User-Centered Design of a Collaborative Genetic Variant Interpretation Tool

Wayne Haowei Liang, MD

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Science

University of Washington

2017

Committee:

Peter Tarczy-Hornoch, Chair

Annie T. Chen

David R. Crosslin

Leslie S. Kean

Program Authorized to Offer Degree:

Biomedical and Health Informatics

University of Washington

Abstract

User-Centered Design of a Collaborative Genetic Variant Interpretation Tool

Wayne Haowei Liang

Chair of the Supervisory Committee:
Peter Tarczy-Hornoch, MD
Professor & Chair, Department of Biomedical Informatics and Medical Education
Professor, Department of Pediatrics (Neonatology)
Adjunct Professor, Department of Computer Science and Engineering

Precision genomic medicine relies upon accurate variant knowledge. However, laboratories continue to arrive at discordant interpretations for the same genomic test. Gaps, inconsistencies, and siloing of variant knowledge may contribute to inter-rater discordance in variant interpretation. Our overall goal is to develop a novel, openly available computerized tool supporting role-based collaboration, knowledge sharing, and consensus-making in variant interpretation. In *Aim 1*, we use literature review and informal expert input to characterize a typical variant interpretation workflow, propose a collaborative workflow, and develop an initial design for a computerized tool supporting collaborative variant interpretation. In *Aim 2*, we use user-centered design methodology to further characterize the typical workflow, define project requirements and user needs, and finalize the design of a tool supporting collaborative variant interpretation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

This thesis would not have been possible without the love, encouragement, and support of my wife Bindu; my children Isaiah, Ariella, and Eliana; my parents Chao Kuang and Showan; and my siblings Jay, Grace, and Ben. Thank you for believing in me.

I am deeply grateful to everyone who contributed to my education, research and professional development during my post-doctoral fellowship, including: Peter Tarczy-Hornoch; David Crosslin; Annie Chen; Gail Jarvik; George Demiris; Bas de Veer; Diane Korngiebel; Brian Shirts; Michael Dorschner; Peter Byers; Julie Park; Sarah Leary; Irv Bernstein; members of the Jarvik/Crosslin laboratory; faculty, trainees, and staff in the Biomedical & Health Informatics and Pediatric Hematology-Oncology programs at University of Washington; and members of the eMERGE EHRI and CSER EMR working groups.

This work was supported in part by the National Institutes of Health, National Library of Medicine (NLM) Biomedical and Health Informatics Training Program at the University of Washington (Grant Nr. T15LM007442). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Chapter 1. BACKGROUND & SIGNIFICANCE

1.1 The Potential of Precision Genomic Medicine

Precision Medicine is the use of prevention and treatment strategies which take into account deeply characterized biologic, environmental, and behavioral variability among individual patients[1]. A major application of precision medicine is the use of genomic test results to personalize clinical care. A real-world example of precision genomic medicine is pharmacogenomics, which is the use of drug-genetic variant interactions to guide drug selection and dosing[2]. Another example is targeted therapies, or the use of drugs targeting specific genetic variants implicated in a disease process[3]. As our understanding of human genomics deepens, the potential for precision genomic medicine to improve care processes and patient outcomes will increase. Therefore, it is critical to understand and optimize the factors that contribute to the effectiveness of precision medicine.

1.2 Inconsistencies in Genetic Test Interpretation

The effectiveness of precision genomic medicine depends upon having reliable genomic test results[4,5]. Unfortunately, multiple studies have reported significant disagreements in the interpretation reports provided by different sequencing laboratories performing the same test[6–9]. For example, laboratories analyzing the same biological specimen may report different sets of sequence variants. Additionally, laboratories given the same sequence variant may report differing interpretations of its clinical significance. Factors contributing to diverse results for the same genomic test may include:

1. Differing sequencing pipeline characteristics;
2. Differing variant interpretation criteria;
3. Gaps and inconsistencies within large variant knowledge bases; and
4. Siloing of knowledge within individual institutions[10].

These discrepancies negatively affect the reliability of genomic test results for clinical use, as well as the quality of submissions to variant knowledge bases. Because variant knowledge bases provide evidence supporting variant classification, this causes a vicious cycle that self-reinforces unreliability for genomic test results.

