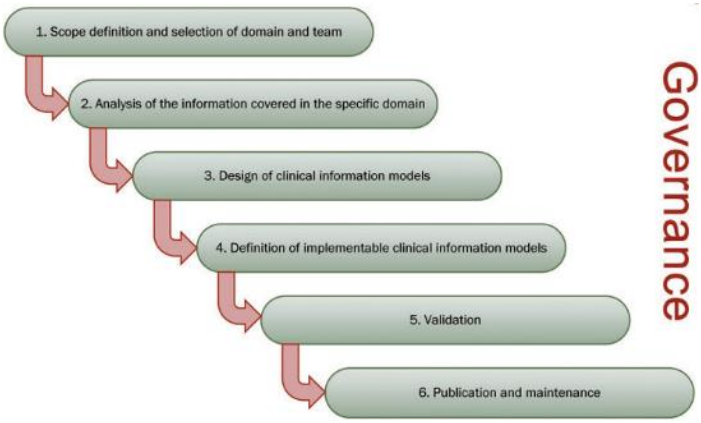


Figure 2.9 Clinical information modelling process steps (Moreno-Conde et al., 2015)

This process is a continual development as there will always be new content and workflows to consider and a final model or platform can be considered a living process, rather than a product.

As the use of high throughput DNA sequencing continues to dramatically increase, so does the need for genetic data interpretation and integration of results into the EHR. The demand for genetic data standards is high and there are several international organisations in the process of developing and improving existing models for genetic data interoperability standards.

OpenEHR, HL7 and GA4GH are all examples of CIMs that are under development for genetic information integration into the EHR.

2.6.1 OpenEHR

OpenEHR is a standards framework for e-health consisting of open technical specifications and clinical models that define an OpenEHR platform for healthcare, constructed by the OpenEHR Foundation (OpenEHR, 2022). The OpenEHR Foundation, established in the UK in 2000, describe OpenEHR as a semantically enabled, vendor independent, health-computing platform.

OpenEHR is based on a multi-level approach that separates knowledge from information. See figure 2.10

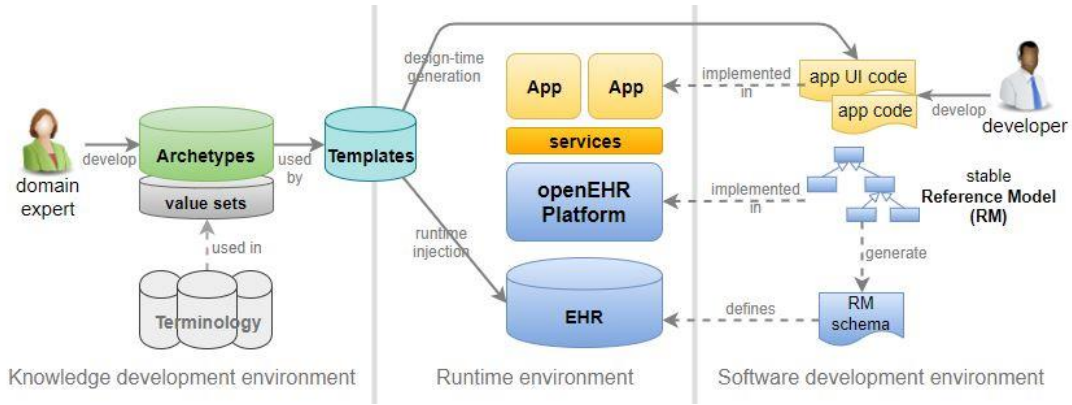


Figure 2.10 OpenEHR architecture

The Reference Model represents the semantics of storing and processing in the system and contains a set of generic data structures that are flexible enough to model most logical structures for knowledge representation. It provides identification, access to knowledge resources, data types and structures e.g. terminologies (Garde, Knaup, Hovenga, & Heard, 2007).

The models built by domain experts are in their own layer, known as archetypes that act as international standards for re-usable clinical content. An archetype is a computable specification of the data points and groups of a specific topic e.g., "Genetic result". These are defined as constraint structures based on the OpenEHR Reference model which ensures that the clinical meta-data (i.e. who, when or where) does not need to be defined in each archetype (Garde et al., 2007). An archetype is similar to a LEGO® instruction sheet (e.g., for a plane) that defines the configuration of LEGO® bricks making up a plane. Archetypes are flexible; one archetype includes many variations, in the same way that a LEGO® instruction might include several options for the same basic object (OpenEHR, 2022). Archetypes can also be specialised and there are several generic structures available which include e.g., Observation, Evaluation, Instructions or Action archetypes (Sachdeva & Bhalla, 2012).

Templates are the means of building clinical data sets specific to a use-cases and are composed of one or more archetypes. Every term in an archetype or template can be linked to a terminology e.g., SNOMED CT, and they are independent of specific messaging or document standards which ensures data interoperability(OpenEHR, 2022).

Garde et al have studied the impact of OpenEHR and archetypes have on health professionals and on semantic interoperability (Garde et al., 2007). They find that archetypes enabled the formal definition of clinical content by domain experts, but without the need for technical understanding. Archetypes enabled information and knowledge to be specified in a precise manner and to the level of granularity required. The two-level approach of separating information structures and clinical knowledge ensured that the archetypes could evolve and deal with changing health knowledge.

To avoid overlap and ensure evidence-based and relevant archetypes are easily accessible, development and maintenance must be systematically organised. Clinical Knowledge Manager[10] is a web-based tool that was developed specifically to support the different processes associated with Open EHR archetype development. It consists of a library over all existing archetypes and templates and enables collaborative development, management, and publishing.

In Norway, a national initiative is currently aiming at standardising the EHR system, DIPS, based on the OpenEHR framework (Ulriksen et al., 2017). Although there are some archetypes that are defined internationally, most will need adaption to fit a national context, and it will also be necessary to develop archetypes from scratch. This entails a large-scale infrastructure that might overlook complex socio-technical issues that although maybe irrelevant in small-scale pilot studies become critical in full-scale implementations. In their study of the challenges involved in developing archetypes in Norway, Ulriksen (Ulriksen et al., 2017) found that reaching a consensus was more complex and time-consuming than expected and that in addition to clinical content,

[10] https://ckm.openehr.org/ckm/

terminologies, semantics and technological issues were also discussed. This implicates a need for a close relationship between clinical and technical stakeholders.

The OpenEHR Genomics project has recently published several archetypes to represent genetic test results, developed based on a Variant calling Format (VCF) file specification (Mascia et al., 2020). DNA variants are described using the nomenclature from the Human Genome Variation Society (HGVS). The *Genomic variant result* is the core archetype designed to be nested within the *Laboratory test result* to report observations and annotations related to DNA variations found as the result of a genetic test, illustrated in figure 2.11. Several archetypes for specific types of DNA sequence changes e.g., "*Genomic substitution variant*", have also been developed, and these are designed to be used in the 'Structured variant' SLOT within the CLUSTER.genomic_variant_result archetype. These archetypes have recently been translated by a multi-disciplinary team into Norwegian and are published in the Norwegian Clinical Knowledge Manager repository[11]. The directorate for e-health recommends OpenEHR as a representation of information models intended for data storage in clinical systems (Directorate for e-health, 2021).

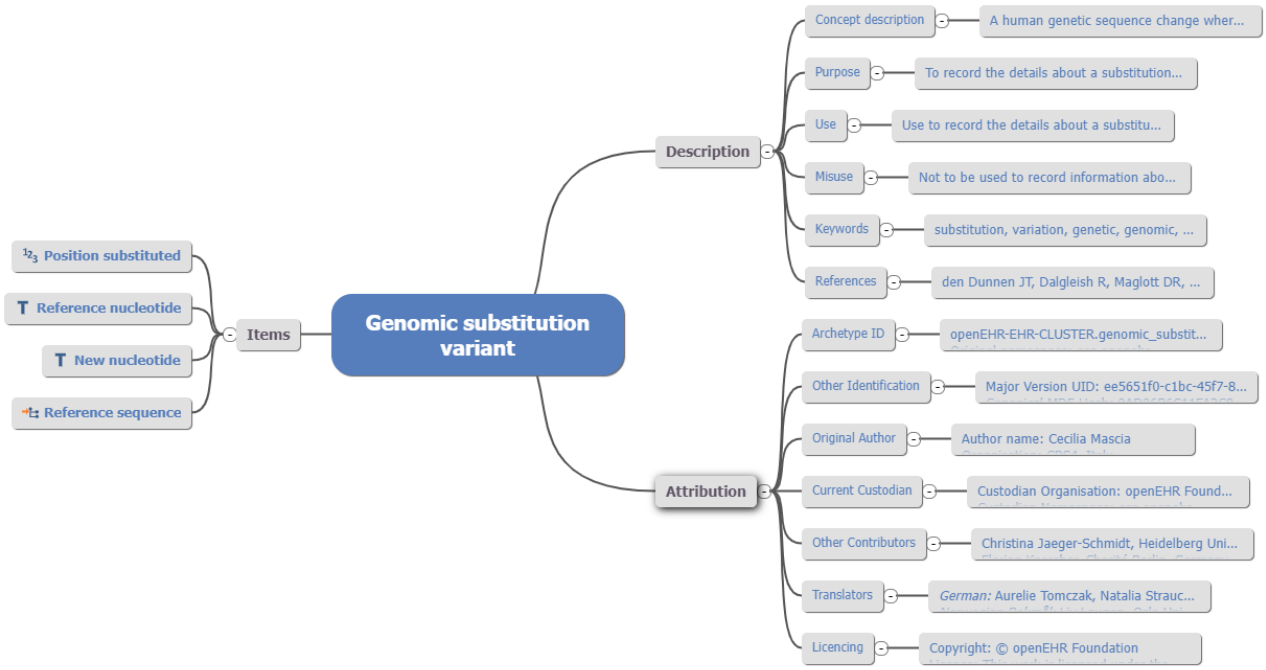


Figure 2.11 Mind map of the archetype Genetic Variant

2.6.2 HL7

Health Level Seven International (HL7) is a non-profit organisation that develops international standards for the transfer of clinical and administrative data. HL7 standards define how information is packaged and communicated from one party to another, setting the language, structure, and data types required for seamless integration between systems.

[11] https://arketyper.no/ckm/archetypes/1078.36.2208

HL7v2, originally created in 1989, is currently the most implemented messaging standard between LIMS and EHR in USA (hl7.org/standards) using a non-XML syntax. In USA EHRs are mandated by law to use HL7 version 2.5.1 for most operations under Promoting Interoperability (The Office of the National Coordinator for Health Information Technology)[12]

The HL7 v3, from 2005, represents a new approach to clinical information exchange based on a model driven methodology that produces messages and electronic documents expressed in XML syntax. It is based on a standard Reference Information Model representing the main business logic of any healthcare environment, from which specific messages and documents can be defined. HL7v3 messages are primarily used in conjunction with HL7 Clinical Document Architecture.

HL7 Fast Health Interoperability Resources (HL7 FHIR) is the latest standard to be developed by HL7, developed to meet the demands for more efficient and flexible development of standard-based integrations and better support for integration with modern technology such as mobile and cloud services. It is a new generation specification that uses modular components called "resources." The resources (definitions of common reusable patterns of clinical information) cover different healthcare scenarios and can be combined or extended to provide solutions for health information systems. FHIR includes resources as well as application programming interfaces (APIs) that exchange these resources. It was developed to become a standard that allows external parties to access information from EHRs using applications and allows third parties to create their own applications that can access these servers. FHIR standards utilise SMART FHIR(Substitutable Medical Applications, Reusable Technologies), an authentication framework for the connection of 3rd party applications. The directorate for e-health recommends FHIR as a representation of information models intended for data exchange (Directorate for e-health, 2021).

HL7 FHIR Genomics uses resource extensions to handle multiple types of genomic data, including genomic data files (e.g., VCF) and messaging services for interpretive genomic reports (Alterovitz et al., 2020). It introduces a new resource, MolecularSequence, designed to hold next generation genetic data in blocks relevant to actionable clinical decision-making and can associate it with repositories for retrieving a patient's full sequence data, such as those defined by GA4GH[13].

FHIR Genomics also presents new profiles in the existing resources Observation and DiagnosticReport to facilitate sharing of genetic test results including the clinical implications and interpretations, as well as representations of VCF. See figure 2.12 and 2.13.

[12] https://www.healthit.gov/topic/meaningful-use-and-macra/promoting-interoperability
[13] https://www.hl7.org/fhir/genomics.html

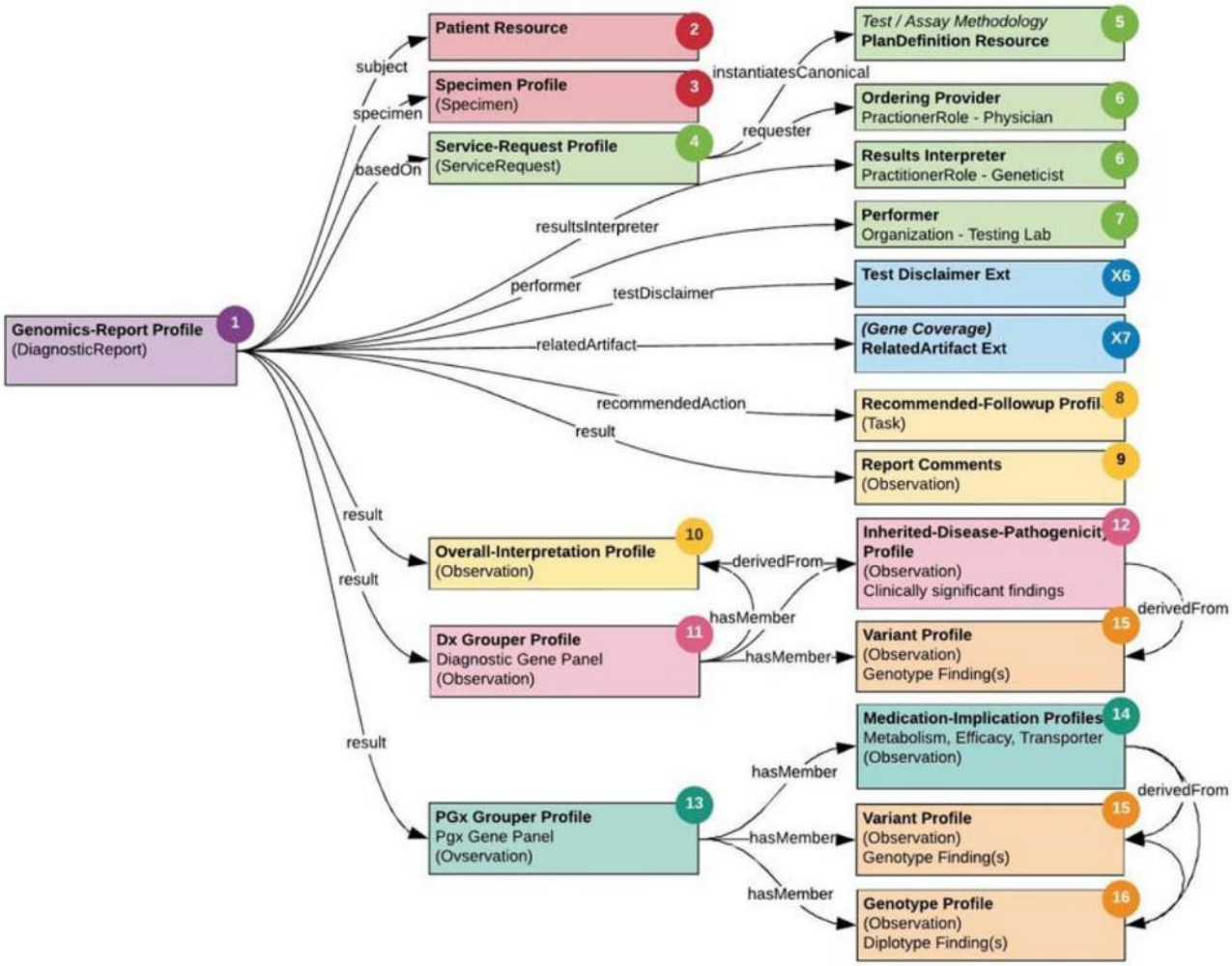


Figure 2.12 Structural design for FHIR Specifications for DiagnosticReport (Murugan et al., 2021)

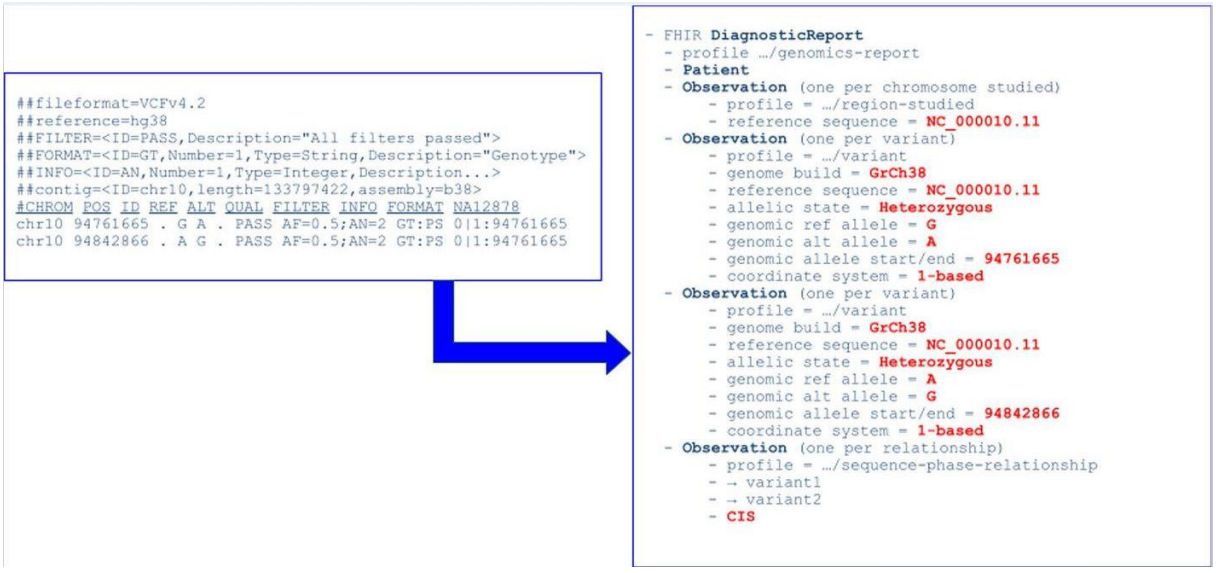


Figure 2.13 From VCF to FHIR Observation profile (Dolin et al., 2021)

2.6.3 GA4GH

The Global Alliance for Genomics and Health (GA4GH), founded in 2013, aims to accelerate biomedical advances by enabling the responsible sharing of clinical and genomic data through harmonised data aggregation (Rehm et al., 2021). They develop standards to tackle challenges in eight distinct areas, including large-scale genomics, genomic knowledge standards and clinical & phenotype data capture. Each GA4GH standard, with more than 30 approved, can be implemented on its own, but when implemented together they support a broader range of clinical activities and enable genetic data sharing (Rehm et al., 2021). GA4GH collaborate with other standard developers and have submitted their standards to ISO's technical committee for approval, as well as translating the GA4GH standards into HL7 FHIR Implementation guides. The standards also interoperate with existing translational models, ontologies and terminologies such as HGVS, LOINC, HPO or SNOMED CT. GA4GH focuses on maintaining and evolving standardised file formats for raw and annotated genetic data, individual variant representation and interpretation and sharing of individual phenotype data and interpreted results, all of which are critical for clinical laboratory workflows to share genomic data and genetic testing results (Rehm et al., 2021).

There are many GA4GH standards that are relevant for genetics but the GA4GH Genomic Knowledge Standards, Variation Representation Specification (VRS) and Variant Annotation (VA) are particularly useful with regards to structured genetic test reports. VRS and VA were developed to address the diverse methods used to access reference sequence, genetic annotation (e.g., genes, variation, regulatory regions, expression) and the capture of associated unstructured metadata. They are based on an information model that links assertions to the evidence supporting it and describes how the evidence was interpreted and applied. The complexity of a given profile can be defined to match the complexity of its source data and use-case.

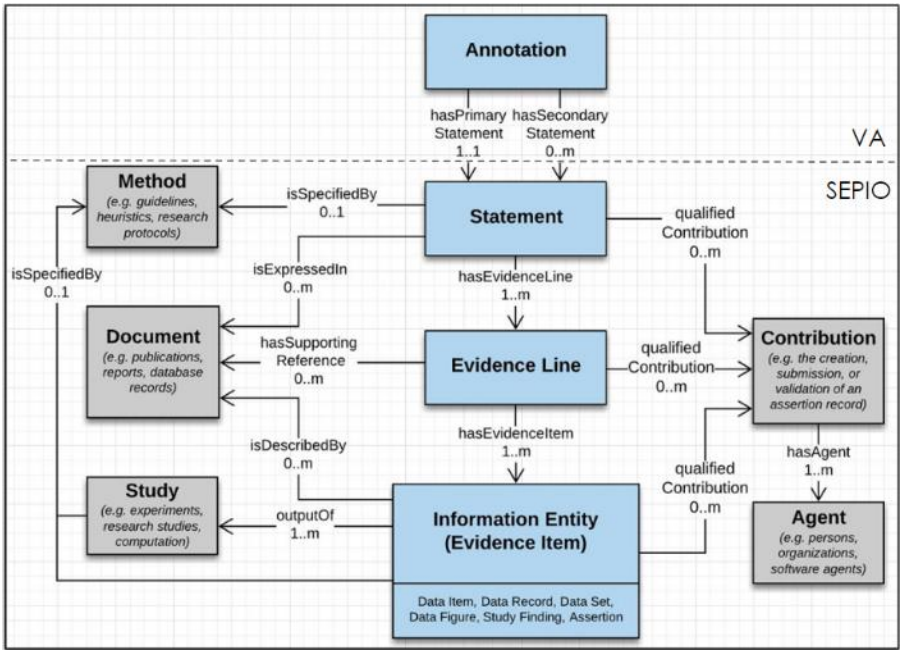


Figure 2.14 Variant Annotation Information Model

GA4GH define a variant annotation as "structured data object that holds a central piece of knowledge (a 'Statement') about a genetic variation, along with metadata supporting its interpretation and use" (GA4GH, 2022).

See figure 2.15 for an example of a variant pathogenicity annotation of "The germline variant 19:45411941:T:C is pathogenic for Breast Cancer". With this structure, it is clear what is primary knowledge, what knowledge represents evidence supporting it, and what is additional metadata.

```
variantID: 17694 (19:45411941:T:C
significance: pathogenic
condition: Breast Cancer
origin: germline variant
submitter: Genologica Medica
date: jun 19, 2022
assertionMethod: ACMG Guidelines 2015
collectionMethod: Clinical study
bibliography: PMID:1730728
functionalImpact: Loss of function
genomicFeatures: BRCA1
equivalentVariants:[NM_007298.3:c.2012T>G,NC
_000017.11:g.43051071A>C,...]
```

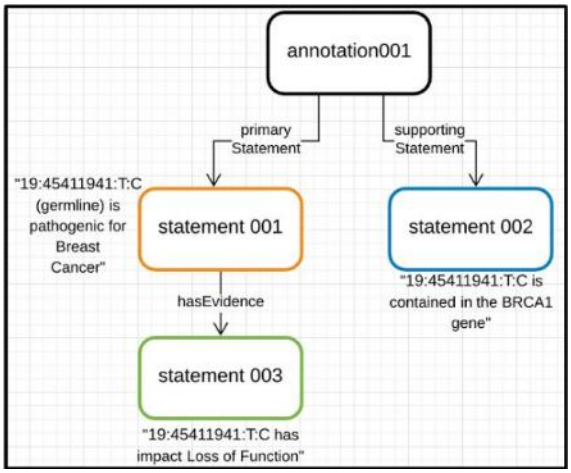


Figure 2.15 Source data for a pathogenicity variant and VA model

The VA standard will provide extensible data models to represent statements made about genetic variations and the evidence supporting them and will also serve as an exchange standard to facilitate sharing and integration of variant annotation data. It defines a set of distinct schema to represent different kinds of statements about genetic variations, each built on a common core model and uses a statement-centric approach, where each assertion of a fact about a variant is identified and structured as a discrete "Statement object". The generic statement object model can be specialised to provide statement classes that can be defined for specific types of annotations and the knowledge is represented by subject, predicate and object. See figure 2.16 for an example of a Pathogenicity Classification Statement.

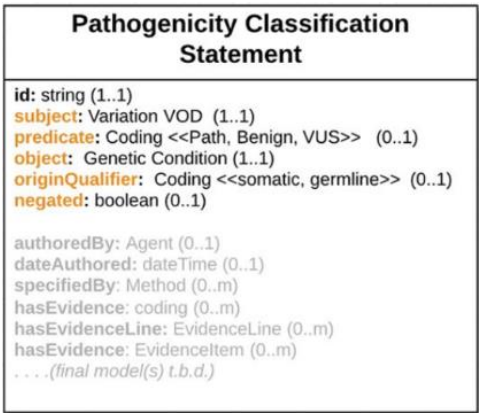


Figure 2.16 Pathogenicity Classification Statement

VRS complements existing variation standard representations e.g., HGVS which is designed for human interpretation, and provide a mechanism for computational representation. It also comprises of several interdependent components, including a terminology and information model, machine-readable schema, sharing conventions and globally unique identifiers. Through precise technical definitions of variation, the standard allows for the semantically accurate representation of many forms of genetic variants.

3 Literature Review

A literature review was performed to investigate the current state of how genetic information is displayed and integrated with other patient information in the EHR, identify research gaps, and provide a rationale for further research and development of the research aims. Literature reviews are an essential tool in the research process as they provide a framework around which the importance of the research can be established and a benchmark by which new findings can be compared (Creswell, 2014). The methodology used to search the literature is described before the results of the review are presented in this chapter, as the results from the literature review determine further method choices.

3.1 Literature Review - Method

A literature review is a search and evaluation of a selection of available literature within a certain topic and aims to provide a summary of relevant findings within the specific domain. The literature also provides solutions and insights. The Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) statement, published in 2020, described the necessary recommendations for a transparent and accurate literature review (Page et al., 2021). A literature review involves three major stages as shown in figure 3.1.

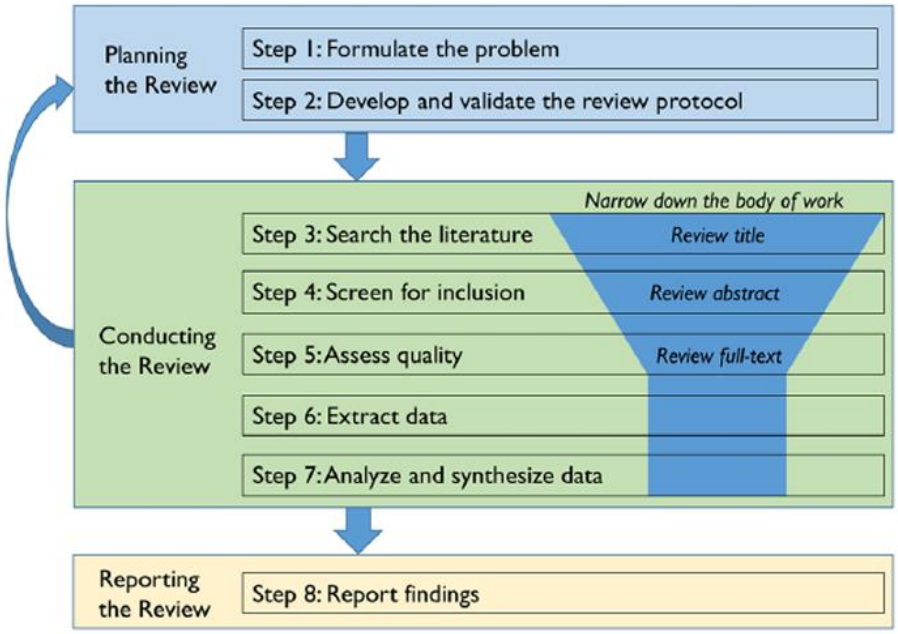


Figure 3.1 Process of a literature review (Xiao & Watson, 2017)

Structured genetic information integrated into the EHR is only a relatively new opportunity, as both the genetic information and the technology to enable integration across HISs has become recently available. Therefore, a structured literature review format was chosen, more specifically a State-of-the-art review. This type of review is

suitable for my purpose as it tends to address more current matters in contrast to other types of reviews. It aims for a comprehensive searching of current literature and provides the current state of knowledge and priorities for future research (Grant & Booth, 2009). However, the time-bound limitations may distort the overall picture of development within a domain. The review has been limited to case reports as these are a useful tool to describe the range of capabilities that an IT infrastructure or a particular technology must support (Taylor et al., 2021).

3.1.1 Planning the Review

As shown in figure 3.1, the first stage of a literature review is planning, which includes determining the review questions.

Review questions: For this literature study, the objectives are to study the current state of integration of genetic test results in the EHR with the following questions:

1. What are the genetic standards adopted in the studies?
2. Where is the genetic data displayed in the EHR?
3. Was genetic data integrated into a CDS?
4. What are the challenges experienced with integration of genetic data into the EHR?
5. Are user involvement and functional requirements described?

3.1.2 Conducting the Review

The next stage is to conduct a search strategy that will narrow down the body of work. A period from 2013 to present was chosen, as this represents the beginning of the development of genetic interoperability standards.

Search strategy: A literature search was performed using Pubmed and Google Scholar. Manual search of references was also conducted.

Inclusion criteria: the following criteria for inclusion were used to select articles to be included in the review.

1. Field of interest- search restricted to human genetics
2. Measure of interest: Case studies that have successfully integrated genetic results into an EHR
3. Setting. All- government, research, private institutions
4. Timeframe: from 2013 to May 2022.
5. Languages: English

Exclusion criteria: review articles, brief reports, rapid communications, posters or abstracts

Search terms and keywords:

1. Genetics OR genomics
2. Interoperability OR integration
3. ("electronic health record*" OR "electronic medical record*" OR "electronic personal health record*" OR "personal health record*" OR "medical record* system*")

The search strategy involved screening titles and abstracts for articles that fulfilled the inclusion criteria. To identify additional studies, a manual search of the reference list of relevant studies retrieved from the electronic database and citing articles was performed.

Quality assessment and data extraction: The review was restricted to peer-review articles which were uploaded to Endnote. The following data was extracted in excel spread sheet: publication year, country of origin, sector, objective, study design, standards used, description of data in the EHR, interoperability summary, challenges, opportunities and future solutions.

3.1.3 State-of-the-Art : Literature Review Findings

The final stage of a literature review is reporting of the findings. 58 articles describing integration of genetic data into the HISs were identified from the databases. Of these, 37 were excluded by abstract and article screening for failing to describe specific case studies relating to genetic data and integration into the EHR. Most of the excluded articles described mapping of genetic elements and methods of obtaining structured genetic test reports but failed to mention use of the structured data within an EHR. The final 21 case studies were predominately studies from USA, the major contributor being eMERGE publications. eMERGE is a national USA network consisting of 2 central sequencing sites and 11 clinical sites and has a goal it to integrate genetic test results to the EHR for clinical care. Table 3.1 shows a summary the 21 case studies.

Table 3.1 Results of 21 articles describing integration of genetic data in the EHR

Year	Country	Name of study	Standards described	Type of genetic data	Results in EHR	CDS available	Ref
2017	USA	PGTIC	HL7 Arden syntax	PGx	PDF report Structured data in external repository	CDS pre-test	(Sissung et al., 2017)
2018	USA	eMERGE	HL7 XML LOINC SNOMED	PGx Results from genetic sequencing	PDF report Structured data in external repository	Not described	(S. Aronson et al., 2018)
2018	Europe	U-PGX	HL7 HL7 FHIR HTTPS	PGx	PDF report Structured data in external repository	CDS pre and post-test	(Blagec et al., 2018)
2018	USA	Prototype EHR	HL7 FHIR Infobutton	Results from genetic sequencing	Structured data in a prototype EHR	Not described	(Crump, Del Fiol, Williams, & Freimuth, 2018)
2018	USA	GACS	HL7 HL7 FHIR CDS Hooks LOINC	PGx	Structured data in external repository, CDS recommendations in EHR	CDS pre-test	(Dolin, Boxwala, & Shalaby, 2018)
2019	USA	Information Management plan	SNOMED CT LOINC HGVS HL7	Results from genetic sequencing	Structured data in EHR	Not described	(Campbell et al., 2019)
2019	USA	CancerLinQ	HL7 FHIR mCODE	Results from genetic sequencing	Structured data in external repository	CDS pre and external post-test	(Conway, Warner, Rubinstein, & Miller, 2019)
2019	Italy	FARMAPRICE	Hl7 IHE DICOM XDS	PGx	Alerts in EHR	External service	(Roncato et al., 2019)
2020	USA	VICC	HL7 FHIR genomics HGVS LOINC SNOMED	Results from genetic sequencing	Results in EHR not described Structured data in external repository,	SMART on FHIR application	(Alterovitz et al., 2020)
2020	USA	Sync for Genes	HL7 FHIR	Results from genetic sequencing	Not described	PGx CDS pre-test	(Garcia, Zayas-Cabán, & Freimuth, 2020)
2020	Italy	OpenEHR Genomics	OpenEHR	Results from genetic sequencing	Not described	Not described	(Mascia et al., 2020)
2020	Korea	Clinical Genomic Sequencing Reports in EHR	HL7 FHIR ISO/TS 20428	Results from genetic sequencing	Structured report in EHR	Not described	(Ryu et al., 2020)
2020	USA	eMERGE	HL7	PGx Results from genetic sequencing	PDF report Structured data in EHR + external repository	PGx CDS pre-test	(Walton, Johnson, Person, Reynolds, & Williams, 2020)

2020	USA	FHIR Lab Reports	SMART on FHIR CDS Hooks HL7 FHIR Genomics	PGx	PDF report Structured report in EHR	CDS pre and external post-test service	(Watkins & Eilbeck, 2020)
2020	USA	eMERGE	XML HL7	PGx Results from genetic sequencing	PDF report Structured data in external repository	CDS pre and external post-test service	(Wiesner et al., 2020)
2021	USA	SMART Cancer Navigator	HL7 FHIR LOINC CDS Hooks	Results from genetic sequencing	Structured report in EHR Structured data in external repository	External service	(Dolin et al., 2021)
2021	USA	eMERGE Neptune	XML HL7, HL7 FHIR	PGx Results from genetic sequencing	PDF report Structured data in external repository	External service	(Eric et al., 2021)
2021	USA	PennChart Genomics Initiative	HL7	PGx Results from genetic sequencing	Structured data in Precison Medicine Tab in EHR	CDS within EHR	(Lau-Min et al., 2021)
2021	USA	eMERGE	HL7 FHIR	PGx Results from genetic sequencing	PDF report Structured data in external repository	External service	(Murugan et al., 2021)
2022	USA	TriNetX	LOINC ICSN HGVS HL7 FHIR-variant profile	Results from genetic sequencing	Structured report in EHR, structured data in external repository	External service	(Hernandez et al., 2022)
2021	China	Pharmacogenomics Clinical Translation Platform	Not described	PGx	Structured report in EHR, structured data in external repository	External service	(Qin, Lu, Shu, Duan, & Li, 2022)

3.1.3.1 Interoperability and nomenclature standards

In the cases where terminologies and nomenclature standards were described, SNOMED CT, LOINC and HGVS were frequently used. Standard nomenclature for PGx variants was often described as problematic, and several studies ((Dolin et al., 2018; Qin et al., 2022) used external resources e.g. PharmCat[14] to extract genomic variants in VCF files and to standardise the process of assigning haplotypes and diplotypes from variants.

Many of the early studies mention the widely used HL7 V2 messaging standard, but the HL7 FHIR Clinical Genomics standards is dominating in the later publications. The few OpenEHR studies available were all based in Europe and focused on the use of OpenEHR for mapping of genetic test reports with little documentation of integration into either prototype or real-life EHR systems. The recently developed GA4GH Variation Representation Standard, although described as promising by Murugan (Murugan et al., 2021) was not used in any of the case studies.

PGx genetic results dominate the type of genetic data in the EHR but several studies are now focusing on clinical actionable genetic variants detected by DNA sequencing. There are predominately descriptions of the process of mapping of either VCF results or genetic

[14] https://pharmcat.org/

reports into computable terms and little focus of how the information is displayed in the EHR. When mentioned, the variant findings were placed in the "Results" section in the EHR or more commonly with PGx variants, the problem list or allergy section. Pennchart genomics (Lau-Min et al., 2021) describe a "Precision Medicine Tab" where all genetic findings can be displayed.

3.1.3.2 Integration of CDS

Several studies have successfully implemented genetic CDS logic. eMERGE have developed a CDS Knowledge Base[15] but due to the complexities of genetic CDS logic, most of the studies have relied on external web services to provide CDS. Many of the cases have designed a system architecture were raw genetic data remains in the LIMS or is stored as structured data in an external genetic data storage repository. Blagec (Blagec et al., 2018) describes an approach using an external genetic system, Genetic Information Management suite as a source for CDS. This translates genetic data into clinically actionable recommendations and creates reports compatible with different reporting standards ie HL7, FHIR and HTTPS, and sends results to the EHR. Dolin (Dolin et al., 2018) has a similar approach, where Genomic archiving and Communication system stores DNA sequencing data and can retrieve information and feed into the EHR through CDS Hooks. Both Dolin (Dolin et al., 2018) and Eric (Eric et al., 2021) have adopted the emerging standard, CDS Hooks, which takes advantage of Application Programming Interfaces (APIs) within the EHR and cloud-based clinical genetic storage. External genetic data storage systems allow for SMART on FHIR solutions which enable applications to be integrated for interoperability across different EHR vendors, which have been demonstrated by Alterovitz (Alterovitz et al., 2015) and Warner (Warner et al., 2016). However, CDS has had limited utility because executing CDS has required manual entry of genetic conditions into the problem list for decision support.

Conway (Conway et al., 2019) states that EHRs seem to be ill suited for managing genetic data and other external solutions are required, including applications that use APIs. Infobutton compliant genetic resources have also been implemented, Crump (Crump et al., 2018) using HL7 Infobutton standards succeeded in configuring eight relevant genetic knowledge resources and integrating into them into EHR next to conditions, medications and genes/variants but found many additional resources that were not infobutton compliant.

3.1.3.3 Challenges

One of the most recurring challenges, described in several studies, is the limited ability to transmit gene and variant information as standards-compliant structured data. This is due to several limitations including lack of adequate standards for representing and communicating complex genetic knowledge. Both Walton (Walton et al., 2020) and Garcia (Garcia et al., 2020) state that inadequate and discordant standards in the HL7 genetic report format and FHIR molecular sequence resource were a major challenge and Qin (Qin et al., 2022) report challenges regarding the computable representation of PGx variants. Murugan noted 21 issues regarding mapping of genetic test reports using FHIR resources and created FHIR extensions as a workaround. The study concludes that the complexity of FHIR® Genomics Reporting Implementation Guide could be a barrier to adoption, as it requires users to have an in-depth understanding of the domain, in addition to resources for implementation. Blagec (Blagec et al., 2018) also experienced

[15] https://cdskb.org/

technical challenges with the implementation demands and resources required to implement even simple CDS systems across different health systems.

There are also inadequate standards for the naming of genetic diseases and phenotypes, resulting in challenges linking genetic information to disease-specific knowledge bases.

Heterogeneity and the differences in the level of detail in reporting genetic results lead to organisational specific solutions that make for non-scalable solutions. There is a lack of a standardised genetic result file format which includes all critical and necessary data elements. Due to the inability of the LIMS to store or transmit structured genetic data, Walton et al (Walton et al., 2020) solved the challenge by writing a custom Python script to convert structured laboratory data into HL7 segments for importing into the EHR but despite intensive coordination between informatics and geneticists, experienced additional difficulties due to lack of standardisation of genetic phenotypes. The solutions demand a team of technically competent IT resources and domain experts, with several groups reporting delays and technical challenges.

Another recurring challenge was the inability of most EHRs to store individual genetic variants in a scalable and standardised manner, resulting in genetic results delivered in PDF and an external storage of structured genetic variants. Several studies (Qin, Conway, Murugan) conclude that current EHRs lag behind the advancement of precision medicine, resulting in workarounds, genetic information in multiple sites in the EHR and custom CDS implementations.

Reported challenges from case-studies:

- Inability of EHR to appropriately receive, store, manage and display discrete genetic information
- EHR not ready to send accurate coded clinical and family history data to LIMS
- Poorly defined role of LIMS in delivering discrete genetic data
- Genetic phenotypes are poorly defined and need standard definition to facilitate delivery of phenotypic information from the laboratory to the EHR
- EHR lacks sufficient information to allow for test ordering for specific genetic variants
- Genetic reports vary in structure and content between laboratories
- Rapid evolution of data types and use cases related to clinical genetics
- No standardised genetic variant data structures between LIMS + EHR
- Interoperability standards use syntax with limited hierarchy and inadequate coding systems
- Mapping between SNOMED CT and Online Mendelian Inheritance in Man (OMIM) terminologies remains as a semantic gap
- No standards for displaying variant data over time
- No consensus on how to request and reclassify variants
- No international standards for CDS rules
- Lack of resources for continual maintenance of evolving complex systems
- Genetic knowledge resources not compliant with HL7 Infobutton standard
- Genetic test reports that are difficult for clinicians to understand
- Engaging patients and clinicians in genetic testing
- Healthcare providers are not investing in HISs necessary to deploy precision medicine

3.1.3.4 Solutions

All the identified challenges and barriers impact the ability to fully use genetic information as a part of healthcare and there are many dependencies between the identified challenges. Some challenges can be overcome using existing platforms and resources while others require changes to EHRs and international standards (Williams et al., 2019).

There are multiple long-term solutions described in the literature. Several entail engaging both LIMS and EHR vendors to discuss future plans for genetic data, including adoption of Infobuttons, CDS Hooks and open API solutions. More specific genetic workflows should be integrated into the EHR e.g., automated workflows for reclassification or reanalysis of existing genetic information. In addition, the genetic community must define and develop the necessary standards required to represent structured genetic information.

Several different short-term methods for integrating data into the EHR are described in the case studies, and can be grouped into three main solutions, as described by Warner (Warner et al., 2016):

- Non-standardised integrations: A custom interface between a genetic laboratory and an EHR. This allows for integration of genetic information but with no interoperability that can be utilised by other institutions
- Middleware: A platform that is not fully integrated with an EHR, often a stand-alone web portal, but shown to be a useful module for conveying genetic information to clinicians. They often require an additional login and have a limited ability to merge clinical data with the genetic information.
- API: Have the potential to be fully integrated in a clinician's workflow and can be launched from within an EHR, due to SMART-FHIR technology. The combination of standardised representation of genetic information and external knowledge base systems results in a powerful tool.

3.1.3.5 User involvement

Only a few of the selected case studies described user involvement in the development of solutions for integrating genetic information in the EHR. Though this does not necessarily mean the others had no contact with end-users or other stakeholders, but it was not described in the case-study. Walton (Walton et al., 2020) included weekly meetings with relevant stakeholders to address challenges and barriers to implementation as they arose in a functional prototype before up-scaling to a fully functional solution and that early in the process, a clinical genetics team were involved to review workflows. The authors also acknowledged the importance of continuous engagement of the clinical genetics team. The PennChart Genomics Initiative (Lau-Min et al., 2021) is a multi-disciplinary effort including clinicians, researchers and software engineers and attribute their success to large-scale, collaborated and coordinated efforts of the whole-team. Others had involved clinicians in parts of the development process, e.g., the choice and prioritising of clinical and genetic terms that are required in a genetic test report. Wiesner (Wiesner et al., 2020) does discuss the need for the involvement of medical geneticists when disclosing results to the patients but expresses concerns about the limited number of geneticists and the increasing number of patients requiring genetic services.

3.1.3.6 Summary of literature review

The literature review identified a small number of early adopters of genetic data integration in the EHR, often small-scale solutions from large academic institutions in the

USA who are involved with a parallel development of standards and pilot implementation studies. Heterogeneity in LIMS and sequencing file formats, variation, and format of genetic reports, and different EHRs result in workarounds, individual protocols, and organisational specific solutions that are difficult to implement and are non-scalable. Several studies describe potential data models but with no real data, and very few mention solutions for the dynamic nature of genetic data e.g., genetic variant reclassification. Solutions for the standardisation of interpretation summary texts, gene/region coverage or aggregated results across different genetic tests was also seldom mentioned.

The integration of genetic data in EHR appears as premature, where early adoption is critical for maturation of standards but where the cost and risks of doing so are high. There are no best practice guidelines, and the current standards are not adequate for genetic data integration leading to process variation across the studies. There is a universal recommendation that genetic data should be displayed in an organised manner, ideally linked to an appropriate CDS.

3.1.3.7 Further choice of method

The selected case studies predominantly describe technical solutions, focusing on system architectures and interoperability standards, without describing clinical involvement from the users and user-interface interactions. In reviewing the literature on integration of genetic data in the EHR, it became apparent that there are several themes which remain unexplored. User involvement, user requirements and how genetic information is best displayed in the EHR are topics that remain unresolved. The results from the literature review made it apparent that qualitative methods with a focus on the needs of the users and their current workflows could fill a gap that was observed in the literature.

4 Method

This chapter looks at different research methods and discusses their strengths and weaknesses before outlining the methodological approach used for this study. The workflow and justification for the choice of methods used to achieve the objectives of the research are described and evaluated.

4.1 Research Setting and Scope

Oslo University Hospital is a highly specialised hospital responsible for both regional and local hospital assignments. The Department of Medical Genetics (DMG) at Oslo University Hospital (OUH) is Norway's largest medical genetics department and provides diagnostics and research within the field of hereditary diseases and focus on genetic diagnostics, genetic counselling, and research. The 7 diagnostic laboratory units, located at Ullevål hospital, offer more than 200 different tests, with a large amount of the samples analysed by sequencing. There are 3 clinical sections located close to Rikshospital, who have direct contact with patients, and perform genetic counselling and diagnostics with a wide variety of indications.

The goal of this study is to investigate how genetic information can be displayed and integrated with other patient information in the EHR and to elucidate the user's main functionality requirements that would support optimised integration and use of genetic information for patientcare. Currently there is no structured genetic information in the EHR that is assumed to complicate the clinician's workflow, making information retrieval time-consuming and clinical decision support near impossible to implement.

The scope of the study is restricted to the EHR currently in use at DMG, OUH and limited to a selection of users of this system at DMG, as well as users of DMGs services within OUH. I have not addressed organisational questions such as user's professional competence and responsibilities. The study was performed within the expected time allowance for a master thesis, together with the resources available to me as a single researcher and the investigations were therefore limited to one genetic department in Norway. Ideally, a feasibility study should have been performed. This would identify if future development of the EHR to facilitate the display of genetic information is justified and is technically possible, but the scope of this study is limited to requirement engineering. However, precise and descriptive user requirements provide the foundation for future software development and are critical for the success of any health information system.

4.2 Research Theory

According to Oates (Oates, 2005), research can be defined as " the creation of new knowledge, using an appropriate process, to the satisfaction of the users of the research". He continues to describe good research is when relevant data resources and sufficient background information are collected, and analysed in a thorough and correct manner, before reaching a conclusion that is presented in a research article or relevant conference. It is important to look at previous research with a critical eye and consider studies in relation to research methodology, validity, and relevance.

The following 6 Ps, purpose, products, process, participants, paradigm and presentation need to be considered for any research study (Oates, 2005).

- Purpose- the reason for doing the research
- Process-the sequence of activities undertaken
- Products – the outcomes of the research
- Participants – those involved in the research
- Paradigm – A model or shared way of thinking
- Presentation – the means of research dissemination

Figure 4.1 below illustrates a research approach, where Participants and Paradigm are present in all stages, while purpose, process, products and presentation are successively completed underway.

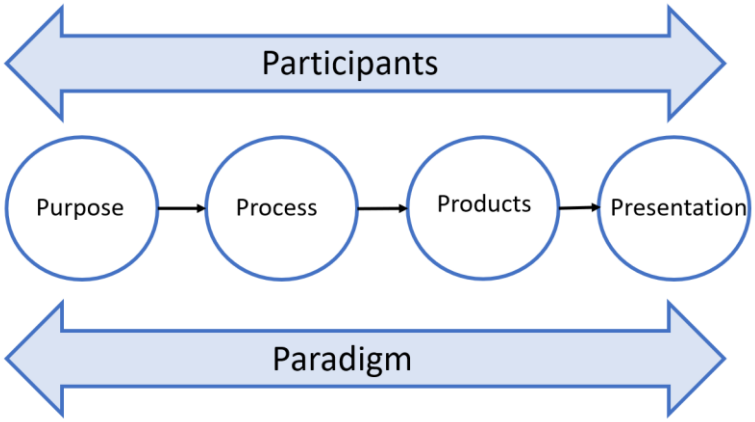


Figure 4.1 6P's of research (Oates, 2005)

4.2.1 Quantitative and Qualitative Methods

There are different research strategies or paradigms, and each have their strengths and weaknesses, what is suitable for one type of research may not be fitting for another.

Quantitative research is based on numerative data, accentuates testing theories and hypotheses, and involves analysing collected data by statistical methods. There are several quantitative methods that could be relevant for this study but perhaps most relevant is a questionnaire. A questionnaire allows for the possibilities of easy distribution to many respondents but there is often no opportunity for reflection or detailed responses. In addition, the results are highly dependent on the questions. It requires excellent knowledge to formulate questions and demands a similar knowledge of those who are going to respond, to avoid the questions being interpreted differently between respondents.

Qualitative research, however, emphasises insight and understanding, is exploratory and can identify problems or uncover answers to questions that the researcher is unaware of. It provides in-depth data from a limited number of participants and helps discover how users think and experience (Tjora, 2017). Qualitative methods are often based on theories of interpretation and human experience and include interviews, observations, and focus groups, followed by a qualitative content analysis of the data e. g. summarising, categorising, and interpreting. Qualitative methods such as interviews and observations result in more reflected responses, that may however, be subjective to the few respondents included in the study while data analysis often incorporates interpretations that are researcher subjective.

4.3 Method Approach and Framework

It is important to choose a research method that is appropriate to answer the research objectives. The observations from the literature review were also essential in the planning and choice of the method approach.

The literature review in chapter 3 established that the integration of genetic information in the EHR is technically challenging with inadequate interoperability standards and lack of available guidelines. One noticeable gap in the reviewed literature is lack of involvement from clinical users. There are numerable descriptions of different technological solutions with the aim to make genetic information readily available in the EHR but little data on the resulting organisation and display of information in the EHR or if these solutions are in fact in agreement with the user's needs.

Healthcare systems require continuous innovation to meet the needs of the users, who are, however, often neglected when new system processes are designed, which results in products with reduced functionality (Altman, Huang, & Breland, 2018). To achieve my goal of understanding the main requirements of clinical users, it is necessary to focus on user involvement. Based on the research objectives for this study, where I want to investigate, understand, and explore the topic of genetic information in the EHR, qualitative research methods were a suitable choice. In addition to academic considerations, the choice of method is also determined by practical considerations. Medical genetics is a highly specialised domain, but from my background from the DMG, I have ample possibility to undertake observation studies and interviews and can recruit a small number of individual domain experts who have relevant experience and who are able to provide insight on how genetic information can be displayed and integrated in the EHR. In addition, the time constraints of the study and the busy time-schedule of the respondents also influenced the choice of method. The results of this research will be made public through the open-source platform of the University of Trondheim, Oria.