Large variant knowledge bases, such as National Center for Biotechnology Information (NCBI) ClinVar[11], Online Mendelian Inheritance in Man (OMIM)[12], and Human Gene Mutation Database (HGMD)[13], depend upon sequencing laboratories to voluntarily publish their results. However, only a small fraction of variant knowledge is published, and the quality of submissions are inconsistent. Gaps and inconsistencies within large variant knowledge bases may contribute to inaccuracies in variant interpretation. For example, Amendola et al. found only 11.4% of 616 variants originally classified as "pathogenic" based on HGMD annotations alone remained classified as pathogenic or likely pathogenic after accounting for allele frequency, segregation, and functional data[9]. Additionally, Dorschner et al. reviewed 239 "pathogenic" variants in HGMD and found that in only 7.5% of the cases was the original publication supportive of the same classification[14]. In 2015, Rehm et al. found that 17% of ClinVar entries with submissions from multiple sequencing laboratories contained significantly divergent interpretations[15]. These examples underscore the importance of variant curation efforts to resolve inconsistencies in existing variant knowledge bases.

Clinical Genome Resource (ClinGen) is a major National Institutes of Health (NIH)-funded program that seeks to build an authoritative, open-access, centralized genomic knowledge base of gene-disease relationships, gene variant pathogenicity, and variant actionability[15–17]. A cornerstone of the ClinGen program is its partnership with ClinVar. ClinGen and ClinVar partnered to develop the Star System, which ranks variant interpretation entries in ClinVar based on level of evidence review, ranging from no stars (i.e. single submitter, no evidence criteria provided) to four stars (i.e. endorse by published practice guideline)[15]. ClinVar entries can be imported into the ClinGenKB platform for secondary review by ClinGen-approved expert panels, supported by a growing suite of curation tools[15,18,19]. Finally, ClinGen serves as an approval body for ClinVar expert panels and a clearinghouse for ClinVar practice guidelines. The ClinGen-ClinVar partnership has steadily improved ClinVar's quality and usefulness as a source of variant knowledge. However, persistent gaps and inconsistencies within ClinVar and other variant knowledge bases continue to limit their usefulness in clinical care.

A major issue facing genomic medicine is siloing of genomic variant knowledge within individual institutions. Variant interpretation relies upon the strength of association between a variant and a

specific disease phenotype. Because pathogenic variants are rare, it is essential that sequencing laboratories share information with large variant knowledge bases, such as ClinVar, to build the evidence base supporting appropriate variant interpretation. However, submissions to publicly available knowledge bases are voluntary, and only a minority of variants are submitted. Barriers to knowledge sharing may include:

1. Low prioritization (i.e. sequencing laboratories are busy with their primary mission of reporting test results);
2. High perceived burden (i.e. preparing ClinVar submissions involves a non-trivial amount of time and effort);
3. Low perceived benefit (i.e. no tangible reimbursement for submission); and
4. Potential competitive benefit in withholding proprietary data;

While commercial entities have established networks for inter-laboratory sharing of variant knowledge, the contents are proprietary and only available for those who pay a subscription, thereby making it inaccessible for the scientific community. The siloing of data within individual institutions and proprietary networks contributes to gaps in publicly available knowledge bases, which contributes to variant misclassifications with potentially adverse clinical outcomes. Pepin et al. found that inadequate sharing of their laboratory's private data contributed to 33% of discrepant classifications their group encountered in routine requests for secondary opinions[20]. They point to the high-perceived burden in curating variants as a significant barrier in sharing results. Interventions that decrease the burden of curating variants may be successful in improving the submission rate.

1.3 The Use of Standards and Guidelines In Variant Interpretation

Historically, a major contributor to inconsistencies within variant knowledge bases is the diversity of methods and terminologies utilized in variant interpretation. Each sequencing laboratory previously used its own pathogenicity criteria to classify variants. In 2015, the American College of Medical Genetics and Genomics (ACMG) and the Association for Molecular Pathology (AMP), two leading groups of professionals involved in genomic test interpretation, published joint guidelines standardizing sequence variant classification criteria for Mendelian disorders[21]. These guidelines defined 28 evidence codes and 20 rules to support classification into five categories: Pathogenic (P), Likely Pathogenic (LP), Variant of Uncertain Significance (VUS), Likely Benign

(LB), or Benign (B) (See **Appendix A** for evidence codes and rules). The "Likely" modifier is used when there is greater than 90% but not full certainty of assignment to that group, and the guidelines recommend that "Likely" variants should be treated similarly as their high certainty counterparts in clinical decision making. A VUS does not have strong evidence either for or against pathogenicity, and should not be used clinically. While these initial guidelines still require further refinement, its establishment is a major milestone in the effort to improve the reliability of genomic tests and variant knowledge bases.