4.3.1 Study Framework

The model of the study process is based on a combination of personal experiences and motivation, literature review, objectives, conceptual framework, data generation methods and data analysis (Oates, 2005). Figure 4.2 illustrates the study approach chosen for this study.

My motivation and research objectives have been described in chapter 1, and the literature review findings in chapter 3. The conceptual framework comprises the combination of strategies and data generation methods together with the data analysis approach and is mainly derived from my experiences and the findings from the literature review.

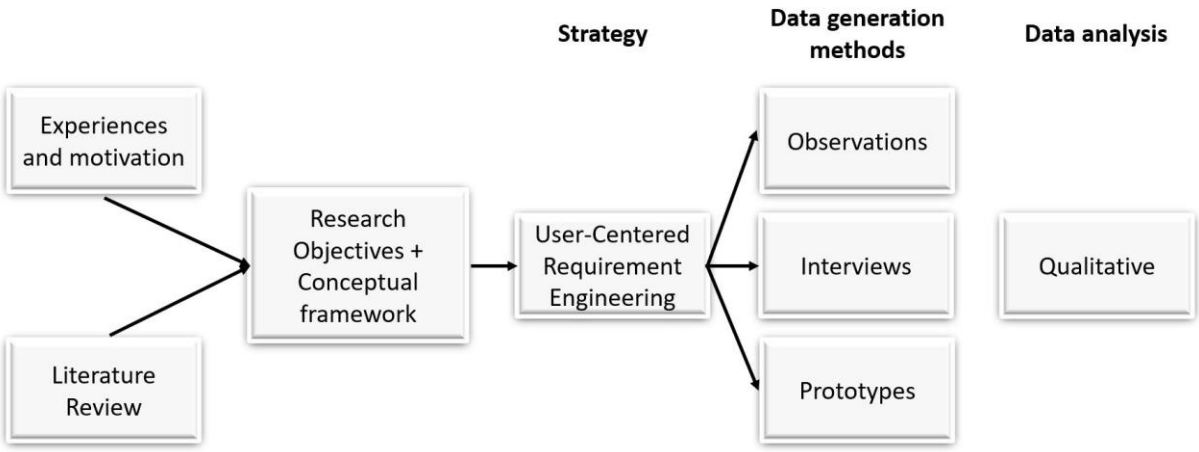


Figure 4.2 Model of the study approach, adapted from (Oates, 2005)

4.4 Study Strategy

A strategy is the overall approach to answer the research objective and examples of different strategies are e.g., survey, experiment, case study or design and creation. User-centred requirements engineering (UCRE) has been used as the strategy to answer the research objectives in this study and provides the basis for the data generation methods. The research objectives in this thesis are derived from user's interactions with a HIS and aims to elucidate user requirements for the interface between EHR and the clinicians utilising genetic information displayed and stored in the system. UCRE, where the users are involved using different techniques of requirement elicitation and validation, together with the context of the study and access to resources makes it a most suitable method for this research.

4.4.1 User-Centred Requirements Engineering

Preece et al (Preece, Rogers, & Sharp, 2019) define a requirement as "a statement about an intended product that specifies what it is expected to do or how it will perform". There are many kinds of requirements in software development but the two most often used are: functional requirements, which describe the constraints of the product and capture what a product will do and non-functional or system requirements which describe the technical specifications of the system being developed. The user requirements for a system can describe both functional and non-functional requirements.

Requirements engineering is a process of gathering and defining what services should be provided by a system. Successful requirement engineering involves understanding the needs of the user in addition to understanding the context in which the software will be used. Identification of the users who may be impacted by the system is critical and will ensure that the needs of all involved are considered. End-users are the driving force behind product development with the goal of a well-designed system that supports rather the constraints the user (Preece et al., 2019). Work in health care is especially dependent on advanced levels of knowledge which often evolves from discussing the issue with a co-worker rather than search for information other places so when designing health IT systems, it is important to give the end-users a prominent position in the requirement engineering process (Das & Svanæs, 2013). Lack of user input during design has shown to be the biggest contributing factor in the adoption failure of IT systems

whereas reported benefits of user involvement include improved system quality due to accurate user requirements, inclusion of needed functions and the avoidance of costly unwanted features, and reduced training needs (A. Kushniruk & Nøhr, 2016).

Focus on the user's knowledge and their workflow to ensure that the resulting product meets the user's real needs, is the central principle of user-centred requirement engineering (UCRE), is especially relevant in the complex domain of genetics. This is, however, frequently neglected in the genomic tool development process(García S et al., 2022), due to the complex data which is rapidly evolving in detail and volume.

4.4.1.1 Requirements elicitation

Requirements elicitation is an iterative process that can be represented with a circle of activities- discovery, classification and organisation, prioritisation and requirements specification, as illustrated in figure 4.3 (Sommerville, 2016).

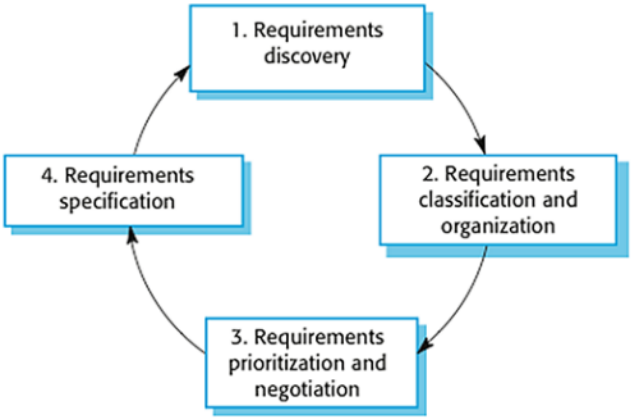


Figure 4.3 Process of requirements elicitation and analysis (Sommerville, 2016)

Each step in the requirement elicitation process shown in figure 4.3 and described below, was included in each iteration of data gathering in this study.

1. **Requirement discovery:** the process of interacting and gathering the requirements from users and stakeholders, using techniques such as observations and interviews. However, gathering and understanding requirements is difficult for several reasons - users often do not know what they want from a system and express requirements in their own terms which software developers struggle to understand, and different users have different requirements which they describe in different ways. One way to capture what a product is intended to do is by UML use-case diagrams. Use-case diagrams present all functional requirements of the system and capture interaction between the user and the product (Preece et al., 2019).
2. **Requirement Classification and organisation:** the process of grouping related requirements together and organising into related topics or themes.
3. **Requirements prioritisation and negotiation:** when multiple users are involved, requirements will conflict. This step is concerned with finding and resolving requirements conflicts through negotiation and compromises.
4. **Requirements specification:** the requirements are documented for further use in the software development process, so that they are understandable by users with no technical knowledge. Quality and accuracy are important, and unclear or confusing requirements can create conflicting interpretations. Well-written

functional requirements should be necessary, concise, feasible, consistent, precise and verifiable (ref: IEEE830-1993), User story format has a simple structure eg As a "role", I want "behaviour" so that "benefit", with a focus on outcomes and describe the interaction between a user and the system. User stories help to eliminate misunderstandings about the scope and functionality of the system as they are structured but easily understood by the users.

The final elucidated requirements must be validated to ensure they actually meet the user's needs. Product defects arise most frequently during the requirement phases in development, due to both misunderstandings or changing needs over time, and are costly to correct late in the development cycle (Kannan et al., 2019). Prototyping is one methodology that can be used to validate requirements.

4.5 Data Generation Methods

Gathering requirements is not always as straightforward as asking the users what they want the system to do, as they are often immersed in the current system and hard for users to detach from the way they currently do things and imagine possibilities that are significantly different from what they use now. No single requirement gathering technique will be able to produce a complete set of requirements, so it is recommended to use a variety of methods. Based on the research setting and objectives, the following data generation methods were chosen to answer the study objectives shown in Table 4.1.

Table 4.1 Methods for data generation

Objective	Method for data generation	Description
What requirements do clinicians have for optimal use of genetic information in the EHR?	Observation	To map current workflow.
	Interview	To identify requirements. Confirm workflow model.
	Prototyping	To validate previously identified and identify new requirements.
Which solutions are available to display and integrate genetic information with other patient information in the EHR?	Interview	To identify solutions from the end-users
	Prototyping	To validate solutions
	Literature (Chapter 3)	To identify solutions from the literature
What are the challenges of integrating genetic information with other patient information in the EHR?	Interview	To identify challenges at DMG
	Literature (Chapter 3)	To identify challenges from the literature

The methods for data generation were used in different iterations of requirement elicitation, as shown in figure 4.4, the results from one method were validated by the results of another. All the methods combined contributed to the requirements engineering process.

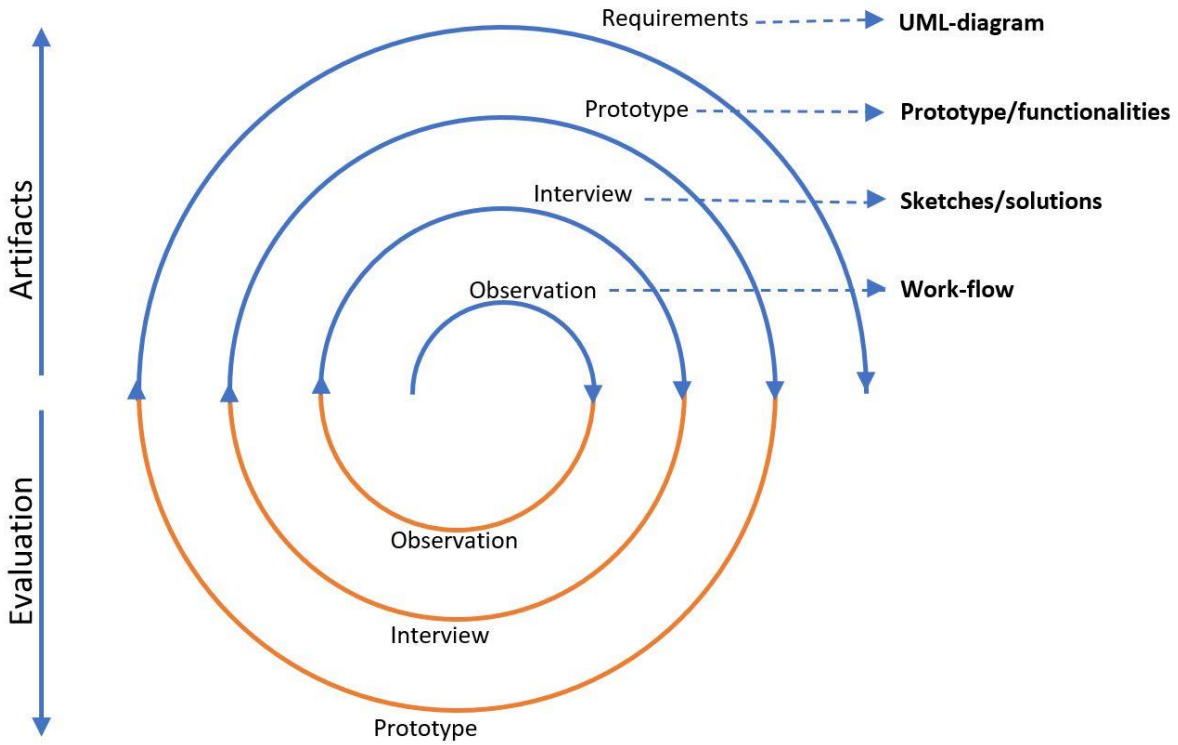


Figure 4.4 Spiral view of requirement engineering method used in this study, adapted from(Teixeira et al., 2012)

4.5.1 Identifying and Selecting Users

UCRE is well-suited for the development of new concepts that meet complex requirements involving different user groups (Ehn, Derneborg, Revenäs, & Cicchetti, 2021). However, successful results are dependent on the selection of respondents and the choice of data generation methods (Das & Svanæs, 2013).

A letter of participation was sent to a selection of users of genetic information in Oslo University Hospital - medical geneticists, genetic counsellors and referring clinicians, see Attachment A. There are a limited number of genetic specialists and users of clinical genetic information at OUH, therefore, the target sample size for the qualitative interviews and prototype testing was ~8–10 participants. Evidence from usability studies, indicates that up to 80–85% of usability issues can be identified by the first eight participants (Virzi, 1992).

The participants were chosen to represent a diversity of user groups, and users with difference levels of experience to ensure a representative selection. The invited users were primarily clinical or laboratory geneticists who are professionals interpreting the clinical implications of a patient's genetic data who may work within the laboratory setting or outside the laboratory. Referring or treating clinicians who make a diagnostic, treatment, or preventative decision or recommendation, based on genetic information, were also chosen to participate, as well as genetic counsellors.

However, user involvement can be a time-consuming process, involving multiple iterations and due to time limitations of this thesis, the number of iterations possible is restricted. By using multiple methods for requirement elicitation, I hope, nevertheless, to

uncover different kinds of relevant information and needs from the users in the available timescale. Based on the methods presented in Table 4.1, a preliminary process was formed, consisting of three different user groups. Some individuals were involved in all 3 phases, but each user group also included participants not previously involved. In this way, I could ensure a representative data collection and avoid the bias of individuals with strong opinions on one certain matter. Each user group could evaluate the identified artifacts from previous steps. A total of 12 participants were involved in the research study.

4.5.2 Observational Studies

User observation involves focusing on what people do, by watching a process carefully and paying attention to significant details with the goal to answering specific questions (Tjora, 2017). My particular focus was to gain insight into the tasks and workflows of the users.

The departments procedures on use of the EHR were first studied to gain a basic understanding of the software to find out what details should be given focus on during the observations. Four participants were visited at their workplace, each with different responsibilities, and an active observation was chosen, where the users have been asked to demonstrate a search for genetic information in DIPS. The participants explained what they were doing and were asked questions underway to ensure an understanding of the process and avoid relying on assumptions. This interaction can change how a user would perform a task, so the dialogue was limited to a minimum and further questions were noted that were then added to the interview guide. Notes were taken underway to document tasks involved, the functions used in the software, if other systems than the EHR were used or the problems the users encountered underway.

4.5.3 Interview

Having gained insights from the literature and observations, there was a need to further investigate topics within the integration of genetic data at the OUH. The most widespread method of generating data within qualitative research is interviews (Tjora, 2017)..

In general, qualitative interviewing emphasises the importance of investigating experiences and perspectives of the interviewees for developing a better understanding of a specific topic (Tjora, 2017). Interviews with users, provides an insight into the users' needs and an understanding of important aspects of the use situation that may affect the user experience. Interviews provide the opportunity for informants to highlight complexities and nuances that may be lost with less flexible collection methods. A lot of knowledge is situational, and interviews can be largely tailored to the informant and the data to be collected. However, the disadvantage of this flexibility is the challenge of managing unlimited discussions, note-taking, and the pursuit of following up on questions takes a substantial effort. Interviews can be described as a "conversation with a purpose" and there are four main types: open-ended/unstructured, structured, semi-structured and focus groups(Sharp, 2019).

Focus group sessions were considered to generate ideas and solutions to the challenges of integration of genetic information into the EHR. This method can generate numerous new ideas and resolve obstacles, is a cost-effective and time-saving method and can provide widespread involvement in a group. However, differences in opinion may emerge that confound the requirement gathering process and it demands good facilitating to avoid outgoing people dominating the session with their ideas. Due to the organisational

challenges of focus groups, and the potential for a single person to dominate the conversation, interviews were chosen as the preferred method.

For this study semi-structured interviews with an interview guide (Attachment 2) were considered the most appropriate as it provides structure which allows them to be replicable but allows the participants to express opinions. Semi-structured interviews enable a discussion focused on the objective and are appropriate for eliciting user requirements from participants who have limited time to available. The interview guide prevents the phrasing of suggestive questions which could lead to the introduction of bias (Sharp, 2019).

Having worked in the field of genetics for many years, I have solid background in the domain that has been studied. However, I am not a user of the EHR myself and both the observations and interviews were necessary to gain an understanding of the clinician's workflow and needs. Seven participants were selected who could speak out in a reflective manner on the topic in question, a so-called strategic choice, as the informants are not chosen randomly (Tjora, 2017). A pilot interview was conducted to test the suitability of the questions and to ensure the given-time frame of one hour was sufficient. The interview guide was adjusted, and several questions removed.

For the interviews, the objectives are as follows:

1. What type of genetic data is registered in DIPS and how is it organised?
2. Identify user processes and workflows
3. Identify challenges with current solution
4. Identify new functionalities with sketching and use examples

Notes were taken under the interview, sketches of simple functions were drawn and directly after each interview the most important findings, identified tasks and apparent user requirements were summarised. No sensitive data was involved, and the interviews were not recorded.

The participants evaluated the workflow identified earlier during the observations and sticky notes were used to describe were challenges occurred.

Sketches are useful in the exploratory stages of a design to propose and communicate ideas in an easy format and can be used to explore many different user requirements. Simple and quick sketches (10 seconds) on paper and sticky notes were made in the first round of interviews. They are very quick and easy to generate during an interview and were drawn both by myself and the participants.

4.5.4 Prototyping

Paper prototyping has been shown to be a highly suitable method for eliciting functional requirements and identifying requirements with high completeness and quality (Rueda, Panach, & Distante, 2020). A prototype can be useful when evaluating ideas and depending on the complexity of the prototype can e.g., test the technical feasibility of an idea or to clarify some uncertain requirements. Low fidelity prototypes, often simple paper prototypes, can be sufficient to show intended functions and the positioning of buttons. They do not look like the final product or contain the same functionalities, but are simple, cheap, quick to modify and support exploration of alternative ideas. Low fidelity prototyping allows for the generation of ideas, answers questions, reveal new information and requirements and can visualise workflows. Although they can provide proof of concept, low fidelity prototypes have limited use with regards to error checking

or navigational issues and sketches may not be refined enough to communicate ideas clearly. High-fidelity prototypes e.g. wire-frame models, provide more functionality and begin to take form of the final product but are resource-intensive to develop and time-consuming to modify (Preece et al., 2019).

As described under earlier, the findings from the interviews, together with the observation study and literature review, formed the foundation for the low fidelity prototyping. A simple paper prototype was made, combining ideas from the sketches and earlier identified requirements and assessed in a final round of evaluation with five users with the purpose to refine the existing user requirements and to identify additional needs and improvements. I attempted to structure the information in a logical and consistent manner, so that the features are presented in a known and understandable format, similar to how the current EHR interface appears at present.

Users often find it easier to respond to a suggested approach demonstrated by a prototype than a text description (Teixeira et al., 2012). The users were encouraged to think aloud and verbalise their thoughts while interacting with the prototype. The evaluation provided feedback that was utilised to generate the final version of functional requirements.

4.5.5 Data Analysis

After each iteration of requirement elicitation, all requirements that had been discovered were classified and organised. Notes and sketches were taken under each observation and interview, and these were analysed immediately after each session. Identified requirements were organised into related groups, based on where in the genetic analysis workflow they were needed. I checked for conflicting requirements that the users would need to prioritise in the next iteration and documented the requirements by using natural language as user stories and UML use-cases.

4.6 Ethics

Ethics revolves around principles, rules and guidelines for assessing whether actions are right or wrong and apply also to research activities. Respondents who participate in research must receive all the information necessary to form a reasonable understanding of the research field, of the consequences of participating in the project and of the purpose of the research. The participants have the right to discontinue their participation at any time without this having any negative consequences for them.

All my respondents were sent an information letter about the purpose of the assignment, as well as information that participation was voluntary. I did not process any personal data or sensitive genetic information in my master's thesis. As I am a colleague and known to all the participants, there is a potential conflict of interest in relation to this research study, but I have had no influence concerning their participation.

5 Results

To explore the primary research objective for this thesis, I have chosen to elucidate the user's main functionality requirements that would support optimised integration and increased use of genetic information and to investigate how genetic information can be displayed and integrated with other patient information in the EHR and identify challenges. In this chapter, I will present the findings from the observation studies, interviews, and prototyping.

5.1 Discovery and Understanding - Observations

User-centred engineering requirements begins with discovery and understanding, which is the process of interacting with the users to discover their needs. It focuses on the users, which needs a system should support, and in what context.

5.1.1 Overview and Context

To understand and be able to describe the context of use for users of genetic information in DIPS, the organisation, the relevant end-users, IT systems that are used in the course of a normal day and general workflow, were mapped. The objectives of the observation study were to gain an insight into the tasks and workflows of the users. Relevant stakeholders involved in the genetic analysis workflow were identified, domain knowledge gathered, an overview of the work routines, time pressure and efficiency issues made and insight into of some of the user's needs gained. The observations assisted the further planning of the interviews.

5.1.1.1 IT-systems

Most of the diagnostic analyses at DMG are based on NGS. In 2021 the department performed 10 000 NGS-based diagnostic analyses and the laboratory is busy, modern, and well-equipped, focusing on automation of all processes. Currently the NGS IT-infrastructure is connected to secure big data storage facilities at University of Oslo, and not to the network of Oslo University hospital. ELLA LIMS is the only LIMS in this network, while the other LIMSs in use are implemented in the hospital IT-network. See figure 5.1.

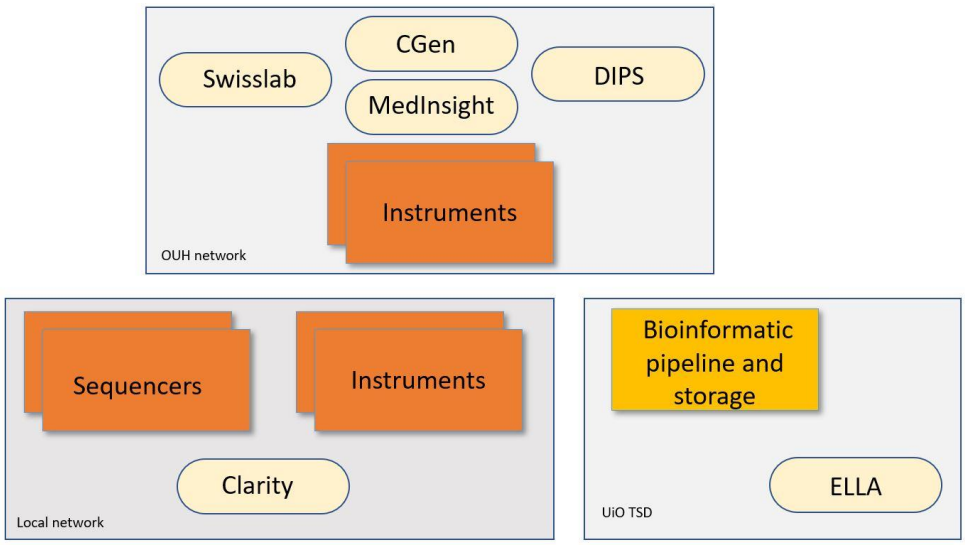


Figure 5.1 Overview of IT systems and networks at DMG

- DIPS:DIPS is the EHR in use at Oslo University Hospital. All genetic consultations and test results are documented here.
- LIMS Swisslab (SWL): This is the main LIMs for DMG and is used to register patient samples with a genetic test order and their results.
- LIMS Clarity + ELLA: These two additional IT systems are involved with the laboratory analysis workflow. All the analysis and sequencing instruments are integrated into Clarity and make it possible to track a sample as it is being analysed. Clarity and the DNA sequencers are on a local network. Interpretation of the genetic test results and DNA variant analysis are done in the ELLA software. ELLA, together with the raw sequencing data, are stored in the University of Oslos', Services for Sensitive data (TSD) infrastructure.
- Clinical systems- Cgen, MedInsight: Additional clinical systems that are used by the clinical geneticists and genetic counsellors e.g., to store patients' family information.

5.1.1.2 Users involved in a genetic analysis workflow

The primary users of DIPS are the referring clinicians, genetic counsellors and the clinical and laboratory geneticists and secondary users are patients, laboratory staff and bioinformaticians.