Despite widespread acceptance of the ACMG/AMP Guidelines, sequencing laboratories continue to report diverging interpretations for the same variants. The Clinical Sequencing Exploratory Research (CSER) Consortium conducted the "Variant Bake-off" study to pilot and evaluate the effectiveness of the ACMG/AMP Guidelines[7]. A "bake off" is a competition between different groups to produce the best results for a specific challenge. In the CSER Variant Bake-off study, ninety-nine previously unseen test variants were classified by at least three of nine Clinical Laboratory Improvement Amendments (CLIA)-certified genomic testing laboratories, including two that were involved in developing the ACMG/AMP Guidelines. Each laboratory first used their own classification criteria, then switched to using the ACMG/AMP Guidelines. This study found that within each laboratory, the two methods produced similar results (21% discordance rate, K-alpha mean 0.91, K-alpha range 0.77-1.00)[7]. In general, the ACMG-AMP Guidelines were more conservative, resulting in more VUS classifications than other classifications (particularly B or LB)[7]. Only 5% of all variants had a clinically significant change in classifications, defined as switching between a pathogenic class (P or LP) and a non-pathogenic class (B, LB, or VUS), which has the potential to change medical management. In contrast, the study demonstrated inadequate concordance when classifications were compared across laboratories, regardless of whether the laboratories used their own protocols (66% discordance, K-alpha 0.76) or the ACMG/AMP Guidelines (66% discordance rate, K alpha = 0.72)[7]. In summary, 22% of all variants had clinically significant discordant classifications, indicating unacceptably low test reliability for medical use.

1.4 Collaboration, Knowledge-sharing, and Consensus-making in Variant Interpretation

The CSER Bake-off study likely over-estimates discordance rates due to its focus on rare variants, as well as the timing of the study soon after the publication of the ACMG/AMP guidelines, when laboratories may have not yet become proficient in using the guidelines. However, lessons learned from this study are instructive in identifying factors which contributing to inter-laboratory discordance. In this study, factors contributing to discordance included:

1. Inadvertent calculation errors;
2. Differing applications of the same guidelines;
3. Interpreting pathogenicity for different disease phenotypes; and
4. Inadequate information sharing[7].

Concordance improved significantly after a post hoc collaborative consensus process, which included the following steps:

1. Calculation verification using the ClinGen Pathogenicity Calculator, a newly-developed decision support tool implementing the ACMG/AMP Guidelines[18]; and
2. Cross-institutional discussion and knowledge-sharing[7].

The ClinGen Pathogenicity Calculator identified that in 4.5% (16 of 353) of the classifications, the original classification was incongruent with those calculated from the evidence codes and guideline rules. Slightly more than half (9 of 353, or 2.5%) were determined to be inadvertent calculation errors and modified. The others (7 of 353, or 2%) were due to professional judgement overriding the guidelines and left unchanged. The remaining disagreements were resolved via cross-institutional discussion. Overall, the collaborative consensus process resulted in the discordance rate dropping from 66% to 29%, and clinically significant discordance rate dropping from 22% to 5%. The dramatic increase in concordance suggests that collaboration, knowledge-sharing, and a consensus process can dramatically improve precision in genomic test interpretation. The remaining 5% cases of unresolved, clinically significant disagreement suggests that in the absence of perfect knowledge, professional judgment continues to play an important role even with guideline use.

While the CSER Variant Bake-off study is an unusual example of multi-institutional collaborative variant interpretation, other studies also support the power of collaboration, knowledge-sharing, and a consensus process in improving classification reliability[7]. Pepin et al. suggested that discrepancies in variant classifications between a laboratory with specialized expertise and other laboratories may be due to the failure of that specialized laboratory to submit their internal variant data to ClinVar[20]. Harrison et al. also demonstrated that a collaborative consensus process among different sequencing laboratories, in this case re-examining their discordant submissions in ClinVar while sharing all available laboratory-specific evidence, also results in significant improvements in inter-laboratory concordance[22]. Therefore, collaboration appears to play a significant role in achieving high inter-laboratory concordance for sequence variant interpretation.