- Patient: Members of the public that use healthcare services and receive a genetic test
- Referring/treating clinicians: Healthcare professionals making a diagnostic, treatment, or preventative decision or recommendation, based on the genetic information. Clinical/laboratory geneticists: Professionals interpreting the clinical implications of a patient's genetic data who may work within the laboratory setting or outside the laboratory.
- Genetic counsellor: a health care professional who provides information to individuals and families about genetic conditions
- Laboratory: Hospital testing laboratory
- Bioinformatician: Individuals responsible for the integration of genetic data into local EHR and other clinical systems

5.1.1.3 Genetic analysis workflow

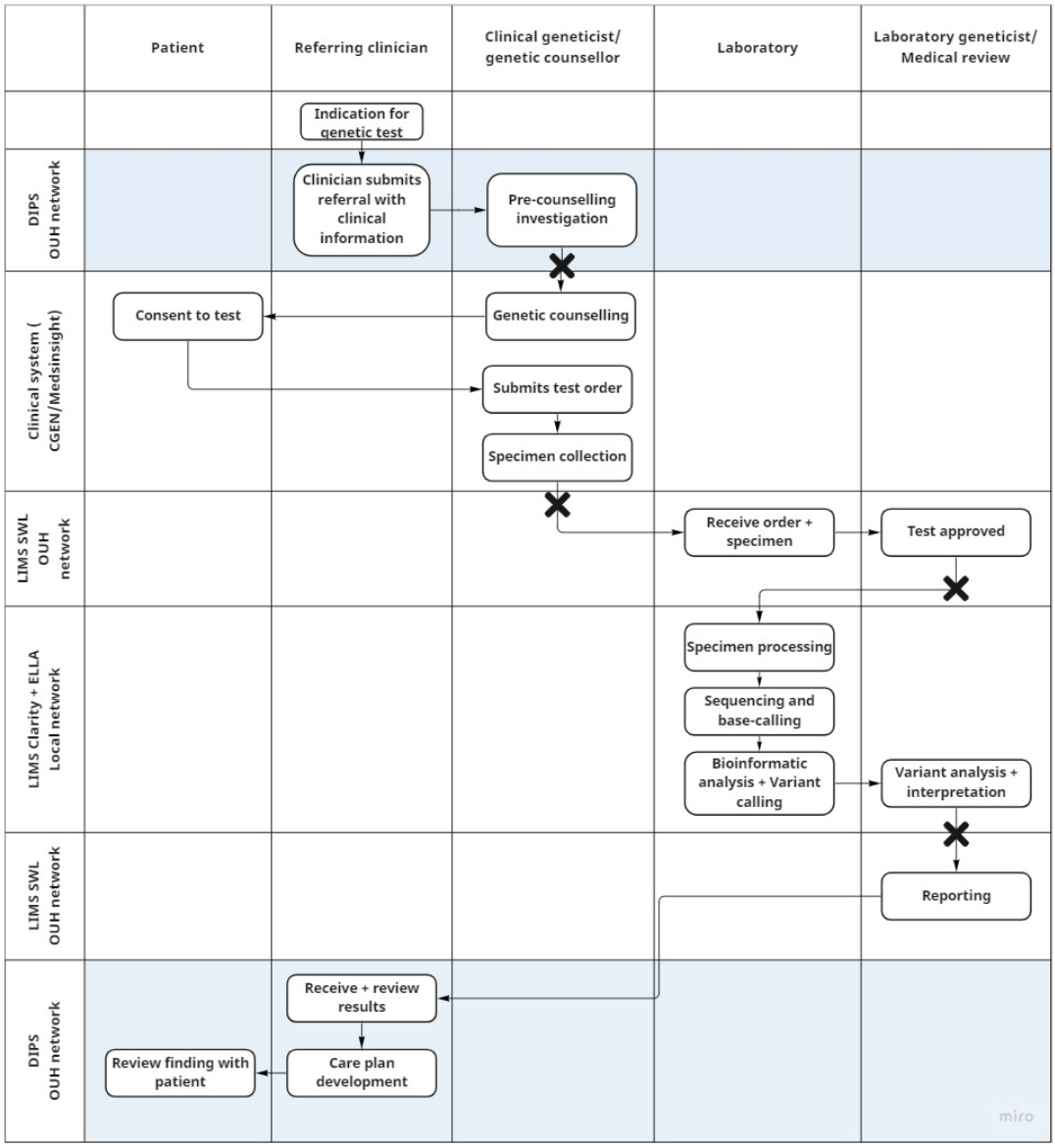


Figure 5.2 Genetic analysis workflow DMG
X :Manual step for data transfer between HIS

As shown in figure 5.2, testing is initiated by the referring clinician who identifies the patient who would benefit from a genetic analysis. The referral is sent electronically in DIPS, in PDF format, with varying degrees of information. The clinical geneticist reviews the referral, completes a family history-based risk-assessment, performs a clinical examination of the patient, before ordering an appropriate genetic test on a paper form, including relevant clinical data to aid in evaluation and interpretation of any findings. Pre-test genetic counselling of patients also includes information about the potential implications of the genetic test results to themselves and their families and consent to the testing. After a blood sample is collected and received by the laboratory, the test order is medically reviewed in LIMS SWL to confirm that the appropriate genetic test has been ordered or to check for duplicate genetic testing. The sample is then processed and sequenced by the laboratory. The results are interpreted and compiled into a report, which is sent from the LIMS SWL to DIPS in PDF format, and the referring clinician can determine relevant patient treatments and care.

The workflow model in figure 5.2 clearly reveals the interoperability challenges between the different HISs, and several steps where patient information has to be copied manually from one system to another. The IT-structure is complicated, susceptible to errors, and inefficient.

5.2 Discovery and Understanding - Interviews

The aim of the semi-structured interviews was to identify how genetic information is organised in DIPS, study the detailed workflows within DIPS and identify challenges and possible solutions. Interviews were conducted between June and August 2022 at OUH.

5.2.1 Current use of DIPS

The respondents described and demonstrated their tasks, provided comments, identified problems, suggested solutions and made sketches regarding the possible future display of genetic information in DIPS. I have grouped the description of tasks according to a simplified workflow, illustrated in figure 5.3.

Figure 5.3 Simplified genetic analysis workflow

5.2.1.1 Pre-schedule investigation and genetic consultation

Reviewing a referral for a genetic test, in addition to examining the patient, involves checking if any other relevant genetic test has previously been performed. This entails searching through DIPS for relevant test reports or reading journal notes to see if any genetic test is mentioned. It is critical to check if any other family member has had a genetic test, which is complicated to find in DIPS. Again, it entails searching through journal notes in DIPS, with the limitation that only results from OUH are available. The clinical geneticists have an additional clinical system, (MedInsight/Cgen) which can be used to find family information if it exists but does not communicate with DIPS. The notes from the genetic consultation are entered into DIPS as text blocks with no structured format. The patient's consent and information that a paper test order has been sent to the laboratory are documented only in the journal notes.

5.2.1.2 Test ordering

There are several different laboratories at Oslo University Hospital that perform genetic tests, in addition to external laboratories both in or outside of Norway. Only a small selection of genetic tests can be ordered electronically in DIPS.

To generate an electronic order in DIPS, the "LabResult" tab is opened, and "New Test order" chosen. Under "Analysis Groups" there is a choice for "Genetic tests". Having chosen the relevant tests, the order form is sent to the respective laboratory. However, only a few simple genetic tests and PGx testing from the Department of Biochemistry and non-invasive prenatal testing from the Department of Genetics can be found in this display.

For all other genetic testing, the relevant order forms must be found elsewhere, e.g., the hospitals webpages, filled out and sent to the respective lab independently of DIPS. This information isn't directly registered in DIPS but maybe found under ordering of blood sample taking with a note that a sample has been sent to DMG, or in the journal notes. It

is not possible to identify which tests have been previously ordered, without manually searching through journal notes to find a mention of genetic testing or to find genetic test reports. As genetic testing is primarily ordered through paper requisitions, the clinical features of the patient and family history are often poorly described, incomplete and often difficult to interpret from hand-written descriptions.

5.2.1.3 Genetic test approval and interpretation

Before approving the genetic test order, the laboratory geneticist will also check if any previous tests have been undertaken. They can search in DIPS for genetic test results among all the other journal notes but can also use the laboratory LIMS SWL to search for earlier test orders. To ensure that the correct genetic test has been ordered, the laboratory geneticists depend on the, often limited, clinical information on the test order form or what they can find in the patient's journal notes. The patient's clinical information is also crucial when interpreting any genetic findings, and the same search process is repeated.

5.2.1.4 Delivering genetic test reports

There are several methods to find and read laboratory results in DIPS, see figure 5.4 and 5.6. For tests that have been ordered electronically through DIPS, the results can be found under "LabResults". Here you can choose "Test Orders" and by choosing the relevant order, the results will be displayed There are also options to visualise the results graphically. Alternatively, "All tests" can be chosen, and all the results from the patients test orders are shown together, with tool-tip functionality for additional information.

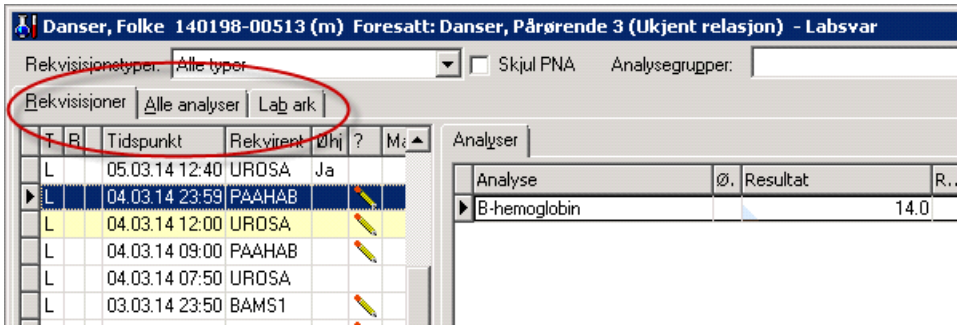


Figure 5.4 Finding test results in DIPS

However, genetic tests that have not been ordered electronically are more difficult to locate. Results from the Department of medical genetics are displayed as journal documents, under "Medical genetics" group but with many other document types e.g. journal notes, telephone notes, discharge summaries, admission notes. See figure 5.5.

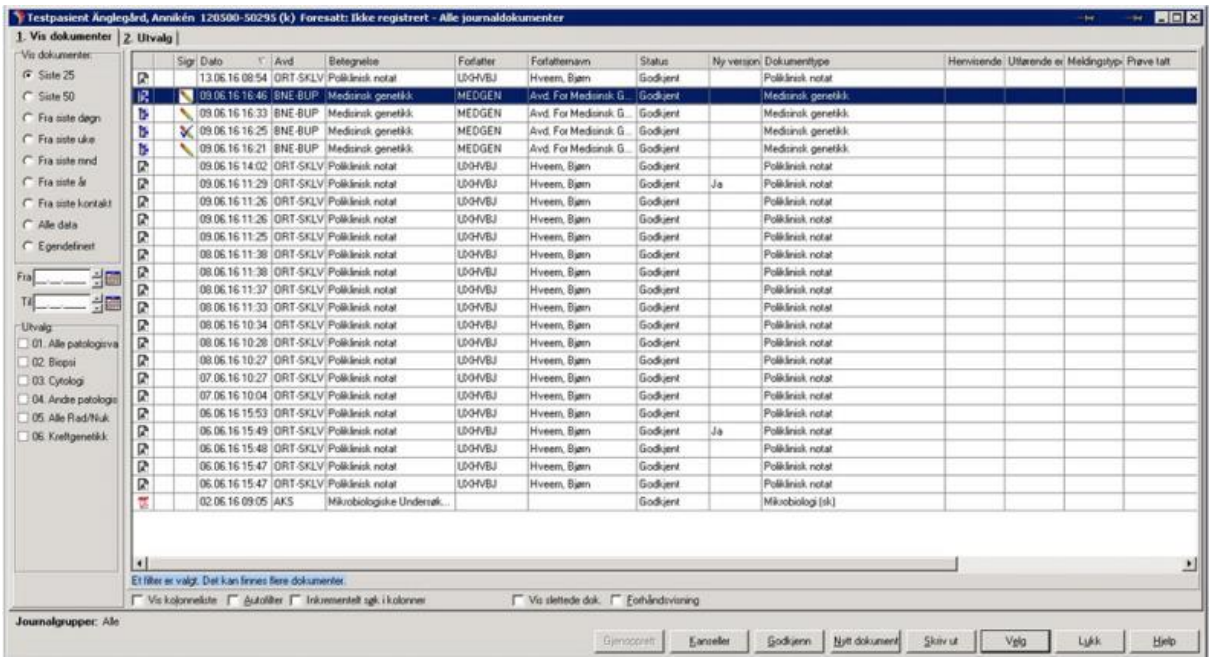


Figure 5.5 Finding genetic test results in DIPS

It is also possible to filter and display only journal documents from the Medical genetic group. There are, however, exceptions, e.g. results from the cardio lab, dept. of medical genetics are displayed as Journal notes. Chosen results are displayed in PDF format in their own window and can consist of several pages.

Genetic results from the department of biochemistry are also displayed as journal documents but under the journal group, "Medical Biochemistry". This includes genetic tests from New-born screening and hormone laboratory. Genetic tests performed in external laboratories are generally scanned under the Medical Genetics group.

All the results, except for the few genetic tests available by electronic ordering, have to be scanned into DIPS under the correct journal group and type, which is mostly an automated process from results generated within OUH. However, external documents are scanned manually and can be placed incorrectly.

5.2.1.5 Location of genetic information in DIPS

The genetic information comes from a range of sources, and it was evident that the laboratory performing the genetic test determines where the genetic information is documented in DIPS. Another factor influencing where and how the genetic information is documented is the ordering clinician's department. All types of genetic information in DIPS are displayed as PDF documents or paragraphs of text, although the document type in DIPS may be different, and no structured genetic data is stored in the system. Most of the participants were unsure where they could find genetic results from PGx, tumor genetics or New-born screening, and had to actively search in DIPs or ask a colleague.

Table 5.1 Ordering and genetic test results in DIPS

Department	Laboratory	Type of genetic test	Electronic ordering	Results under lab analyses	Document type in DIPS
Medical Genetics	Cancer	Sequencing	No	No	Medical genetics
	Cardio	Sequencing	No	No	Journal note
	General	Sequencing	No	No	Medical genetics
	Foster diagnostics	NIPT	Yes	Yes	Not relevant
Medical Biochemistry	Pharmacology	Pharmacogenetics	Yes	Yes	Not relevant
	Hormone	Sequencing	No	No	Medical biochemistry
	Newborn screening	Sequencing	No	No	Medical biochemistry
Pathology	Molecular pathology	Sequencing	No	No	Pathology

5.2.1.6 Use of standards and vocabularies

Most of the respondents did not know of any standards that were in use except for the HGVS nomenclature to describe DNA variants. LOINC is not used for the genetic analysis test names, and SNOMED-CT not used to represent a genetic diagnosis. ICD-10 codes are available in DIPS, but not used actively in relation to genetic information. However, the Norwegian laboratory coding system was used for billing. HPO terms are not available in DIPS but are greatly missed by the clinical geneticists. None of the respondents were aware of other genetic standards in use such as HL7 FHIR, GA4GH or OpenEHR.

5.2.1.7 Level of interoperability in current workflow

There are very low levels of interoperability in the current systems. Using the interoperability described in Table 2.2, the workflow can be described at level 2, where the information is added manually to SWL and can only be read within the laboratory. For genetic purposes, DIPS can be defined as Level 3, with information e.g., external lab reports, having to be manually scanned and registered into the system. There are no functionalities that can satisfy levels 5 or 6, such as defined information being forwarded to the correct location in DIPS.

5.2.2 Identification of Challenges

From both the observations and the interviews, many challenges were identified which are categorised in this section, shown in figure 5.6.

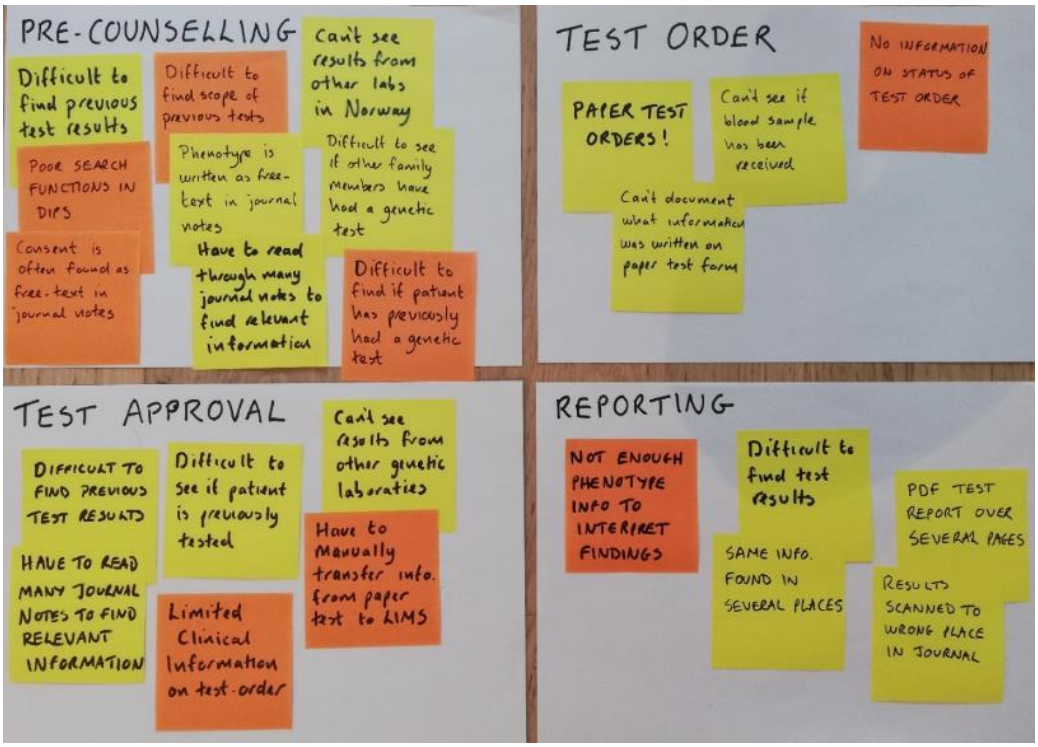


Figure 5.6 Challenges identified during interviews

5.2.2.1 Inadequate information in referrals and order forms

Neither the referral or test order forms contain any mandatory information sections which results in lack of relevant information for the clinical and laboratory geneticists. Referrals are generated within DIPS as a PDF document, but test orders are on paper order forms. This results in the search of information, either in PDF documents in DIPS, accessing auxiliary systems such as Cgen or MEdInsight or e-mails/telephone with the referring clinician. Once sent, a referring clinician has no documentation of what was written on the order form. *"Sometimes I have information about the patient, but I am unsure if I mentioned it on the test order-form",* answered one respondent when asked about the routines for ordering genetic tests. Most of the respondents searched for clinical symptoms or patient phenotype, and family information, either in regards in choosing the correct genetic test or in the process of DNA variant interpretation. Searching for results of earlier genetic tests was also frequently mentioned.

5.2.2.2 Organisation of genetic information in DIPS

All the respondents stated that finding the genetic information that was relevant to their workflow was difficult, labour-intensive, time-consuming, and potentially error prone, as stated by one participant, *"There is never-ending number of clicks and if I'm lucky, I'll find the information I'm looking for".* The laboratory geneticists experience the identical challenges as the clinical geneticists in finding relevant evidence and in contacting referring clinicians or external laboratories for necessary information.

All the relevant information that the respondents searched for are located in PDF documents or as text blocks in journal notes, and there are only limited methods for searching. During the observations and interviews, it was noted that there several different methods for searching were used, and most respondents used a combination of approaches to find the information they needed. During one of the interviews, a genetic

test result from a European laboratory describing critical genetic results was found in the journal notes from the referring department just by chance.

Genetic information is not one of the types of critical information that can be registered in DIPS and, therefore, disease defining, or risk actionable genetic variants are not displayed under the Critical Information tab but remain hidden in a PDF document. One respondent stated, *"you wouldn't see any genetic information if you didn't know it was there".* Several of the respondents stated that the same genetic information could be found in multiple places in DIPS and there have been no efforts to consolidate genetic information into one place. Genetic information is also displayed within genetic clinics notes, as journal notes in DIPS.

5.2.2.3 Lack of system integration with LIMS

Paper order forms received in the laboratory have to be manually registered in the SWL LIMS. Difficult to read handwriting and abbreviations are transferred to SWL and this is naturally a labour-intensive and potentially error-prone action. The referring clinicians have no means to follow the status of an order, one respondent said, *"I have no idea if the laboratory has received my patients' blood-sample, and often have to resort to a phone-call to check if everything is OK".*

Genetic test results are delivered from the LIMS to DIPS as a PDF document. Therefore, no structured genetic information is available in the DIPS with details of genetic variants, nomenclature, and interpretation. There is, consequently, no possibility for flagging of different genetic categories e.g., disease defining, PGx, carrier recessive or variants of uncertain significance or any forms of clinical decision aids and support. If the patient has had a genetic test earlier, it is challenging to find the scope of the test e.g., which genes were included, or which methods have been used.

5.2.2.4 Lack of national EHR system

In Norway, there are 5 genetic departments who provides genetic counselling and diagnostics within the field of hereditary diseases. However, the geneticists have no access to genetic results from other genetic departments out with their own. It is quite surprising that within the same health region, Health South-East, there is no integration between the two genetic labs located there, despite both using the same EHR platform, DIPS. Patients can have earlier performed a genetic test from another genetic department but unless this information is documented in journal notes, there is no alternative than to contact other genetic departments in search of information which must then be sent by fax or letter or communicated over the phone.

5.2.3 Mapping of Oppurtunities

Having discussed the challenges that each respondent experience during their work-tasks in a genetic analysis workflow, they were given the opportunity to describe which functionalities or data-elements they consider would be most useful or that they regard as missing in today's solutions. During the interview, the respondents were asked to describe or sketch how they thought genetic information could be displayed in DIPS, with the aim that the respondents could start to reflect on their functional needs. See figure 5.7.

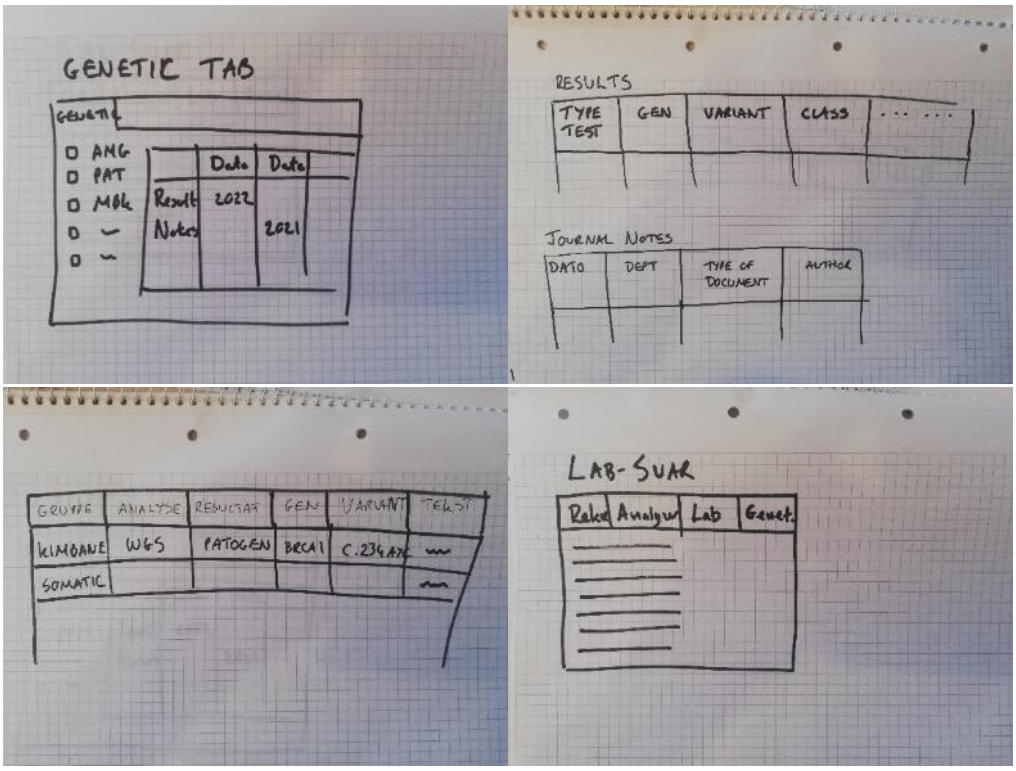


Figure 5.7 Sketches from interviews

It soon became apparent that this was not a straightforward task, due to the complexity of genetic testing and the amounts of information that could potentially be included. Many of the respondents were not aware of the possibilities that structured data can provide and the opportunities that this could give for clinical decision support. It was also difficult for several to imagine anything other than the current system and the limitations that they are fully aware of, as described by one respondent "*The functionalities I want are just not possible!".* To overcome some of these obstacles, the participants were presented with proposed clinical use examples from the literature to illustrate possible situations, see Appendix 2.

The needs and solutions from the interviews were summarised as follows:

- Electronic test ordering: This was a functionality wanted by all the respondents.
- Structured information: One particular functionality that was also mentioned by all the participants was coded representations of phenotypes. Providing family indexing and pedigrees was also mentioned, but primarily by clinical geneticists.
- Location of genetic information: There was a great need to have all genetic information from all providers located in one section of DIPS.
- Display genetic test results: All the respondents would like to see a detailed list over medically actionable DNA variants with nomenclature and interpretation, with links to more detailed information
- Clinical decision support: A mechanism for medically actionable genetic information to trigger an alert to the referring clinician or to trigger an alert if patient has previously had a genetic test was suggested.

Although there were many identical functionalities and needs that were identified from all the participants, there were a few differences. There were, for example, varying interests in the possibilities of Infobuttons in the test report to external knowledgebase systems.

Despite a genetic test report including detailed information of the genes involved and the reasons why a particular DNA variant is interpreted as medically actionable, the respondent told that *"I always google the gene mentioned in the test report to see if there is anything else I can find out. It would be helpful to have a link to a reputable source."* Other respondents were more sceptical to Infobuttons, as they felt that they would need to verify the quality of the contents.

5.3 Classification and Organisation - Prototype

After requirements discovery and understanding, the next step is classification and organisation of the unstructured collection of requirements, where related requirements were grouped under the relevant steps in the workflow. Using the sketches and ideas for solutions provided from the interviews, a simple paper prototype was made that included many of the concepts and functionalities described from the users.

The prototype consisted of several "screens" as shown in figure 5.8. See Appendix 3 for detailed pictures of each screen.