Collaboration similarly plays an important role in the routine function of individual sequencing laboratories. Multiple laboratory personnel play specific essential roles in the genomic testing analytic pipeline, from sample acquisition to test report generation, and must collaborate to ensure proper function of this pipeline. Multiple laboratories within the same institution may collaboratively perform certain specialized tests by sharing analytic pipeline components. Laboratories may collaborate with outside experts when specialized knowledge is required to ensure accurate interpretation[20]. Finally, laboratories collaborate externally by voluntarily publishing variant interpretation knowledge to shared knowledge bases, such as ClinVar.

1.5 Computer Supported Collaborative Variant Interpretation

Enhanced collaboration, knowledge-sharing, and consensus-making in variant interpretation may help to improve the reliability of individual genomic test results, as well as the quality of shared variant knowledge bases such as ClinVar[7,15,20,22]. Computerized tools may have a role to play in supporting these processes. The CSER Bake-off study demonstrated the utility of a computerized tool supporting variant classification[7]. In the study, a pathogenicity calculator was able to detect a 2.5% manual calculation error rate. However, most discordant classifications did not involve calculation errors. Rather, they arose from inter-rater differences in access to information and professional judgment, and were resolved only after a collaborative consensus process involving inter-laboratory knowledge sharing and discussion. Therefore, a tool that supports both guidelines-based calculations and inter-rater collaboration, knowledge-sharing, and consensus-making may

be more useful in collaborative variant interpretation than a stand-alone calculator. Based on the lessons learned from the CSER Bake-off study, we propose the following functionality specifications for an openly available tool supporting collaborative variant interpretation (see **Table 1.1)**:

Table 1.1. Specifications: Computerized Tool Supporting Collaborative Variant Interpretation

Functionality	Short Description
1. Variant evidence management	Retrieve, organize, and display all variant evidence required for pathogenicity classification.
2. Variant classification decision support	Support adherence to the ACMG/AMP Guidelines for variant classification.
3. Role-based collaboration support	Provide support for individuals to collaborate on variant interpretation based on their role in the interpretation workflow.
4. Consensus-making support	Provide support for a consensus-making process to resolve differences in variant classification.
5. Variant knowledge base management	Capture variant classifications and supporting evidence in a local knowledge base for reuse, analytics, discovery, and external sharing.
6. Variant knowledge-sharing	Export and sharing of variant data with other systems, such as NCBI ClinVar.

1.5.1 Variant Evidence Management

Variant classification requires reviewing all relevant data, including clinical and phenotypic data, variant annotation data generated by the genomic testing analytic pipeline, and other sources of evidence, in order to synthesize an interpretation. The AMCG-AMP Guidelines categorizes the diverse array of 28 evidence criteria required for evaluating variant pathogenicity into eight categories, including: 1) Population data; 2) Computational and predictive data; 3) Functional data; 4) Segregation data; 5) de novo data; 6) Allelic data; 7) Other database; and 8) Other data[21]. Evidence is obtained from a variety of internal and external sources, including online databases, computer software, from clinical information systems, and from the patient. Therefore, end users may desire functionality supporting retrieval, organization, and display of all variant annotation information required for pathogenicity classification into one consolidated virtual space, to facilitate evidence review for variant interpretation.

1.5.2 Variant Classification Decision Support

The ACMG/AMP Guidelines were developed to standardize sequence variant classification practices for Mendelian disorders, and are now widely accepted within the field of genomic medicine[7,21,22]. Raters must systematically evaluate 28 evidence criteria, then apply 20 logical rules, to calculate an guidelines-based variant interpretation (see *Section 1.3*)[21]. The classification criteria were derived from existing, long-standing, commonly agreed-upon practices within the clinical genomics community[21]. However, formalizing these existing practices into logical rules introduced a new layer of complexity and created opportunities for manual calculation errors, even by experienced raters[7,18,22]. Therefore, end users may desire decision support functionality applying the ACMG/AMP Guidelines for variant classification.