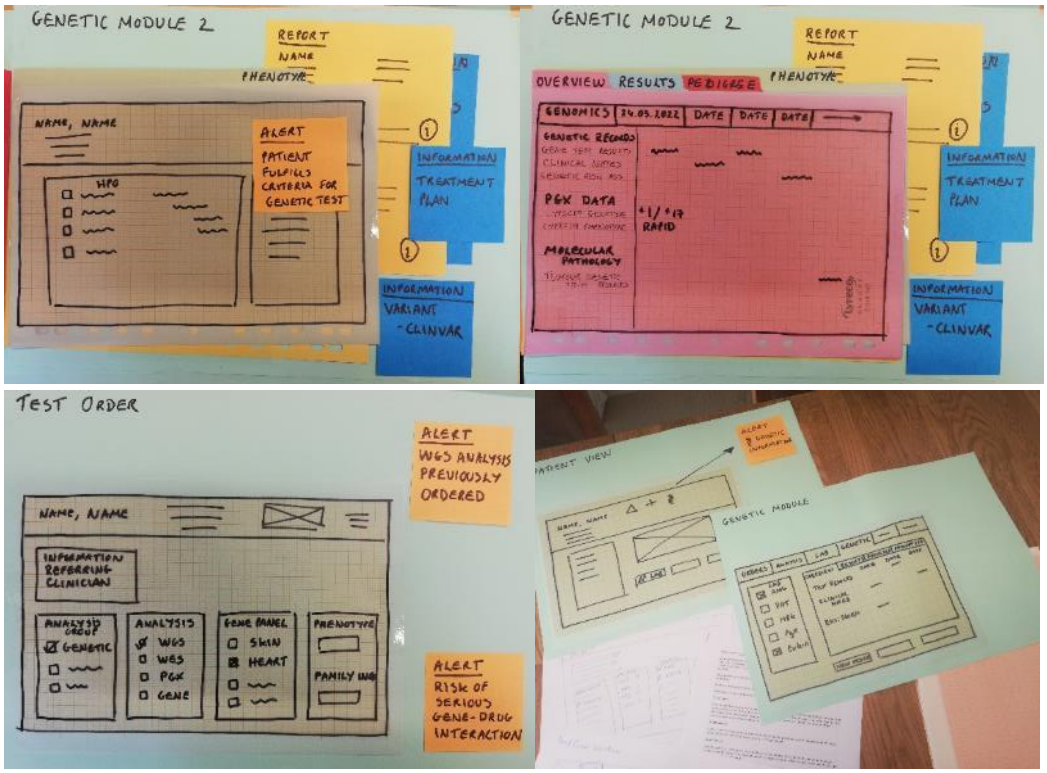


Figure 5.8 Paper prototype

The prototype was loosely based on the current EHR, but I deliberately didn't include any colours or unnecessary details, as the aim was to focus on functionalities and not design. The prototype doesn't include any "buttons" or moving parts, but the user could describe the actions they would take, and then move over to the next "screen".

The prototype was designed with the assumption that there was full interoperability between the LIMS and DIPS. I presumed that all genetic information was in a structured format using standard terminologies, vocabularies, and genetic interoperability standards such as HGVS, LOINC and HL7-FHIR. This would allow for all elements in a genetic test report to be mapped to structured data elements, including test names, DNA variant descriptions, variant interpretations, and clinical phenotypes.

Functionalities included in the paper prototype:

Patient view: An indicator to show that there is documented genetic information for this patient.

Genomic module, located in the "Lab" section of the EHR, with several tabs:

- Overview: An overview of all genetic information, test results and genetic consultation notes, from all providers, sorted after date and genetic category, e.g., germline, somatic or PGx. There are possibilities to filter which type of genetic category is of interest, and the user can navigate directly to the most relevant information.
- Results: In this tab, the patients genetic test results are shown as a table with information of gene and variant location, and interpretation class. Infobuttons that will take the user directly to external knowledge base systems for gene and DNA variant. Link to structured genetic report with detailed sections on methodology and test scope. Again, with Infobuttons to relevant treatment plans or more detailed test methodology. By clicking on a DNA variant, a visual representation of variant class over time will be shown.
- Pedigree: Representation of patient pedigrees and relevant family information
- Phenotype: Overview of patient's phenotype in structured format, e.g., HPO terminology that have been previously registered by the treating clinician. An alert will show if the patient fulfils criteria for a genetic test, depending on the registered clinical information. Possibilities to register new clinical information.

Test Order: Electronic ordering of genetic tests, sorted under different type of genetic analysis e.g., Genomic or single gene, or single family variant. The user can choose from an up-to-date list of available genetic tests, and a dynamic choice of phenotypes, depending on which test was chosen. If a patient has a previously registered genomic test, an alert will appear, and the user can choose to reanalyse the original raw data.

5.3.1.1 Level of interoperability in prototype

The intended interoperability in the prototype is the highest level, level 6 (see Table 2.2), where information is forwarded automatically to the EHR, into the correct location and with defined data elements, avoids all re-typing of information. This allows for automatic categorising and interpretation of the genetic and clinical information. The structured exchange of data and the use of standard terminologies and nomenclatures allows for full semantic interoperability.

5.4 Prioritisation and Negotiation: Prototype Evaluation

When multiple end-users are involved, requirements will often conflict, and it is important to prioritise and resolve requirement disagreements. To validate the findings from the interviews and to evaluate if the paper prototype was according to the respondents' requirements and needs, 5 end-users, not all previously involved in the requirement elicitation process, were invited to assess the system. Not all the suggested functionalities were incorporated into the prototype but enough to understand the participant's main requirements. They were first informed briefly about the limitations of the prototype and were then asked to follow their typical routines for searching for genetic information in the EHR, test ordering and interpretation of the final test report.

5.4.1.1 Overall impression

The prototype was met with much enthusiasm, "this would be fantastic!" was said several times. The participants found it logical, simple to find information, and easy to get an overview of the genetic information. The prototype was considered to be extremely timesaving. Having "clicked" through the different tabs one participant exclaimed, "This would save me hours every day".

Several of the participants first understood the possibilities that structured data and high levels of interoperability can give when evaluating the prototype, with comments such as, "Lot of functions that I didn't know were available". However, there was also sceptism that such a solution would ever be developed.

Many of the identified requirements were common for both pre-counselling and test approval, and the same information was needed but in different steps of the genetic workflow.

5.4.1.2 Patient view

All the participants found the icon in the patient view, indicating that there is genetic information registered for this patient, very useful. It was sufficient with an indicator, and not necessary for an alert for all clinicians, as this is often irrelevant information for most of the treating clinicians. However, alerts for risk of gene-drug interactions were judged important. One participant responded, "with just one click, I can be in a particular patient's genetic workspace".

5.4.1.3 Genomic module

Overview tab: The participants thought that the overview of all genetic test results, from all providers was extremely practical and efficient. One participant thought that the possibility to see both germline and somatic genetic testing together was particularly useful. In this view they could quickly get an overview of the patient's history and which test reports and clinical notes were most relevant. It was noted that the clinical notes should be limited to genetic consultations or where a tentative genetic diagnosis was concluded to limit the number of irrelevant documents. Two participants suggested flagging of genetic results where a medically actionable variant was detected and the associated clinical notes to enable easy identification of relevant information. Solutions for flagging of not only medically actionable variants but also VUS, risk alleles or PgX findings were also considered.

Results tab: This tab was very interesting for all the participants and created quite a few discussions. Discrete genetic variants with gene name, variant nomenclature and class are listed, which also gives the possibility to see aggregated results from multiple genetic tests overtime, which was of particular importance for many of the participants. This display allows for tracking of which specific genetic test has been previously ordered, and the possibility, via Infobuttons, to see the scope of the test. This functionality however is possibly only of interest for genetic specialists. It was also suggested that a variant should be flagged if the laboratory had reclassified the variant and changed the variant classification, again a function that perhaps only special interested clinicians would use, but important all the same. By using structured data from the LIMS, there is almost unlimited amounts of metadata from genetic testing which could be included in the EHR, and there were several discussions of which data elements were important and relevant. Details of the quality of the genetic results was example of a data element that was found to be too detailed to include in this display.

By "clicking" on a genetic test result, a full-text genetic test report is shown with Infobuttons to relevant external knowledge systems and relevant treatment plans, which were considered very useful by the participants.

Pedigree tab: All the participants would use this tab, both to register information and to find relevant details of family history.

Phenotype tab: Good clinical information is important both for the choice of genetic test and for the interpretation of the results. Structured phenotype information was a critical need for all the end-users and the functionality to register this information was much-admired by all the participants. An alert if a patient fulfills criteria for a genetic test based on phenotype, but is not tested, is built-into the system.

5.4.1.4 Test-order

This is a functionality that has been greatly missed in the current EHR and the participants all mentioned that paper test orders are a source of many difficulties. One participant suggested that instead of choosing a genetic test panel, that the registered phenotype could dynamically suggest which genetic test would be most relevant. This was confirmed as a very appropriate modification from the other participants. The importance that an up-to- date list over available genetic tests was critical for this function was commented on by one of the participants. To be warned that the patient already has been tested was also very useful as several of the users had experienced unnecessary multiple testing of the same patient but from different clinicians.

5.4.2 Refinement of requirements

From the evaluation and testing of the prototype most of the functions and needs identified from the interviews were confirmed, indicating that they are necessary and core requirements. However, several new functional requirements were also identified, some were modified or down prioritised.

5.4.2.1 New identified requirements

- Structured referrals with mandatory data elements
- Register and display patient consent
- Possibility to search for genetic variants
- Take out reports
- View genetic related diagnoses in the genomic module and flag clinical notes that are associated with diagnosis.
- Flagging of genetic categories in overview tab
- Flagging if genetic variant has been reclassified

5.4.2.2 Modified requirements

- The granularity of DNA variant interpretations in the Results view was modified due to concerns that there was too much complex information on display. Instead, the flagging of medically actionable variants was suggested, and the clinician could then click further on to the full report to see more detailed information.
- Initial worries of use of Infobuttons by some participants from the interviews were not confirmed in the evaluation, and these functions were considered important. There was agreement that an alert for available genomic information would not be useful for many treating clinicians and could lead to alert fatigue. This function was modified to the insertion of a visual icon on the patient view.

- Test ordering: changed from clinician choosing genetic test, then registering phenotype, to a dynamic choice of suitable genetic test after phenotype choices.

5.4.2.3 Removed requirements

No functionalities were judged as unnecessary or unsuitable.

5.4.3 Prioritised Requirements

When the participants were asked which requirements were most important or to prioritise the requirements, there was a wide range of choices, depending on their particular tasks in the genetic analysis workflow. Electronic test ordering was very important for the referring clinicians, so they could follow the progress of the analysis and ensure that they had provided the necessary information. The referring clinicians were also most enthusiastic about clinical decision support functionalities such as infobuttons and links to external knowledge bases, which could assist in understanding a complex genetic test result. The medical geneticists however prioritised the overview tab, were all the genetic information was displayed into one place, with flagging of critical results and notes. Structured phenotype information was also stated as a very crucial functionality, to aid in both choice of test and variant interpretation.

5.5 Documentation – Use-Cases and User Stories

The user requirements were written in a natural language, supplemented by an UML-Use case diagram.

The UML use-case diagram (Figure 5.9) represents the individual interactions between the users and the EHR genetic module. The use-cases capture high level communication needs and depicts the major activities the various actors can accomplish when using the system.

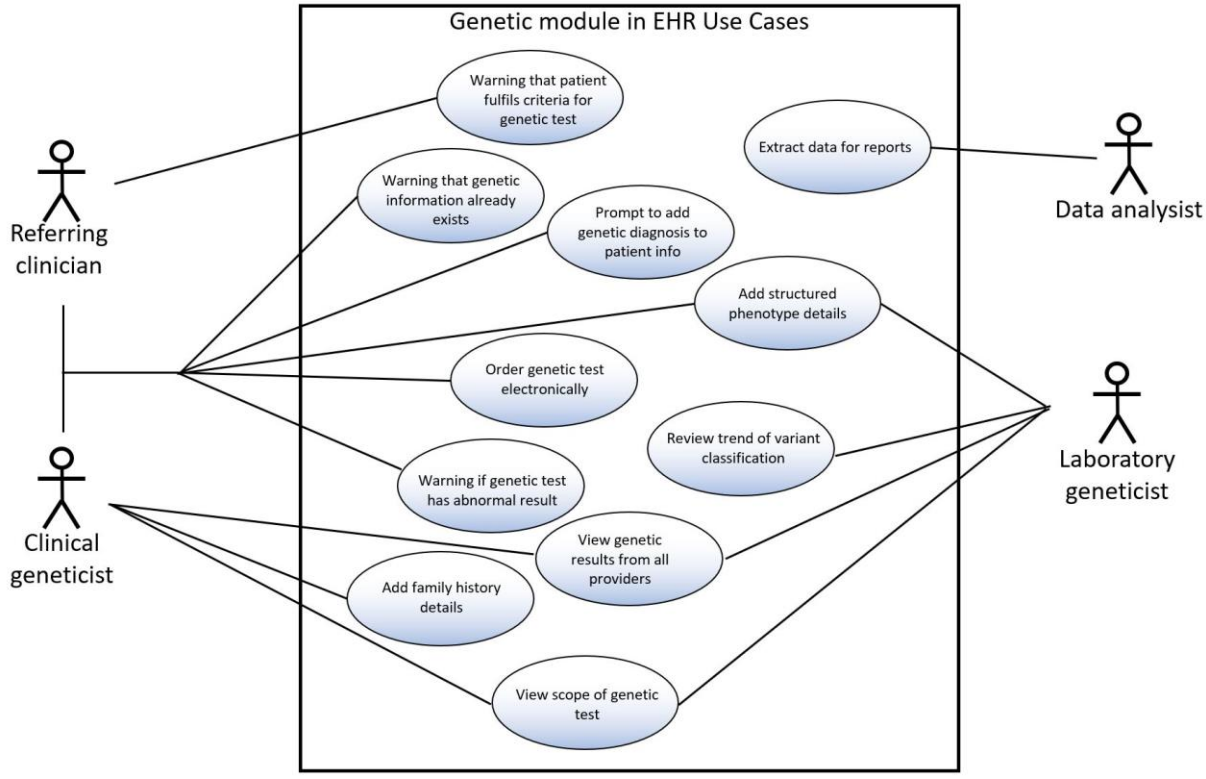


Figure 5.9 Use cases

User stories, listed in Table 5.2, were written in the end-users voice and in a structured natural language, without software jargon, to describe the identified requirements. The format used is simple to read and the single sentence that includes who, what and why of a feature, provides a shared understanding of the desired goal.

Table 5.2 User stories

	As a <type of user>	I want to <objective>	So that <benefit>
Test order			
1	Clinician for patients with rare disease	track which genetic tests have been previously ordered	I can avoid duplicate testing
2	Clinician for patients with rare disease	order a genetic test electronically	I avoid writing information manually and introduce errors in the test order.
3	Clinician for patients with rare disease	find an updated list of available genetic tests electronically	I can easily choose the most relevant test
4	Clinician for patients with rare disease	order a genetic test electronically	I can include structured clinical information
5	Clinician for patients with rare disease	order a genetic test electronically	I can document which tests have been ordered
6	Clinician for patients with rare disease	retrieve clinical data/phenotype	I can order the correct genetic test
7	Clinician for patients with rare disease	Search for patients with the same genetic faults	Patients receive optimal treatment
Pre-Counselling/Test approval			
8	Clinical geneticist	Easily see if there is genetic information registered for my patient	I can efficiently find relevant results
9	Clinical geneticist	See genetic test results from all health providers	I can avoid duplicate testing
10	Clinical geneticist	See genetic test results from all health providers	I can efficiently find relevant results
11	Clinical geneticist	see which genes have been included in previous genetic tests	I can order appropriate additional genetic testing
12	Clinical geneticist	See a display of detailed genetic test results with variant descriptions	I can efficiently find relevant results
13	Clinical geneticist	See a display of detailed genetic test results with visual flagging of genetic categories	I can efficiently find relevant results
14	Clinical geneticist	View genetic related diagnoses with visual flagging of clinical notes associated with the diagnosis	I can efficiently find relevant information
15	Clinical geneticist	See a display of aggregated results from multiple genetic tests over time	The patient receives optimal care
16	Clinical geneticist	order a genetic test electronically	I avoid writing information manually and introduce errors into the LIMS

17	Clinical geneticist	see which genetic variants have new interpretations	The patient receives optimal care
18	Clinical geneticist	retrieve clinical data/phenotype	I can interpretate relevant genetic findings
19	Clinical geneticist	Register patient consent	I know that the patient has received necessary information
20	Clinical geneticist	Receive structured referrals, with mandatory information elements	I avoid searching for information
Test report			
21	Clinician for patients with rare disease	genetic test report to include links to relevant external knowledge bases	I understand the implications of the findings
22	Clinician for patients with rare disease	genetic test report to include links to relevant treatment plans,	The patient receives optimal care
23	Clinician for patients with rare disease	view genetic related diagnoses in a problem list	The patient receives optimal care
Clinical decision support/aids			
24	Clinician for patients with rare disease	Be alerted if my patient has documented genetic information available	I can efficiently find relevant information
25	Clinician for patients with rare disease	be alerted if my patient fulfils the criteria for a genetic test but has not been tested	The patient receives optimal care
26	Clinician for patients with rare disease	be alerted before I order a genetic test if a patient has previously been tested	I avoid taking a new blood sample
27	Clinical geneticist	be alerted before I order a genetic test if a patient has previously been tested	I can order reanalysis of existing data
28	Clinician for patients with rare disease	be alerted before I administer a drug to a patient before Pgx testing	The patient can receive safe and optimal medications
Administration			
29	Administrator	be able to take out reports for the health directorate	they can follow the requirements from the Biotechnology Act.

5.6 Summary of Key Findings

A list of 29 high-level requirements obtained from multiple methods have been identified and grouped by actor and work-task.

5.6.1 Genetic Analysis Workflow

Mapping of the workflow revealed an IT-structure that is complicated, susceptible to errors, and inefficient. There are several different networks and HISs in use, with interoperability and integration challenges and low adoption of genetic interoperability standards. Patient data is transferred manually in at least 4 steps, with the loss of granularity and details of critical information at each stage.

5.6.2 Current State for Display of Genetic Information in DIPS

From the results, it is clear that the present EHR has no functionality for ordering, displaying or organising genetic information. Genetic information is found in multiple places, under multiple document types and only as free-text or in PDF format. No structured genetic information is available in DIPS. Genetic information comes from a range of sources and the laboratory performing the genetic test most often determines where the genetic information is documented in DIPS. There is no means of differentiating between different categories of genetic information and it is not possible to differentiate between genetic test reports with or without findings. The genetic and clinical information found in text strings and PDF documents is unavailable to trigger CDS and there are, therefore, no alerts or warnings related to genetic information in the current EHR.

5.6.3 User Needs and Solutions for Genetic Information in the EHR

A simple prototype representing the functional requirements identified from the users and applying the possibilities that structured representations of genetic information can provide, was constructed. The resulting prototype included functionalities for electronic test ordering, with linking to appropriate structured phenotype information. All genetic information was collected and displayed one place, the genomics tab, where the users could easily find relevant test results and genetic consultation notes, from all providers, both within the hospital and external laboratories. The use of structured genetic information allowed for both clinical decision support in the form of relevant alerts and warnings, infobuttons, links to disease-specific knowledge bases and the extraction of data for reports.

The participants concluded that the functionalities in the prototype would be extremely timesaving and ensure improved patient care. It was not necessary to include all available genetic information from the LIMS to convey the main requirements of the clinicians.

6 Discussion

In this chapter the implications of the results and their relevance to the research questions and existing knowledge will be presented. It also outlines the quality of the study and limitations of the work.

The objective of this study was to identify the requirements for the organisation and display of genetic information and investigate the possibilities for representation of this information in the EHR to ensure genetic data is readily available to the users and for clinical decision support.

The following research questions provide a basis for answering the main research objective:

1. What requirements do clinicians have for optimal use of genetic information in the EHR?
2. Which solutions are available to display and integrate genetic information with other patient information in the EHR?
3. What are the challenges of integrating genetic information with other patient information in the EHR?

To assess the research questions of this study, a literature search has been carried out and a framework of user-centred requirement engineering methods was used to identify end-users needs that would support the use of genetic information and demonstrated how genetic information can be displayed and integrated in the EHR.

Integration of genetic information into the EHR has been described as one of the major hurdles in the implementation of precision medicine (S. J. Aronson & Rehm, 2015). In this study, the focus has been on user involvement and utilising their knowledge to better understand factors that can contribute to poor usability and to gain insights on how the display and integration of genetic information can be improved and used with increased efficiency.

6.1 What requirements do clinicians have for optimal use of genetic information in the EHR?

6.1.1 Genetic Analysis Workflow

Laboratory methods have rapidly advanced from testing for a few DNA variants in one or two genes to evaluating thousands of variants across hundreds of genes, resulting in increased complexity of data analysis pipelines, interpretation, and the need for many HISs. The genetic analysis workflow at DMG is no exception, and the complicated workflow identified is typical for most genetic laboratories. Walton (Walton et al., 2020) describes a similar situation, having to write scripts to convert structured laboratory data into HL7 segments for import into the EHR. However, at DMG, the LIMS is unable to store any structured genetic information, so the laboratory must resort to manual data transfer, with the reduced quality that this entails. The pre-existing design and constraints of the LIMS and data-transfer methods are also commented on by Carter (Carter et al., 2022) who concludes that implementation of solutions such as HL7-FHIR,

which would provide a solution to the problem, are limited in laboratories. DMG are in the process of implementing a new LIMS, with innovative functionalities for structured genetic information and data transfer, which will allow for new possibilities and improvements.

6.1.2 Genetic Information in DIPS

A clinicians' ability to use and interpret genetic information depends on how the data is displayed in the EHR. The work in this thesis shows shown that it is not the clinical content that determines where the genetic information is documented in the EHR, but the laboratory performing the genetic test, in line with a previous study (Shirts et al., 2015). The authors also find that genetic information is displayed in different places in the EHR, with no effort to consolidate the information into one place, and most often in PDF documents and text blocks. In DIPS there is no possibility to display disease defining genetic information, whereas others (Ohno-Machado, Kim, Gabriel, Kuo, & Hogarth, 2018) have reported that the presence of clinically significant variants are entered in the EHR as "allergies" or as problems in a "problem list". This information can could then trigger a warning to the clinician, but this functionality is not available in DIPS. There is no possibility to differentiate a normal genetic test result from one containing medically actionable information, which results in the clinicians having to open and read many documents in search of relevant information.

6.1.3 User Requirements for Genetic Information in the EHR

The participants' sketches and input from the interviews and observations were reformulated into user requirements and categorised related to their workflow. The user requirements formed the basis for the paper prototype that was designed and then evaluated by the participants. One of the principles of the UCRE process was then ensured by the involvement of users through all stages of the method. Teixeira et al. (Teixeira et al., 2012), emphasises that user involvement early in the process is important for understanding user requirements for solutions that reflect the healthcare worker's complex everyday life.

This study showed that there was high accordance between the core requirements identified from observations, interviews and sketching, but there were identified several more detailed requirements through the evaluation of the prototype. This is also in accordance with other studies that have shown that prototyping is a most effective method for UCRE (Teixeira et al., 2012).

There is very little literature available that describe or mention user requirements of clinicians with regards to use of genetic information in HISs but overall the identified requirements are in accordance with findings reported by Ayatollahi (Ayatollahi, Hosseini, & Hemmat, 2019). The authors identified requirements for integrating genetic data into the EHR from the literature and then invited medical geneticists to determine the most important requirements. However, the requirements were very vaguely described such as visible genetic test results, information sharing, usable interface, and use of standard terms. Other studies (T. M. Herr et al., 2015; Aly Khalifa et al., 2021) have focused on the importance of the interpretation of genetic test results and clinical decision support systems. Contrary to these findings, most of the requirements identified in this study are related to the organisation and display of genetic results in DIPS. These results could provide evidence that the current status for integration and display of genetic information has great potential for improvement. This result also casts light on the importance of the

user's knowledge of genetic clinical decision support and being fully aware of the opportunities that exist when implementing structured genetic information.

6.2 Which solutions are available to display and integrate genetic information with other patient information in the EHR?

The simple paper prototype illustrates how genetic information can be displayed and integrated with other patient information in DIPS. The model is based on solutions found in the literature, with functionalities derived from the interviews. It included functionalities for electronic test ordering, structured phenotype information and structured genetic information from the LIMS. The display of all genetic information was organised in one location, a genetic module, defined by clinical content and not the providing laboratory.

Electronic test ordering enables the EHR to document which tests have been ordered and the structured representation of the genetic test name e.g., LOINC terminology, allows the information to be organised in the genetic module/tab in the EHR. Implementing HPO terminology in the EHR would facilitate structured phenotype descriptions, which would ensure the laboratory receives the relevant clinical information they are so dependent on for choice of genetic test and interpretation of the results. The return of structured results, using HL7 FHIR genomics representation, in the form of discrete genetic variants and their interpretations from the LIMS to EHR is also a vital functionality. This allows, not only for the display of genetic variants in the results tab in the genetic module but also enables flagging of medically actionable variants and the implementation of many CDS functionalities. Ineffective navigation and genetic information that is difficult to locate, makes it more likely that clinicians will be unable to find and act upon important genetic information sometime after initial testing (Shirts et al., 2015).

Interestingly, the results of this study showed that it was not a requirement to display all possible discrete genetic elements in the EHR. The end-users concluded that the display of genetic information would be too complex for non-genetic specialists if too many details were included. This is a novel point that is not described in the literature, which is often technology driven.

The overwhelming positive responses from the paper prototype indicates that there is a huge room for improvement from current solutions. Many of the participants were under the impression that the functionalities that were illustrated by the paper prototype, were not possible in a real-life setting. However, several of the case-studies reviewed in the literature study describe successful implementation of similar solutions. The PennChart Genomics Initiative (Lau-Min et al., 2022) have recently reported that they have successfully linked orders and results from the genetic testing laboratories with discrete data in the EHR. They have used the EHR vendor Epic Genomic Module which integrates ordering and genetic test reports into the user's workflow, can display genetic results, translate them for a variety of users and make them actionable at point of care.

6.3 What are the challenges of integrating genetic information with other patient information in the EHR?

The challenges identified in this study are consistent with what has been reported in the literature. The literature review revealed an extensive list of challenges that were common to many of the case-studies in the review.

Lack of interoperability between the HISs in use today is one of the major challenges, both in the literature and in this study. Any improvements of transfer of genetic information to and from the EHR must also recognise the pre-existing constraints of the LIMS and data transfer methods available. The only information that is sent electronically between DIPS and the LIMS today is a PDF genetic test report and all test ordering is done with paper test orders. This demonstrates a very low level of interoperability.

The ability to store individual genetic variants in a standardised and reliable manner is equally important as interoperability. The LIMS in use at DMG is unable to deliver discrete genetic information and DIPS is not ready to receive and display the information. However, there are no international recommendations for the display of genetic information in EHR that could specify what level of granularity should be included or not. Several of the participants were concerned that the genetic information displayed in the prototype could be too complex for clinicians with limited genetic knowledge.

DIPS is not able to send accurate clinical and family information to the LIMS. A major hurdle that restricts the use of genetic information in the clinic is the overwhelming amounts of data that must be analysed and interpreted. Describing which tools and methods were used to identify whether a variant is pathogenic or not is also critical information, as different labs may come to different conclusions depending on the information available to them at the time. Effective variant interpretation is dependent on large datasets with phenotyped-linked information, and incorporating phenotypic data information in the variant classification process is essential to solve the bottleneck of variant interpretation (Furness, 2017). Genetic phenotypes are also poorly defined and standard definitions are lacking to facilitate delivery of phenotypic information from the EHR to the laboratory. It is therefore not possible to implement accurate electronic ordering with well-defined phenotype information.

The complexity of genetic information, the lack of necessary data standards and genetic functionalities have all contributed to the gaps between genetic data generation, interoperability and usability (Carter et al., 2022). The same challenges are found in this study and pose a significant challenge to all the end-users in the ability to generate, communicate and use genetic test results.

During this study, it has become apparent that the situation of integrating genetics into the EHR is even more complex than originally anticipated. This study has focused on one LIMS and one EHR, but in reality, there are numerous different LIMS and EHRs that would require integration with each other to enable successful integration of genetic information with patient information in a seamless manner. In addition, the systems should also allow sharing of variant and gene-level interpretations and be linked to external knowledge repositories. Aronson and Reim (S. J. Aronson & Rehm, 2015) describe a precision medicine HIS which illustrates some of the complexities involved, figure 6.1.

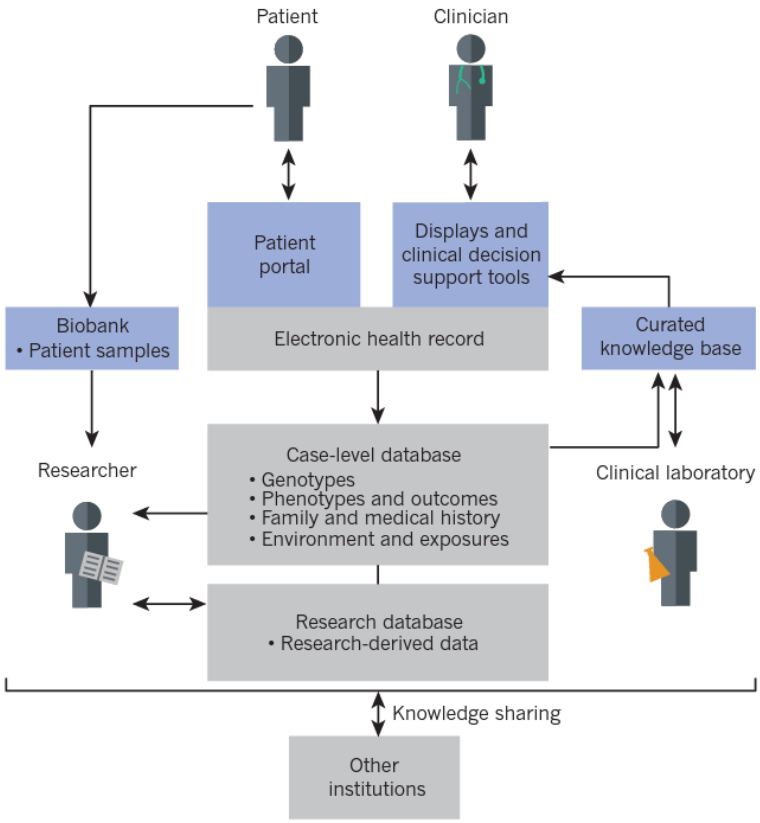


Figure 6.1 Precision-medicine ecosystem (S. J. Aronson & Rehm, 2015)

6.4 The potential of improved integration of genetic informatin in the EHR

Having identified user needs, investigated how genetic information can be displayed and integrated in the EHR and gained an overview of the challenges involved, the opportunities that improved organisation might result in, can be discussed.

The current complexity of ordering, viewing, and acting on genetic testing can be so overwhelming some clinicians may choose not to consider it at all. The complexity, however, can be rationalised by implementing some form of CDS to guide clinicians in how to use and interpret personalised data. It is necessary to move away from PDF reports where the genetic information is stored in such a way that it is unavailable to any form of secondary use. Genetic results in PDFs make it difficult to cross-check variants between different reports and make it impossible for CDS alerts based on genetic test results. Structured genetic information is the first step to precision medicine, with the aim to create a seamless interface with sequencing laboratories so test results integrate into the patient chart as discrete data – presented in a similar way to results from any other laboratory test.

Structured genetic information in the EHR can support many applications within precision medicine at the point-of care, as shown in Table 6.1.

Table 6.1 Examples of genetic CDS applications (Sen et al., 2019)

Genetic CDS application	Definition
Personalised medicine	Gene-drug interactions are used to inform the clinician of the patient's response to drug therapy. Alerts the clinician of dangerous gene-drug interactions
Risk analysis	Uses the patient's genetic information to assess the possibility of developing certain diseases. High risk triggers warnings and prompts increased screening or avoiding agents.
Diagnosis	Known genotype-phenotype relations are used to alert the clinician to disease causing variations and this increases diagnosis accuracy and reduces diagnosis time.
Newborn screening	Genomic analysis is used to enable the early detection and intervention for a select group of conditions. This facilitates the prevention of developmental impairments, delayed physical growth, severe illness, and death.
Somatic/tumour treatment	Knowledge about the mutations in the genetic makeup of tumour cells may have immediate therapeutic implications, which may alter cancer treatment choices.

Successful integrated decision tools within the workflow are necessary to enable users to search tests and relevant diagnoses and clinician alerts should be sent selectively at the appropriate step in the workflow to prevent alert fatigue. The ability to identify patients with specific genetic results within the EHR will also be important to recontact potentially affected patients and improving access to genetic testing for patients who are likely to benefit.

For DMG, the benefits of structured, organised genetic information in the EHR would be "a life-changer", as stated by one of the participants. As identified in the study, the users spend hours searching for relevant information both in DIPS and external systems, ordering tests with inadequate access to genetic knowledge, interpreting test results with insufficient phenotype information and manually transferring data with loss of information and risk of error. The participants commented that the functionalities in the prototype would increase referring clinician's awareness of genetic information and would make ordering a genetic test easier to integrate in their workflow. There were discussions regarding an increase in treatment choices based on genetic information in the future and that it required a maturation process before all clinicians are comfortable with ordering genetic tests. Even so, there was agreement that a visual display of genetic information would lead to improved use.

Structured, organised genetic information within a robust HIS infrastructure would allow for efficient and high-quality genetic testing for the right patient at the right time. Lau-Min et al (Lau-Min et al., 2022) have recently published their experiences of integrated genetic information in the EHR and conclude that it has "substantially streamlined the delivery of genomic medicine". However, it required multidisciplinary collaboration and EHR vendor development.

6.5 Study Quality and Limitations

In any research project, it is important to be critical of the sources, the quality, and the limitations of the chosen methods. Reliability, validity, and generalisability can together give an indicator of research quality (Tjora, 2017).

6.5.1 Reliability

Reliability refers to the consistency of the results and the extent to which the results can be reproduced if the research is repeated under the same conditions and can be assessed by checking the consistency of results over time. The interview guide enhanced the reliability of the study, and although recordings were not used, it ensured that all the interviews followed the same structure. Reliability would have been improved if two researchers had been involved in the data gathering methods and data analysis. However, the strong correlation between findings from the different data gathering methods strengthens the reliability of the study. The results are also consistent with what has been found in previous studies in the literature.

6.5.2 Validity

Validity refers to how accurately a method measures what is intended. This can include whether the choice of methodology is relevant for answering the research objectives, the sampling and data analysis is appropriate, and finally the results and conclusions are valid for the sample and context.

An explorative approach was chosen to answer the research questions, using qualitative methods within a UCRE framework. Appropriate sampling methods were used and a clearly defined representative choice of participant and an adequate amount of respondents were included in the study. Although there are relatively few users of genetic information in the EHR, different users were chosen to be included in the different iterations to ensure that results could be validated during the process. My personal knowledge of genetics was used to make certain that the interview guide had relevant questions and correct terminologies to ensure it was relevant for the research questions. The methodological choices were constrained by both time and practicalities but by using purposeful sampling and data-analysis combined with several iterations and triangulation of methods, a high degree of validity of the results was ensured in this study.

6.5.3 Generalisability

Most qualitative research studies study a specific issue in a certain population, of a focused locality in a particular context, so generalisability of the research findings is usually not an expected attribute (Leung, 2015). One approach for assessing generalisability for qualitative studies is to adopt same criteria for validity e.g. use of systematic sampling, triangulation and constant comparison of results, and correct documentation. This enables others to assess if the findings are valid for other situations.

The identified workflow for a genetic analysis at DMG, OUH consists of many of the same steps and challenges that many others have described in the literature. And although it would seem that DIPS is less ready for genetic information than other EHRs, identical challenges as reported in the literature have been recognised in this study.

6.5.4 Limitations

I have worked in the field of medical genetics for many years and, as a laboratory leader, have been involved in system innovation, improving workflows and understand the importance of involving the users in any developments and change. I have tried to distinguish between the information that came from data generation in this study and ideas that come from my own personal experience. This subjectivity, however, may have influenced my interpretations of the data. I know all the participants in this study, but we do not work in the same unit at the department, and although I have had no intentions on influencing the choice to participate in the study, there is a possibility that my role as a leader in the department has had an effect.

This study included a state-of-the-art literature review that excluded many articles that an extensive literature review might have included. All possible solutions that might exist for integration of genetic information may not have been revealed, but one can assume that most available aspects have been covered and discussed.

The quality of the results is also limited by the choice of participants, and it could have been appropriate to include a more varied group of end-users. One limitation of the approach in this study is the over representation of clinical and laboratory geneticists and fewer referring clinicians. The results from the UCRE confirm that the different professions have unique needs and requirements with regards to genetic information in the EHR and the final list of requirements might have been more extensive if a fuller representation of end-users in the study had been included. It would be interesting to involve additional laboratory providers of simpler genetic tests at OUH, e.g., department of Biochemistry or Dept. of pathology, who use different LIMSs then DMG, but also send genetic results to DIPS. Due to the time restrictions of the study this was unfortunately out of scope. Requirement elicitation is an iterative process, and it would be possible to achieve higher levels of details in the functional needs of the users if additional iterations of the prototype had been continued and a functional prototype developed.

7 Conclusion

This chapter will conclude the study by summarising the key research findings in relation to the research objectives, discuss the value and contribution of the study and present proposals for further work.

7.1 Research Objective and Key Findings

This study aimed to investigate the organisation of genetic information in the EHR and explore the potential the improved organisation can have on the use of genetic test results in clinical use. To answer the research objective and identify relevant requirements, I have used a user-centred requirement engineering approach. Through qualitative and exploratory methods such as observations, interviews, and paper prototyping, I have shown the importance of end-user involvement while identifying user requirements for optimal use of genetic information in the EHR. The results indicate that HIS developments should be driven by user needs identified in clinical settings and involve close collaboration with end users as opposed to technology driven innovation.

To realise precision medicine, it is necessary to define data requirements, develop robust IT infrastructures and integrate genetic information into the EHR. However, this study also reveals the importance of improving clinician knowledge regarding the potential of structured genetic information in the EHR, and how it would allow for a streamlined, efficient workflow and increased clinical use.

The results demonstrate that there is a huge potential for organised and improved integration of genetic information with other clinical patient information in the EHR that is currently in use at OUH. Currently all genetic information is organised in PDF documents or text blocks, difficult to locate and unavailable for any computable recommendations.

The developed prototype in this study represents a possible solution to a complex problem. The prototype included only the necessary complexity to meet the user requirements using existing standards and available technology. The prototype indicated that enhanced organisation and display of genetic information using standardised representations of genetic data, can result in highly efficient workflows and increased quality of genetic test results. Improved organisation and integration of genetic information in a genetic module provided an easily accessible overview of all genetic information, and perhaps most importantly, allowed for CDS.

This study has provided a list of identified user requirements from clinicians using genetic information at the point of care, all of which would lead to a more effective workflow than the present. A qualitative understanding of the complexities of integrated genetic information in the EHR is also presented. Together this can be used as a starting point for innovation towards bridging the genetics-HIS gap within precision medicine.

Having all the genetic information collected in one place in the EHR, together with the genetic decision support within the clinician's workflow, has a high impact on the clinical use and allows for many applications within precision medicine at the point of care.

7.2 Future work

In future work, investigating a functional prototype with a broader selection of end-users might prove to be useful. This study, though including referring clinicians, was very much from the perspective of medical geneticists. The involvement of non-genetic specialists from other laboratories and clinical departments would confirm the identified requirements, lead to new suggestions and provide information for further development.

Future research should investigate the possibilities of both the new LIMS and EHR systems that are planned for implementation within the next couple of years at DMG. In the implementation and configuration phase of these new systems, it is critical to have focus on user involvement and ensure that the necessary requirements are conveyed to the software developers. It would also be interesting to study how it could be possible to increase awareness of the potential for genetic CDS for healthcare professionals.

In Norway, the Directorate for Health has established a national strategy for personalised medicine where large-scale genetic analysis plays a major contribution (Directorate for e-health, 2016)Although the report concludes that the development of functionalities for the handling of genetic information in the EHR will be pivotal for successful implementation of personalised medicine, no national solutions have yet been published.

However, the Norwegian national competence network for precision medicine[16] (NorPreM) hope to facilitate and contribute to increased harmonisation and standardisation and are currently working on how to ensure that the genetic laboratories can receive the necessary clinical information to ensure high quality genetic diagnosis of hereditary diseases. This is important work which health providers, EHR vendors and the genetic community should prioritise.

[16] https://spesialisthelsetjenesten.no/nasjonalt-kompetansenettverk-for-persontilpasset-medisin

Bibliography

Alterovitz, G., Heale, B., Jones, J., Kreda, D., Lin, F., Liu, L., . . . Ramoni, R. (2020). FHIR Genomics: enabling standardization for precision medicine use cases. *NPJ genomic medicine, 5*(1), 1-4.

Alterovitz, G., Warner, J., Zhang, P., Chen, Y., Ullman-Cullere, M., Kreda, D., & Kohane, I. S. (2015). SMART on FHIR Genomics: facilitating standardized clinico-genomic apps. *Journal of the American Medical Informatics Association, 22*(6), 1173-1178. doi:10.1093/jamia/ocv045

Altman, M., Huang, T. T. K., & Breland, J. Y. (2018). Design Thinking in Health Care. *Preventing chronic disease, 15*, E117-E117. doi:10.5888/pcd15.180128

Aronson, S., Babb, L., Ames, D., Gibbs, R. A., Venner, E., Connelly, J. J., . . . e, M. N. E. W. G. (2018). Empowering genomic medicine by establishing critical sequencing result data flows: the eMERGE example. *Journal of the American Medical Informatics Association, 25*(10), 1375-1381. doi:10.1093/jamia/ocy051

Aronson, S., Mahanta, L., Ros, L. L., Clark, E., Babb, L., Oates, M., . . . Lebo, M. (2016). Information Technology Support for Clinical Genetic Testing within an Academic Medical Center. *Journal of Personalized Medicine, 6*(1). doi:10.3390/jpm6010004

Aronson, S. J., & Rehm, H. L. (2015). Building the foundation for genomics in precision medicine. *Nature, 526*(7573), 336-342. doi:10.1038/nature15816

AS, D. Historien om DIPS. Retrieved from https://www.dips.com/om-oss/historien-om-dips

Ayatollahi, H., Hosseini, S. F., & Hemmat, M. (2019). Integrating Genetic Data into Electronic Health Records: Medical Geneticists' Perspectives. *Healthcare informatics research, 25*(4), 289-296. doi:10.4258/hir.2019.25.4.289

Bernasconi, A., Canakoglu, A., Masseroli, M., & Ceri, S. (2021). The road towards data integration in human genomics: players, steps and interactions. *Briefings in Bioinformatics, 22*(1), 30-44. doi:10.1093/bib/bbaa080

Bick, D., Jones, M., Taylor, S. L., Taft, R. J., & Belmont, J. (2019). Case for genome sequencing in infants and children with rare, undiagnosed or genetic diseases. *J Med Genet, 56*(12), 783-791. doi:10.1136/jmedgenet-2019-106111

Blagec, K., Koopmann, R., Crommentuijn-van Rhenen, M., Holsappel, I., van der Wouden, C. H., Konta, L., . . . Samwald, M. (2018). Implementing pharmacogenomics decision support across seven European countries: The Ubiquitous Pharmacogenomics (U-PGx) project. *Journal of the American Medical Informatics Association : JAMIA, 25*(7), 893-898. doi:10.1093/jamia/ocy005

Campbell, W. S., Carter, A. B., Cushman-Vokoun, A. M., Greiner, T. C., Dash, R. C., Routbort, M., . . . Campbell, J. R. (2019). A Model Information Management Plan for Molecular Pathology Sequence Data Using Standards: From Sequencer to Electronic Health Record. *The Journal of Molecular Diagnostics, 21*(3), 408-417. doi:https://doi.org/10.1016/j.jmoldx.2018.12.002

Cancer Research UK. Family History and Inhertited cancer genes. Retrieved from https://www.cancerresearchuk.org/about-cancer/causes-of-cancer/inherited-cancer-genes-and-increased-cancer-risk/family-history-and-inherited-cancer-genes

Carter, A. B., Abruzzo, L. V., Hirschhorn, J. W., Jones, D., Jordan, D. C., Nassiri, M., . . . Roy, S. (2022). Electronic Health Records and Genomics: Perspectives from the Association for Molecular Pathology Electronic Health Record (EHR) Interoperability for Clinical Genomics Data Working Group. *The Journal of Molecular Diagnostics, 24*(1), 1-17. doi:https://doi.org/10.1016/j.jmoldx.2021.09.009

Coiera, E. (2009). Building a national health IT system from the middle out. *Journal of the American Medical Informatics Association, 16*(3), 271-273.
Coiera, E. (2015). *Guide to health informatics* (3rd ed.). Boca Raton, Fla: CRC Press.
Conway, J. R., Warner, J. L., Rubinstein, W. S., & Miller, R. S. (2019). Next-Generation Sequencing and the Clinical Oncology Workflow: Data Challenges, Proposed Solutions, and a Call to Action. *JCO Precis Oncol, 3*. doi:10.1200/po.19.00232
Crump, J. K., Del Fiol, G., Williams, M. S., & Freimuth, R. R. (2018). Prototype of a Standards-Based EHR and Genetic Test Reporting Tool Coupled with HL7-Compliant Infobuttons. *AMIA Joint Summits on Translational Science proceedings. AMIA Joint Summits on Translational Science, 2017*, 330-339.
Cutting, E. M., Overby, C. L., Banchero, M., Pollin, T., Kelemen, M., Shuldiner, A. R., & Beitelshees, A. L. (2015). Using Workflow Modeling to Identify Areas to Improve Genetic Test Processes in the University of Maryland Translational Pharmacogenomics Project. *AMIA ... Annual Symposium proceedings. AMIA Symposium, 2015*, 466-474. (Accession No. 26958179)
Das, A., & Svanæs, D. (2013). Human-centred methods in the design of an e-health solution for patients undergoing weight loss treatment. *International Journal of Medical Informatics, 82*(11), 1075-1091. doi:https://doi.org/10.1016/j.ijmedinf.2013.06.008
Directorate for e-health. (2016). *Nasjonal strategi for persontilpasset medisin i helsetjenesten*. Retrieved from https://www.helsedirektoratet.no/rapporter/strategi-for-persontilpasset-medisin-i-helsetjenesten
Directorate for e-health. (2019). *Felles språk i helse- og omsorgssektoren* (IE-1052). Retrieved from https://www.ehelse.no/publikasjoner/felles-sprak-i-helse-og-omsorgssektoren-malbilde-versjon-1.0
Directorate for e-health. (2021). *Plan for internasjonale standarder 2021-2024* (IE-1079). Retrieved from https://www.ehelse.no/publikasjoner/plan-for-internasjonale-standarder-2021-2024
Dolin, R. H., Boxwala, A., & Shalaby, J. (2018). A pharmacogenomics clinical decision support service based on FHIR and CDS hooks. *Methods of Information in Medicine, 57*(S 02), e115-e123.
Dolin, R. H., Gothi, S. R., Boxwala, A., Heale, B. S. E., Husami, A., Jones, J., . . . Alterovitz, G. (2021). vcf2fhir: a utility to convert VCF files into HL7 FHIR format for genomics-EHR integration. *BMC Bioinformatics, 22*(1), 104. doi:10.1186/s12859-021-04039-1
Ehn, M., Derneborg, M., Revenäs, Å., & Cicchetti, A. (2021). User-centered requirements engineering to manage the fuzzy front-end of open innovation in e-health: A study on support systems for seniors' physical activity. *International Journal of Medical Informatics, 154*, 104547. doi:https://doi.org/10.1016/j.ijmedinf.2021.104547
Eric, V., Yi, V., Murdock, D., Kalla, S. E., Wu, T. J., Sabo, A., . . . Gibbs, R. A. (2021). Neptune: an environment for the delivery of genomic medicine. *Genet Med, 23*(10), 1838-1846. doi:10.1038/s41436-021-01230-w
Furness, L. M. (2017). Bridging the gap: the need for genomic and clinical -omics data integration and standardization in overcoming the bottleneck of variant interpretation. *Expert Review of Precision Medicine and Drug Development, 2*(2), 79-89. doi:10.1080/23808993.2017.1322897
GA4GH. (2022). VA-SPEC. Retrieved from https://github.com/ga4gh/va-spec
Gansel, X., Mary, M., & van Belkum, A. (2019). Semantic data interoperability, digital medicine, and e-health in infectious disease management: a review. *Eur J Clin Microbiol Infect Dis, 38*(6), 1023-1034. doi:10.1007/s10096-019-03501-6
Garcia, S. J., Zayas-Cabán, T., & Freimuth, R. R. (2020). Sync for Genes: Making Clinical Genomics Available for Precision Medicine at the Point-of-Care. *Appl Clin Inform, 11*(02), 295-302.
García S, A., Iñiguez-Jarrín, C., Lopez, O. P., Gonzalez-Ibea, D., Pérez-Román, E., Borredà, C., . . . Talón, M. (2022, 2022//). *CitrusGenome: Applying User Centered*