1.5.3 Role-Based Collaboration Support

The process of variant interpretation involves multiple personnel playing various distinct roles at different time points in the variant interpretation workflow. For example, case intake specialists, bioinformatics technicians, senior molecular geneticists, and ad hoc external expert consultants (e.g. other medical specialists, other laboratories) all perform different roles in the collaborative interpretation process[20,23]. Recent publications have focused on high level collaboration among senior personnel of multiple sequencing laboratories in resolving differences in variant interpretation opinions[7,22]. These inter-laboratory collaborations, which generally occur in the research context, are critical for curating the public variant knowledge base. However, in routine practice, inter-laboratory collaborations occur at much less frequency than intra-laboratory collaboration among various laboratory personnel (e.g. case intake specialists, bioinformatics technicians, senior molecular geneticists) playing different roles in the variant interpretation process. For example, each participant upstream to the senior molecular geneticist obtains or generates data necessary for variant interpretation (e.g. clinical data, phenotype, variant annotation data), and may play a role in filtering and prioritizing variants for additional review. External collaborators, such as other medical specialists or other laboratories, may be invited in a small subset of cases to contribute expertise for narrowly defined questions. Therefore, end users may desire functionality supporting role-based, internal or external collaborative work in variant interpretation.

1.5.4 Consensus-Making Support

Evidence suggests that consensus processes are effective in increasing inter-rater concordance for variant interpretation[7,22]. The CSER Bake-off study describes a consensus process in which different laboratories corresponded via email and phone to discuss the rationale and supporting evidence for their interpretation opinion[7]. Harrison et al. describes a similar consensus process for inter-laboratory collaboration[22]. In both cases, the ACMG/AMP Guidelines provided a structure for each laboratory to present their rationale and the evidence supporting their specific interpretation, including when they disregarded the guidelines due to professional judgment. However, both studies do not describe using tools or processes designed specifically to provide a structured process for multiple experts to arrive at a consensus. End users may desire functionality that supports a structured consensus-making process to resolve differences in variant interpretation.

1.5.5 Variant Knowledge Base Management

Laboratories rely upon both internal and external data as evidence sources in variant interpretation[7,20,21]. Publicly available data may be incomplete or inconsistent (see *Section 1.2*). Certain data may only be available internally, particularly for specialized laboratories with a particular expertise[20]. Examples of internal laboratory data that are useful for variant interpretation include institutional allele frequency, segregation analysis results, and functional data. Electronic tools can capture and organize variant data in an internal variant database, to support evidence generation as well as knowledge sharing and data export. Additionally, having a database of previously authored interpretation text can save time when interpreting a previously encountered variant. Therefore, end users may desire functionality that captures variant classifications and supporting evidence in a local knowledge base for reuse, analytics, discovery, and external sharing.

1.5.6 Variant Knowledge-Sharing

Public knowledge-sharing is critical for establishing the evidence base necessary for reliable variant classification (See *Section 1.2*). Large, publicly available variant knowledge bases such as ClinVar are dependent upon voluntary submissions from sequencing laboratories. While participation is actively encouraged by organizations such as ClinGen[15], external knowledge base

submissions are an uncompensated activity and may receive lower prioritization compared with the primary mission of the clinical sequencing laboratory (i.e. providing clinical genomic test results). Even if laboratories desire to share knowledge externally, laboratories may lack tools supporting transfer of structured variant information from laboratory systems into external knowledge bases. Therefore, end users may desire functionality that supports the export and sharing of internal variant data with other systems, such as ClinVar.

Chapter 2. RELATED WORKS & GAPS

Based on the specifications defined in *Section 1.5,* we conducted a review of existing, openly available tools that provide decision support implementing the ACMG/AMP Guidelines for variant interpretation. See **Table 2.2** for a summary of the existing tools' functionality and gaps. While existing, commercial-available tools may perform some of the functionality described in our specifications, we did not include these tools in this review due to lack of access. The scope of our review is limited to tools that are openly available to the entire clinical genomics community and could be incorporated into existing workflows without significant upfront and ongoing cost.

Table 2.2. Functionality and Gaps in Three Existing Tools (N=4)

"Yes" indicates partial to full support of the specifications defined in Table 1.1.

	Genetic Variant Interpretation Tool (GVIT) [24,25]	**InterVar**[26,27]	**ClinGen Pathogenicity Calculator**[18,28]	**ClinGen Variant Curation Tool**	**Number of Tools with Functionality**
1. Variant evidence management	**NO**	**YES**	**YES**	**YES**	3 (75%)
2. Variant classification decision support	**YES**	**YES**	**YES**	**YES**	4 (100%)
3. Role-based collaboration support	**NO**	**NO**	**NO**	**NO**	0 (0%)
4. Consensus-making support	**NO**	**NO**	**YES**	**YES**	2 (50%)
5. Variant knowledge management	**NO**	**NO**	**YES**	**YES**	2 (50%)
6. Variant knowledge-sharing	**YES**	**YES**	**YES**	**YES**	4 (100%)