Design for Evaluating the Usability of Genomic User Interfaces. Paper presented at the Evaluation of Novel Approaches to Software Engineering, Cham.
Garde, S., Knaup, P., Hovenga, E. J. S., & Heard, S. (2007). Towards Semantic Interoperability for Electronic Health Records. *Methods Inf Med, 46*(03), 332-343.
Grant, M. J., & Booth, A. (2009). A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal, 26*(2), 91-108. doi:https://doi.org/10.1111/j.1471-1842.2009.00848.x
Greenes, R. A. (2014). Chapter 1 - Definition, Scope, and Challenges. In R. A. Greenes (Ed.), *Clinical Decision Support (Second Edition)* (pp. 3-47). Oxford: Academic Press.
Greenhalgh, T., Stones, R., & Swinglehurst, D. (2014). Choose and Book: a sociological analysis of 'resistance' to an expert system. *Soc Sci Med, 104*, 210-219. doi:10.1016/j.socscimed.2013.12.014
Guarrera, T. K., McGeorge, N. M., Ancker, J. S., Hegde, S., Zhou, Y., Lin, L., . . . Bisantz, A. M. (2014). Characterising the effect of interoperability on healthcare work: a novel framework. *Theoretical Issues in Ergonomics Science, 15*(6), 578-594. doi:10.1080/1463922X.2013.838318
Heart, T., Ben-Assuli, O., & Shabtai, I. (2017). A review of PHR, EMR and EHR integration: A more personalized healthcare and public health policy. *Health Policy and Technology, 6*(1), 20-25. doi:https://doi.org/10.1016/j.hlpt.2016.08.002
Heeks, R. (2006). Health information systems: Failure, success and improvisation. *International Journal of Medical Informatics, 75*(2), 125-137. doi:https://doi.org/10.1016/j.ijmedinf.2005.07.024
Hernandez, S., Fairchild, K., Pemberton, M., Dahmer, J., Zhang, W., Palchuk, M. B., & Topaloglu, U. (2022). Applying FHIR Genomics for Research - From Sequencing to Database. *AMIA Annu Symp Proc, 2022*, 236-243.
Herr, T. M., Bielinski, S. J., Bottinger, E., Brautbar, A., Brilliant, M., Chute, C. G., . . . Starren, J. (2015). Practical considerations in genomic decision support: The eMERGE experience. *J Pathol Inform, 6*, 50. doi:10.4103/2153-3539.165999
Herr, T. M., Bielinski, S. J., Bottinger, E., Brautbar, A., Brilliant, M., Chute, C. G., . . . Starren, J. (2015). A conceptual model for translating omic data into clinical action. *Journal of pathology informatics, 6*, 46-46. doi:10.4103/2153-3539.163985
HIMSS. (2022). What is interoperability? Retrieved from https://www.himss.org/resources/interoperability-healthcare
Hooker, G. W. (2021). Building an infrastructure to enable delivery of genomic medicine. *American Journal of Medical Genetics Part C: Seminars in Medical Genetics, 187*(1), 95-99. doi:https://doi.org/10.1002/ajmg.c.31881
IT6103. (2019). In *IKT organisasjon og samarbeide i helsetjenesten, semesteroppgaven.* NTNU: Own work.
IT6122. (2020). In *CLINICAL DECISION SUPPORT: IMPLEMENTATION AND DISSEMINATION*. Own work.
Kannan, V., Basit, M. A., Bajaj, P., Carrington, A. R., Donahue, I. B., Flahaven, E. L., . . . Toomay, S. M. (2019). User stories as lightweight requirements for agile clinical decision support development. *Journal of the American Medical Informatics Association, 26*(11), 1344-1354. doi:10.1093/jamia/ocz123
Kawamoto, K., & Greenes, R. A. (2014). Chapter 21 - The Role of Standards: What We Can Expect and When. In R. A. Greenes (Ed.), *Clinical Decision Support (Second Edition)* (pp. 599-615). Oxford: Academic Press.
Khalifa, A., Mason, C. C., Garvin, J. H., Williams, M. S., Del Fiol, G., Jackson, B. R., . . . Huff, S. M. (2021). Interoperable genetic lab test reports: mapping key data elements to HL7 FHIR specifications and professional reporting guidelines. *Journal of the American Medical Informatics Association, 28*(12), 2617-2625. doi:10.1093/jamia/ocab201
Khalifa, A., Mason, C. C., Garvin, J. H., Williams, M. S., Del Fiol, G., Jackson, B. R., . . . Huff, S. M. (2021a). A qualitative investigation of biomedical informatics interoperability standards for genetic test reporting: benefits, challenges, and

motivations from the testing laboratory's perspective. *Genet Med, 23*(11), 2178-2185. doi:10.1038/s41436-021-01301-y
Khalifa, A., Mason, C. C., Garvin, J. H., Williams, M. S., Del Fiol, G., Jackson, B. R., . . . Huff, S. M. (2021b). A qualitative study of prevalent laboratory information systems and data communication patterns for genetic test reporting. *Genet Med, 23*(11), 2171-2177. doi:10.1038/s41436-021-01251-5
Kushniruk, A., Monkman, H., Borycki, E., & Kannry, J. (2015). User-Centered Design and Evaluation of Clinical Information Systems: A Usability Engineering Perspective. In V. L. Patel, T. G. Kannampallil, & D. R. Kaufman (Eds.), *Cognitive Informatics for Biomedicine: Human Computer Interaction in Healthcare* (pp. 141-161). Cham: Springer International Publishing.
Kushniruk, A., & Nøhr, C. (2016). Participatory Design, User Involvement and Health IT Evaluation. *Stud Health Technol Inform, 222*, 139-151.
Lau-Min, K. S., Asher, S. B., Chen, J., Domchek, S. M., Feldman, M., Joffe, S., . . . Nathanson, K. L. (2021). Real-world integration of genomic data into the electronic health record: the PennChart Genomics Initiative. *Genetics in Medicine, 23*(4), 603-605. doi:10.1038/s41436-020-01056-y
Lau-Min, K. S., McKenna, D., Asher, S. B., Bardakjian, T., Wollack, C., Bleznuck, J., . . . Nathanson, K. L. (2022). Impact of integrating genomic data into the electronic health record on genetics care delivery. *Genetics in Medicine*. doi:https://doi.org/10.1016/j.gim.2022.08.009
Lehne, M., Sass, J., Essenwanger, A., Schepers, J., & Thun, S. (2019). Why digital medicine depends on interoperability. *npj Digital Medicine, 2*(1), 79. doi:10.1038/s41746-019-0158-1
Liyanage, H., Krause, P., & De Lusignan, S. (2015). Using ontologies to improve semantic interoperability in health data. *J Innov Health Inform, 22*(2), 309-315. doi:10.14236/jhi.v22i2.159
Mascia, C., Frexia, F., Uva, P., Zanetti, G., Pireddu, L., Giacomelli, G., . . . Tomczak, A. (2020). *The openEHR Genomics Project* (Vol. 270).
Masys, D. R., Jarvik, G. P., Abernethy, N. F., Anderson, N. R., Papanicolaou, G. J., Paltoo, D. N., . . . Levy, H. P. (2012). Technical desiderata for the integration of genomic data into Electronic Health Records. *J Biomed Inform, 45*(3), 419-422. doi:10.1016/j.jbi.2011.12.005
McCudden, C. R., Henderson, M. P. A., & Jackson, B. R. (2020). Chapter 18 - Laboratory information management. In W. Clarke & M. A. Marzinke (Eds.), *Contemporary Practice in Clinical Chemistry (Fourth Edition)* (pp. 301-321): Academic Press.
Moreno-Conde, A., Moner, D., Cruz, W. D. d., Santos, M. R., Maldonado, J. A., Robles, M., & Kalra, D. (2015). Clinical information modeling processes for semantic interoperability of electronic health records: systematic review and inductive analysis. *Journal of the American medical informatics association, 22*(4), 925-934.
Murugan, M., Babb, L. J., Overby Taylor, C., Rasmussen, L. V., Freimuth, R. R., Venner, E., . . . Gibbs, R. A. (2021). Genomic considerations for FHIR®; eMERGE implementation lessons. *Journal of Biomedical Informatics, 118*, 103795. doi:https://doi.org/10.1016/j.jbi.2021.103795
NHS. (2015). *Interoperability Handbook*. Retrieved from chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.england.nhs.uk/digitaltechnology/wp-content/uploads/sites/31/2015/09/interoperabilty-handbk.pdf
Oates, B. J. (2005). *Researching information systems and computing*: Sage.
Ohno-Machado, L., Kim, J., Gabriel, R. A., Kuo, G. M., & Hogarth, M. A. (2018). Genomics and electronic health record systems. *Human Molecular Genetics, 27*(R1), R48-R55. doi:10.1093/hmg/ddy104
OpenEHR. (2022). What is OpenEHR? Retrieved from https://www.openehr.org/about/what_is_openehr
Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., . . . McKenzie, J. E. (2021). PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ, 372*, n160. doi:10.1136/bmj.n160

Paxton, A. (2022). A gap in need of a fix: EHRs and genomics. In: Cap Today.
Preece, J., Rogers, Y., & Sharp, H. (2019). *Interaction Design: Beyond Human-Computer Interaction*: Wiley-Blackwell.
Pronovost, P., M. M. E. Johns, S. Palmer, R. C. Bono, D. B. Fridsma, A. Gettinger, J. Goldman, W. Johnson, M. Karney, C. Samitt, R. D. Sriram, A. Zenooz, and Y. C. Wang,. (2018). *Procuring Interoperability: Achieving High-Quality, Connected, and Person-Centered Care*. Retrieved from Washington, DC:
Qin, W., Lu, X., Shu, Q., Duan, H., & Li, H. (2022). Building an information system to facilitate pharmacogenomics clinical translation with clinical decision support. *Pharmacogenomics, 23*(1), 35-48. doi:10.2217/pgs-2021-0110
Rehm, H. L., Page, A. J. H., Smith, L., Adams, J. B., Alterovitz, G., Babb, L. J., . . . Birney, E. (2021). GA4GH: International policies and standards for data sharing across genomic research and healthcare. *Cell Genomics, 1*(2). doi:10.1016/j.xgen.2021.100029
Roncato, R., Dal Cin, L., Mezzalira, S., Comello, F., De Mattia, E., Bignucolo, A., . . . Cecchin, E. (2019). FARMAPRICE: A Pharmacogenetic Clinical Decision Support System for Precise and Cost-Effective Therapy. *Genes, 10*(4). doi:10.3390/genes10040276
Rowley, J. (2007). The wisdom hierarchy: representations of the DIKW hierarchy. *Journal of Information Science, 33*(2), 163-180. doi:10.1177/0165551506070706
Rueda, S., Panach, J. I., & Distante, D. (2020). Requirements elicitation methods based on interviews in comparison: A family of experiments. *Information and Software Technology, 126*, 106361. doi:https://doi.org/10.1016/j.infsof.2020.106361
Ryu, B., Shin, S. Y., Baek, R. M., Kim, J. W., Heo, E., Kang, I., . . . Yoo, S. (2020). Clinical Genomic Sequencing Reports in Electronic Health Record Systems Based on International Standards: Implementation Study. *J Med Internet Res, 22*(8), e15040. doi:10.2196/15040
Sachdeva, S., & Bhalla, S. (2012). Semantic Interoperability in Standardized Electronic Health Record Databases. *J. Data and Information Quality, 3*(1). doi:10.1145/2166788.2166789
Sen, A., Al Kawam, A., & Datta, A. (2019). Emergence of DSS efforts in genomics: Past contributions and challenges. *Decision Support Systems, 116*, 77-90. doi:https://doi.org/10.1016/j.dss.2018.10.011
Sharp, H. (2019). *Interaction design : beyond human-computer interaction*. In Y. Rogers & J. Preece (Eds.), (5th ed.).
Shirts, B. H., Salama, J. S., Aronson, S. J., Chung, W. K., Gray, S. W., Hindorff, L. A., . . . Tarczy-Hornoch, P. Z. (2015). CSER and eMERGE: current and potential state of the display of genetic information in the electronic health record. *Journal of the American Medical Informatics Association, 22*(6), 1231-1242.
Sissung, T. M., McKeeby, J. W., Patel, J., Lertora, J. J., Kumar, P., Flegel, W. A., . . . Goldspiel, B. R. (2017). Pharmacogenomics Implementation at the National Institutes of Health Clinical Center. *The Journal of Clinical Pharmacology, 57*(S10), S67-S77. doi:https://doi.org/10.1002/jcph.993
SNOMED International. Retrieved from http://browser.ihtsdotools.org/
SNOMED International. Introduction. Retrieved from https://www.snomed.org/snomed-ct/five-step-briefing
Sommerville, I. (2016). *Software engineering* (10th , global ed.). Boston Mass.: Pearson.
Standardization, I. O. f. (2019). Health Informatics - Electronic Health Records. In.
Sutton, R. T., Pincock, D., Baumgart, D. C., Sadowski, D. C., Fedorak, R. N., & Kroeker, K. I. (2020). An overview of clinical decision support systems: benefits, risks, and strategies for success. *npj Digital Medicine, 3*(1), 17. doi:10.1038/s41746-020-0221-y
Taylor, C. O., Rasmussen, L. V., Rasmussen-Torvik, L. J., Prows, C. A., Dorr, D. A., Samal, L., & Aronson, S. (2021). Facilitating Genetics Aware Clinical Decision Support: Putting the eMERGE Infrastructure into Practice. *ACI open, 05*(02), e54-e58.

Teixeira, L., Ferreira, C., & Santos, B. S. (2012). User-centered requirements engineering in health information systems: A study in the hemophilia field. *Computer methods and programs in biomedicine, 106*(3), 160-174.

Tjora, A. H. (2017). *Kvalitative forskningsmetoder i praksis* (3. utg. ed.). Oslo: Gyldendal akademisk.

Ullman-Cullere, M., Clark, E., & Aronson, S. (2018). Implications of Genomics for Clinical Informatics. In L. Liu & M. T. Özsu (Eds.), *Encyclopedia of Database Systems* (pp. 1814-1818). New York, NY: Springer New York.

Ulriksen, G.-H., Pedersen, R., & Ellingsen, G. (2017). Infrastructuring in healthcare through the openEHR architecture. *Computer Supported Cooperative Work (CSCW), 26*(1), 33-69.

Virzi, R. A. (1992). Refining the test phase of usability evaluation: How many subjects is enough? *Human factors, 34*(4), 457-468.

Walton, N. A., Johnson, D. K., Person, T. N., Reynolds, J. C., & Williams, M. S. (2020). Pilot implementation of clinical genomic data into the native electronic health record: challenges of scalability. *ACI Open, 4*(02), e162-e166.

Warner, J. L., Jain, S. K., & Levy, M. A. (2016). Integrating cancer genomic data into electronic health records. *Genome Med, 8*(1), 113. doi:10.1186/s13073-016-0371-3

Watkins, M., & Eilbeck, K. (2020). FHIR Lab Reports: using SMART on FHIR and CDS Hooks to increase the clinical utility of pharmacogenomic laboratory test results. *AMIA Joint Summits on Translational Science proceedings. AMIA Joint Summits on Translational Science, 2020*, 683-692.

Welch, B. M., Eilbeck, K., Fiol, G. D., Meyer, L. J., & Kawamoto, K. (2014). Technical desiderata for the integration of genomic data with clinical decision support. *Journal of Biomedical Informatics, 51*, 3-7. doi:https://doi.org/10.1016/j.jbi.2014.05.014

Wiesner, G. L., Kulchak Rahm, A., Appelbaum, P., Aufox, S., Bland, S. T., Blout, C. L., . . . Leppig, K. A. (2020). Returning Results in the Genomic Era: Initial Experiences of the eMERGE Network. *J Pers Med, 10*(2). doi:10.3390/jpm10020030

Williams, M. S., Taylor, C. O., Walton, N. A., Goehringer, S. R., Aronson, S., Freimuth, R. R., . . . Del Fiol, G. (2019). Genomic Information for Clinicians in the Electronic Health Record: Lessons Learned From the Clinical Genome Resource Project and the Electronic Medical Records and Genomics Network. *Frontiers in genetics, 10*, 1059-1059. doi:10.3389/fgene.2019.01059

Appendices

Appendix 1: Letter of invitation

Letter of invitation

Master thesis – Organisation of genetic data in the EHR

Dear participant, you are invited to take part in a research study entitled " User requirements for integrated genetic information in the electronic health record". This research study is being undertaken in part fulfilment of an MSc in Health Informatics at NTNU, Trondheim.

Background and study aims:

- To facilitate use of genetic information at point of care, it should be available in standardised, computable formats.
- EHR (Electronic Health Record) is an essential component of a clinician's workflow, so utilising the EHR to present information to support the use of genomic medicine in clinical care to improve outcomes represents a tremendous opportunity.
- Effectively linking and integrating clinical information with personal genetic information in the EHR, will enable the shift to personalised medicine.
- Inconsistencies in how genetic information is shared and delivered can make it difficult for clinicians to interpret and use.

The purpose of this study is to define the requirements for the organisation and display of genetic information in the EHR and to investigate the possibilities for representation of this information to ensure genetic data is readily available to the users and for clinical decision support. This research requires the gathering, evaluation and analysis of qualitative data relating to integration of genetic information in DIPS

You have been chosen to participate in this study in your role as a domain expert. Your participation in this study is voluntary and you are free to withdraw at any time without providing a reason. If you choose to take part in this study, I will contact you to arrange an interview and/or evaluation of a simple prototype. The time taken to complete the interview is anticipated to be approximately 45 minutes and will not be recorded, but notes will be taken underway. All information collected during the research will be kept strictly confidential. In my thesis I may use direction quotations (when they are contextually appropriate) but you will remain anonymous. The data collected will be aggregated for the purpose of the research and no participants will be individually named in the thesis.

Appendix 2: Interview guide

Context:

1. Name of department:
2. Position/Title:
3. What type of genetic information in DIPS do you use and why?

Presentation of workflow and use examples

Comments

Test ordering

1. Are genetic test orders handled in a way that is different from other laboratory tests?
2. Is it possible for electronic ordering of genetic tests?
3. Is it possible to identify which genetic tests have been ordered?
4. How are phenotype/clinical features/family history data included in the test order? Entered by whom?
5. Do you experience any challenges with the current solution wrt test ordering?
6. What functionalities or data elements would be most useful or are missing with regards to test ordering? E.g. triggers/warnings that a genetic test has been earlier ordered or that the results of a genetic test are registered in DIPS?, CDS for test ordering or interpretation? Post-test or pre-test.
7. How and where is patient consent documented?
8. Are there any functionalities which could lead to increased use of genetic testing?

Reporting of genetic results to Dips

1. What are the different sources that enter genetic information into DIPS (eg: local hospital laboratory, reference laboratory, independent genetic testing laboratories, physician notes)?
2. What factors determine the location in DIPS where genetic information is displayed? Eg source laboratory, department of the requisitioner, the way info will be used).
3. Are there some instances where genetic info is displayed in multiple places?
4. Do you have systems other than DIPS to store genetic info where clinicians can view genetic results?
5. Do you experience any challenges with the current solution wrt genetic results?
6. What functionalities or data elements would be most useful or are missing with regards to display of genetic results?
7. Are there any functionalities which could lead to increased visibility of genetic results?

Genetic category questions – see below for categories

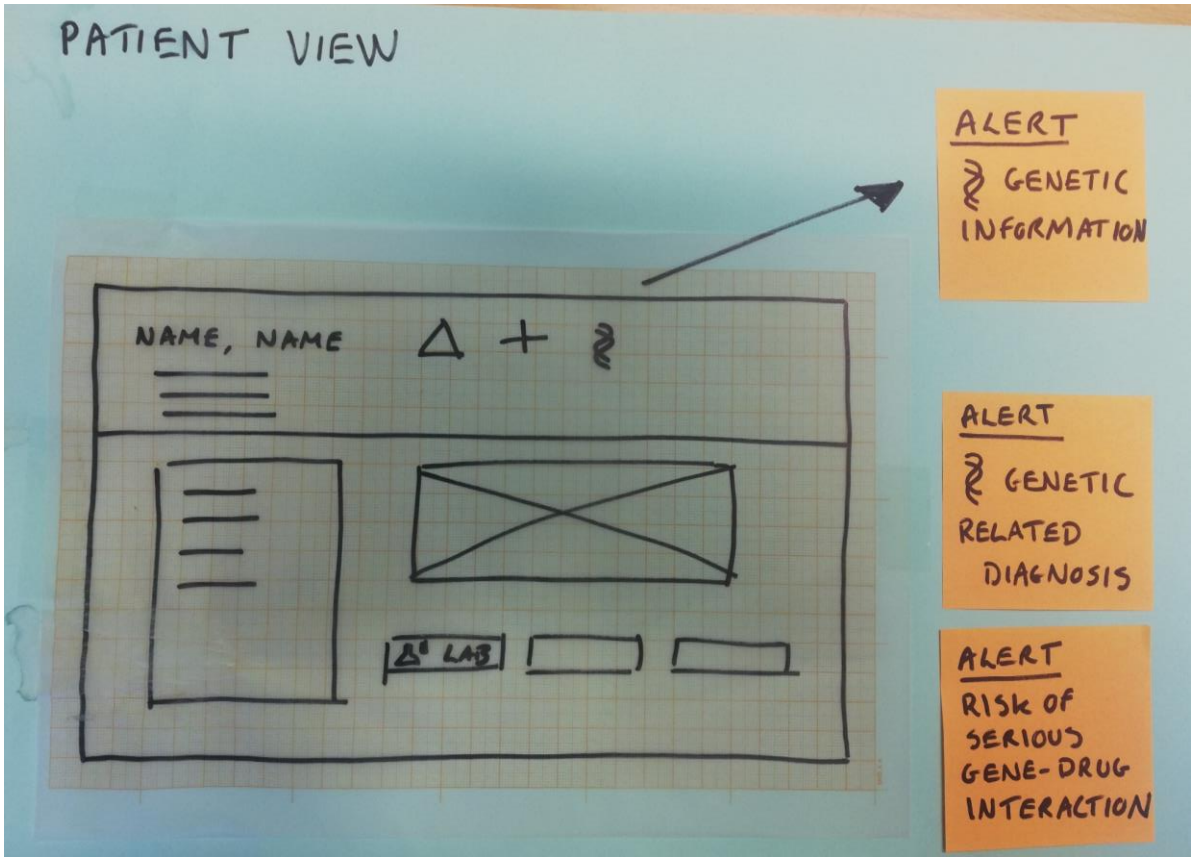
1. Are any genetic test results displayed otherwise than in text blocks or PDF documents?
2. Are any genetic test results annotated in DIPS eg. problem list, allergy list, diagnosis section?
3. Is there a mechanism for high risk medical actionable info to trigger an alert in the EHR?
4. What genetic terminology standards or interoperability standards are currently in use?
5. Finally, is there are anything else that would you like to change from today's DIPS wrt genetic information?

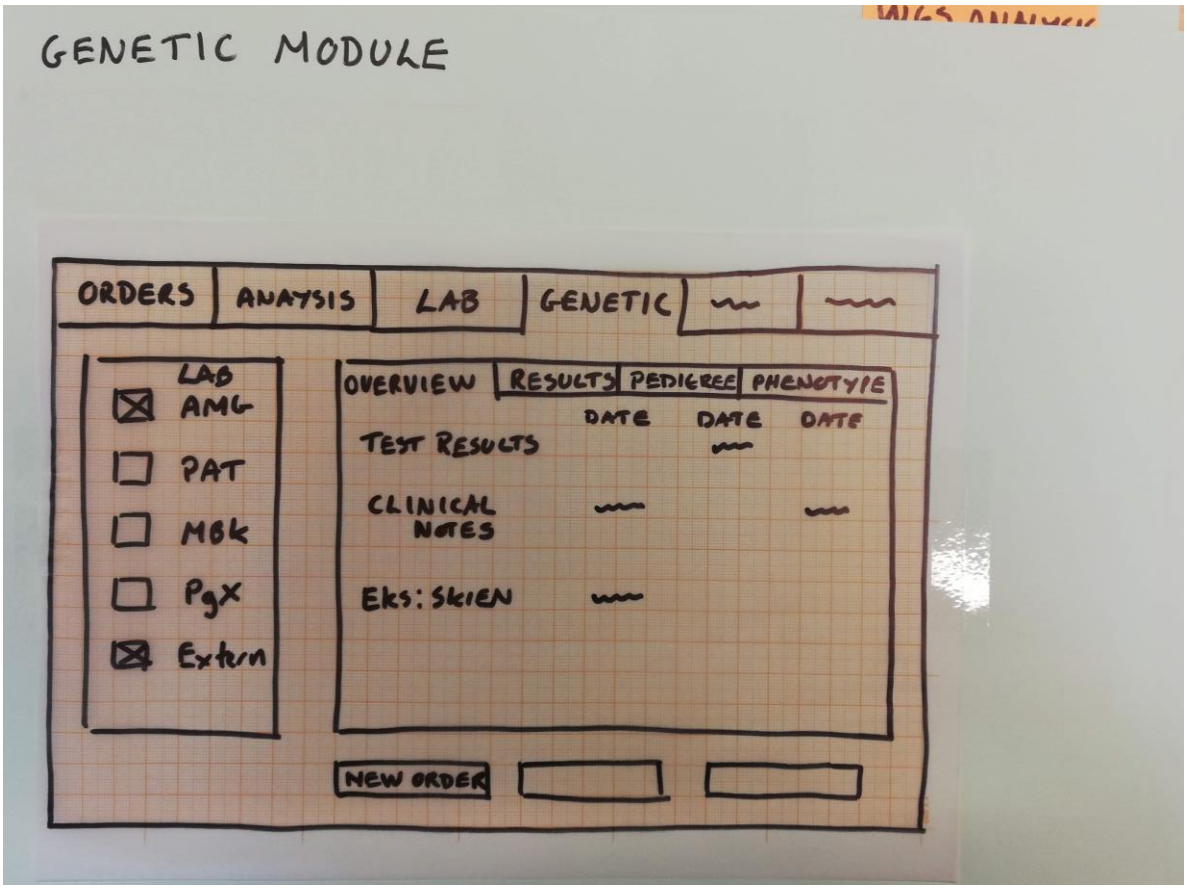
Genetic category	**Displayed in DIPS?** Where?	Where would the ideal location/ways to display this info?
Diagnostic/disease defining		
Risk actionable		
Carrier recessive		
VUS		
Uninterpreted variants		
Pharmacogenetics		
Somatic/tumor genetics		
Newborn sreening		

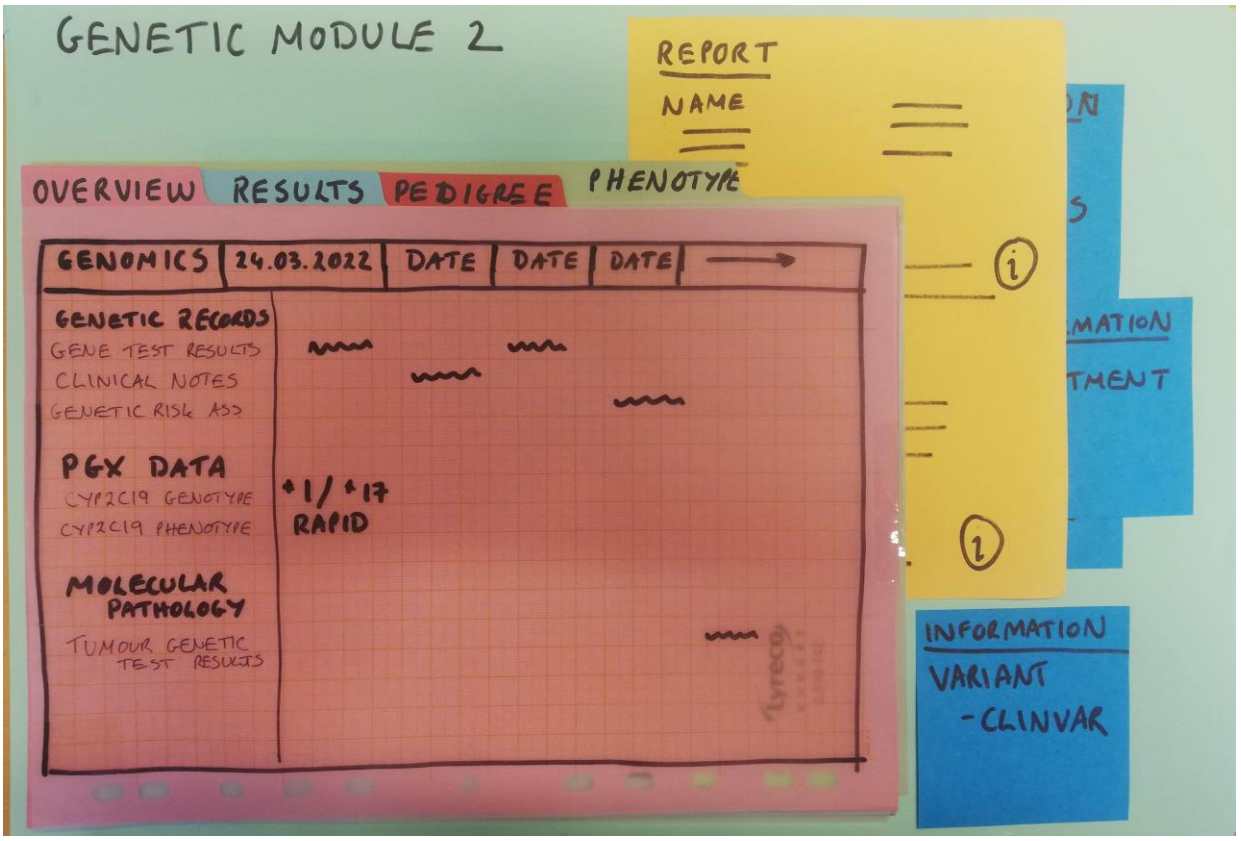
Use examples (Shirts et al, 2015)

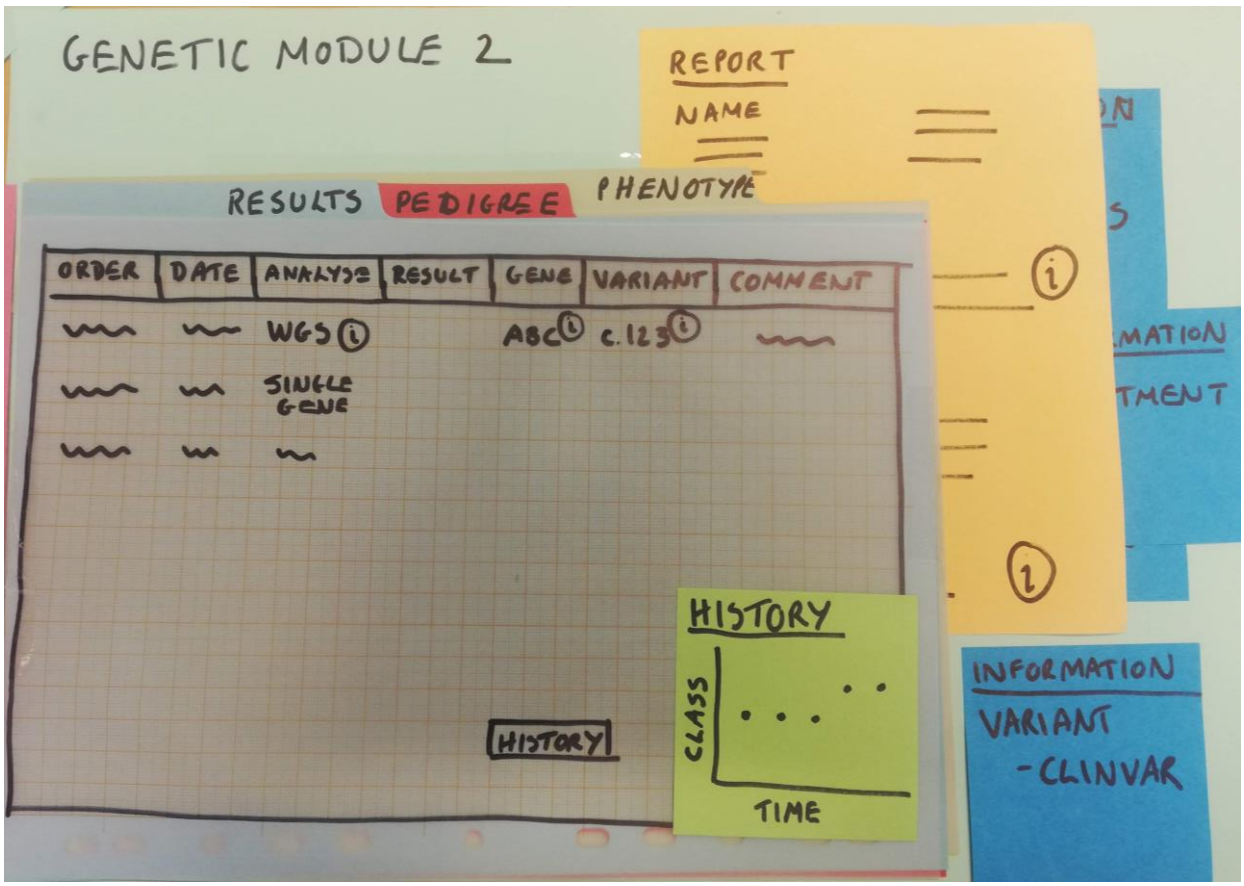
- A patient presents with hypertrophic cardiomyopathy. A genetic cardiac panel is ordered, and a pathogenic cardiomyopathy variant is identified. The overall test result, laboratory report, structured variant data, and associated interpretations are placed in the DIPS's genetic summary screen. An alert is sent to the ordering clinician indicating the result is available. The ordering geneticist reviews the report and determines that it is appropriate to add "Genetic Predisposition to Cardiomyopathy" to the patient's problem list.
- An individual undergoes exome testing and is found to be heterozygous for one CFTR mutation. The laboratory result is noted in the EHR and flagged as carrier status results.
- A cancer risk panel of 40 genes is ordered for a male patient with a personal history of colon cancer. The patient is found to have a BRCA2 truncating mutation. The result is reported in the EHR and to the clinician as an incidental finding. There is a mechanism to alert the clinician about additional, unexpected follow up that may be necessary such as genetic counselling about increased cancer risk and appropriate evaluation of family members.
- A 43-year-old female patient with a personal and family history of breast cancer undergoes sequencing analysis of BRCA1 and BRCA2. A missense VUS is reported in BRCA1 and reported as a VUS. Nine months later, a revised laboratory report reclassifies the variant as pathogenic based on additional evidence. The EHR is updated to now follow the recommendations found in Diagnostic and Actionable categories.

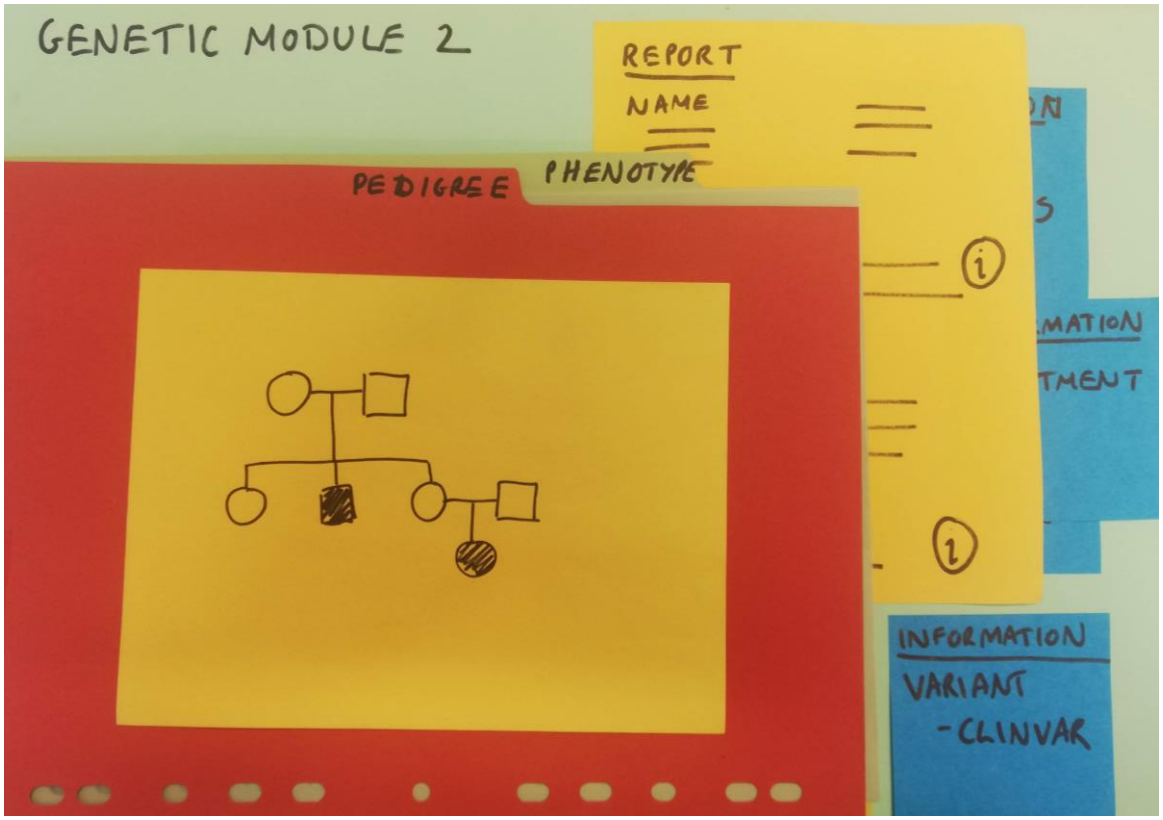
Appendix 3: Paper Prototype

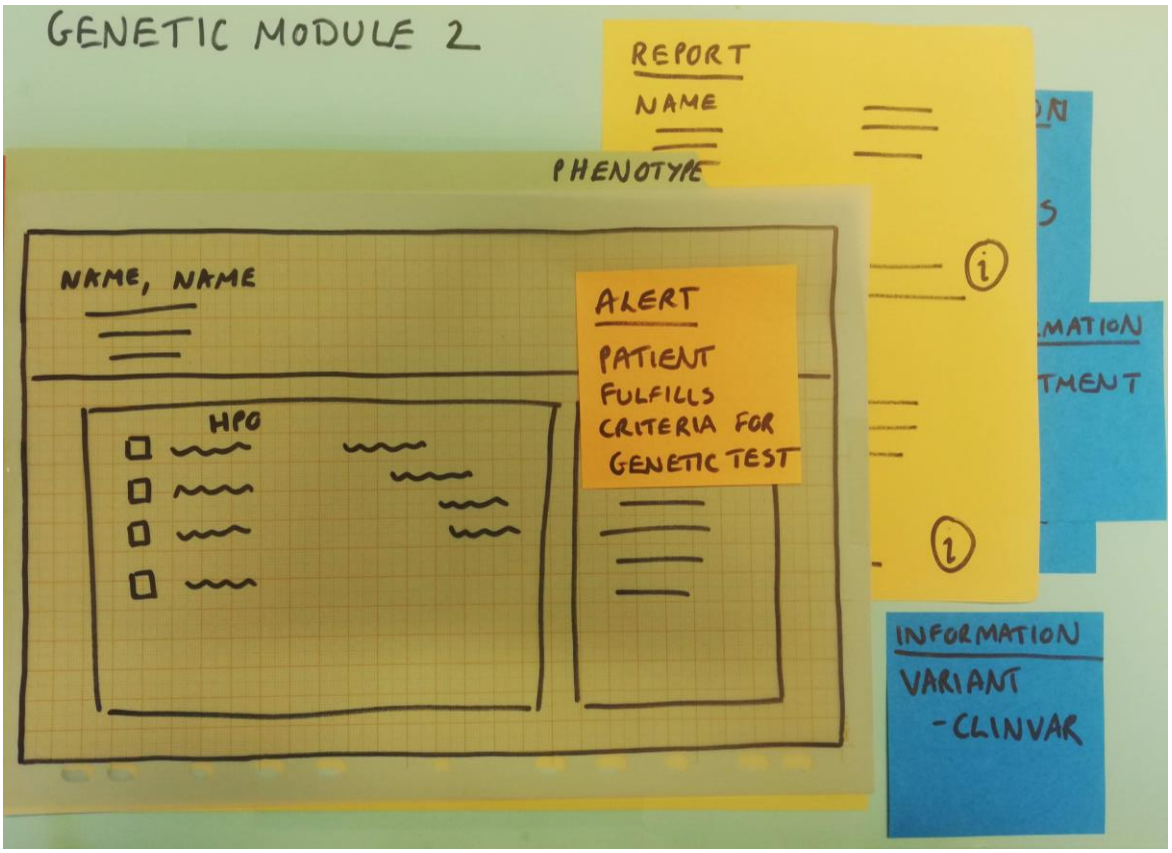


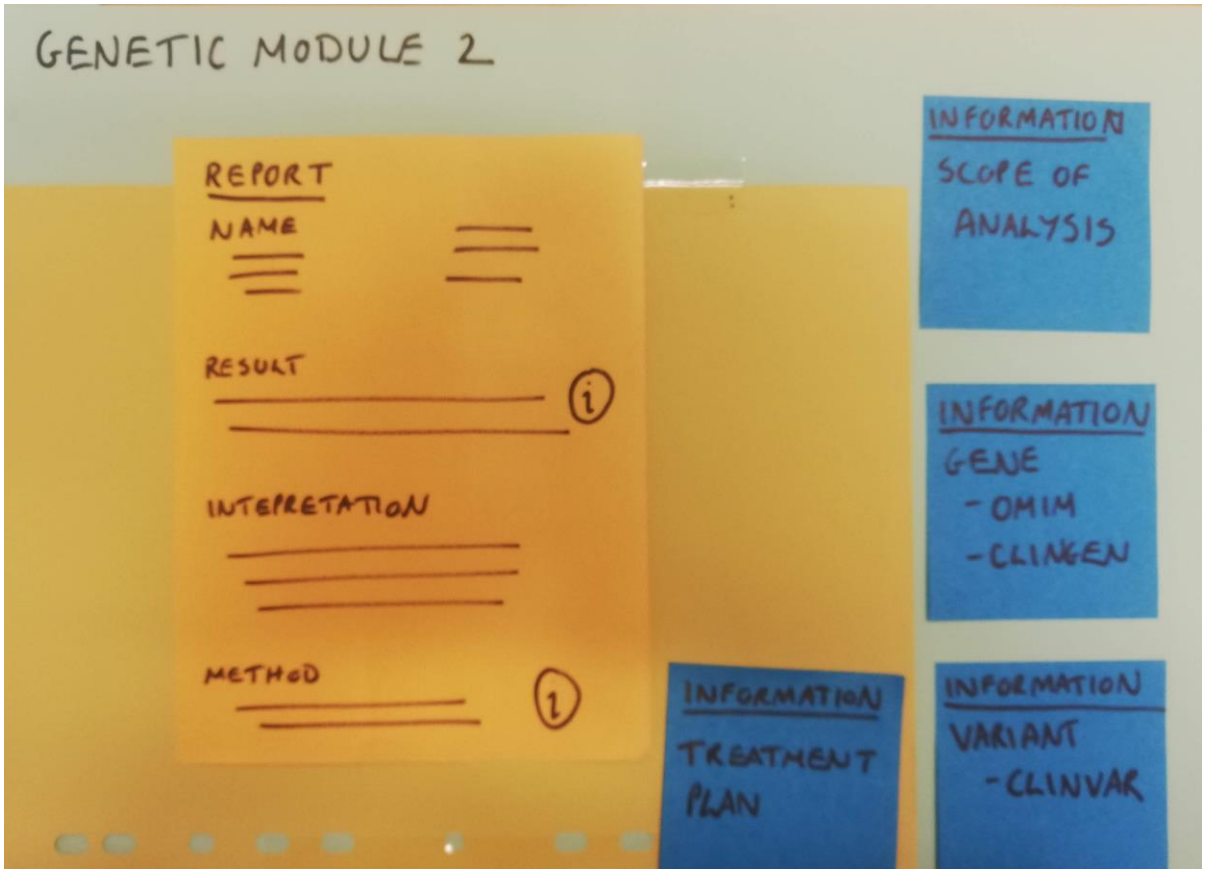


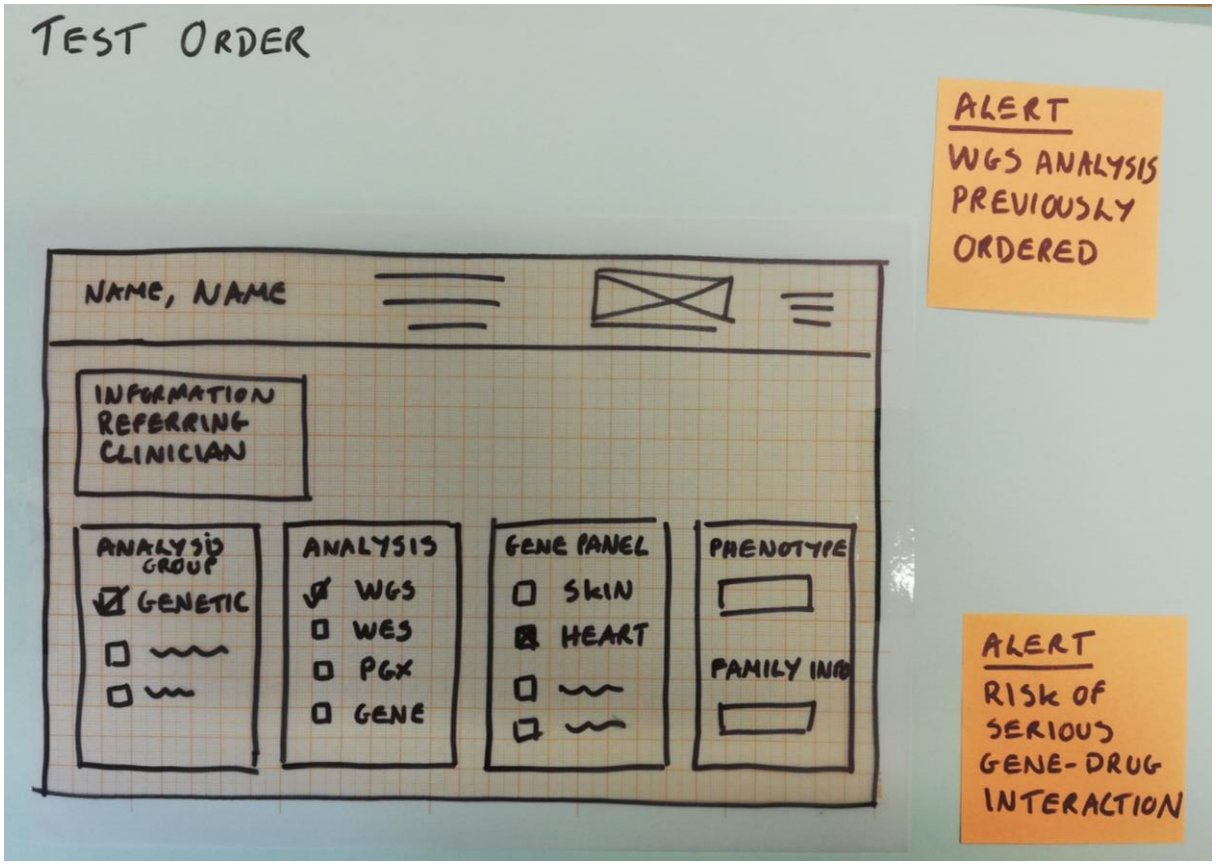

GENETIC MODULE 2
REPORT
NAME
OVERVIEW
RESULTS
PEDIGREE
PHENOTYPE
GENOMICS
24.03.2022
DATE
DATE
DATE
GENETIC RECORDS
GENE TEST RESULTS
CLINICAL NOTES
GENETIC RISK ASS
PGX DATA
CYP2C19 GENOTYPE
CYP2C19 PHENOTYPE
*1/*17
RAPID
MOLECULAR PATHOLOGY
TUMOUR GENETIC TEST RESULTS
INFORMATION
VARIANT
-CLINVAR

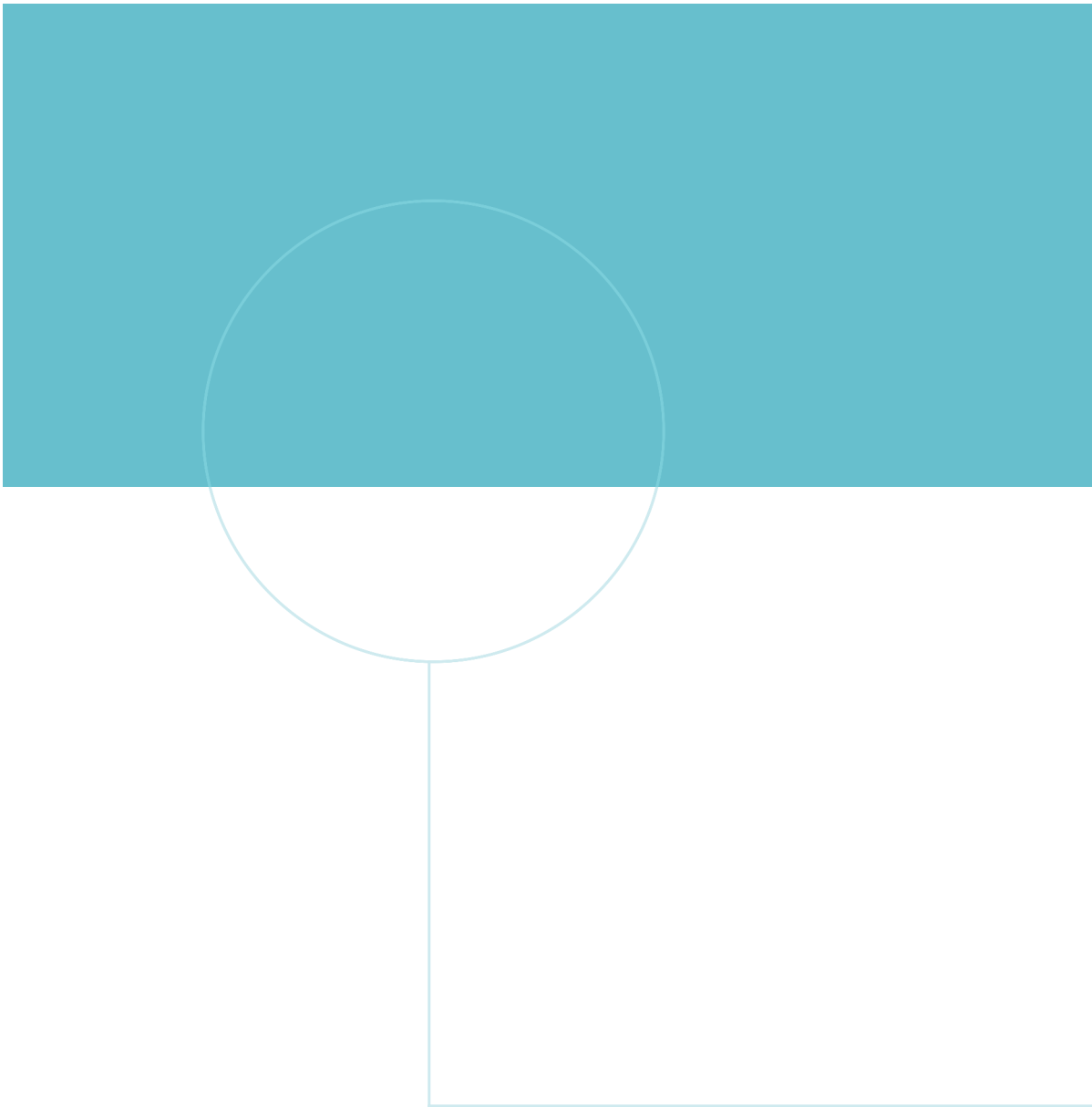

GENETIC MODULE 2
REPORT
NAME
RESULTS
PEDIGREE
PHENOTYPE
ORDER
DATE
ANALYSIS
RESULT
GENE
VARIANT
COMMENT
WGS
ABC
c.123
SINGLE GENE
HISTORY
HISTORY
CLASS
TIME
INFORMATION
VARIANT
-CLINVAR


GENETIC MODULE 2
REPORT
NAME
PEDIGREE
PHENOTYPE
INFORMATION
VARIANT
-CLINVAR


GENETIC MODULE 2
REPORT
NAME
PHENOTYPE
NAME, NAME
HPO
ALERT
PATIENT FULFILLS CRITERIA FOR GENETIC TEST
INFORMATION
VARIANT
-CLINVAR


GENETIC MODULE 2
REPORT
NAME
RESULT
INTERETATION
METHOD
INFORMATION
SCOPE OF ANALYSIS
INFORMATION
GENE
- OMIM
- CLINGEN
INFORMATION
TREATMENT PLAN
INFORMATION
VARIANT
- CLINVAR


TEST ORDER
ALERT
WGS ANALYSIS PREVIOUSLY ORDERED
NAME, NAME
INFORMATION REFERRING CLINICIAN
ANALYSIS GROUP
GENETIC
ANALYSIS
WGS
WES
PGX
GENE
GENE PANEL
SKIN
HEART
PHENOTYPE
FAMILY INFO
ALERT
RISK OF SERIOUS GENE-DRUG INTERACTION



NTNU

Norwegian University of
Science and Technology