2.1 EXISTING TOOLS SUPPORTING GUIDELINES-BASED INTERPRETATION

2.1.1 Genetic Variant Interpretation Tool (GVIT)[24,25]

The Genetic Variant Interpretation Tool (GVIT), developed by the University of Maryland Program for Personalized and Genomic Medicine, is an easy to use, freely available, web-based passive decision support tool that implements the ACMG/AMP Guideline using formal reasoning rules. Of the three tools examined, this tool had the simplest feature set. The tool utilizes a web-based GUI. There is no user registration and log in requirement to use the tool. There is no

Application Programming Interface (API). The tool does not support variant annotation management; it does not support importing variant data. The tool supports variant classification decision support. Users interact with the tool by manually select the appropriate ACMG/AMP evidence criteria from a long checklist. The application then dynamically calculates and displays the appropriate classification. The tool is intended for a single user and does not support role-based collaboration or consensus-making. The tool does not perform variant knowledge base management. The tool captures calculated interpretation results for download, but does not capture the supporting evidence criteria or other rationale for the calculated classification. The tool partially supports variant knowledge-sharing. Users can download the processed variants as a custom Comma Separated Value (CSV) file containing three columns: Patient ID, Variant ID, and Interpretation. While this file is human- and machine-readable, the file would require additional data transformation and processing in order to export into downstream systems such as ClinVar. In conclusion, the Genetic Variant Interpretation Tool is a basic, straightforward passive rules calculator for the ACMG/AMP Guidelines. See **Figure 2.1** for screenshots.

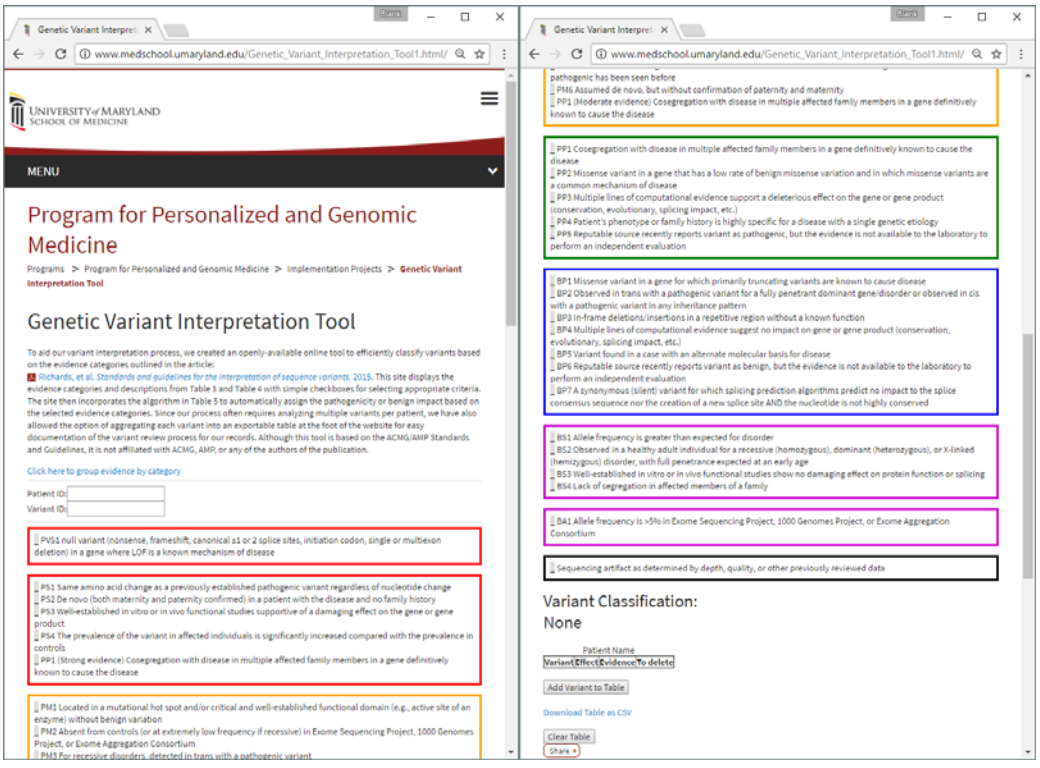

Figure 2.1. Screenshots: Genetic Variant Interpretation Tool.

Left: Top half. Right: Bottom half.

2.1.2 InterVar[26]

InterVar, developed by the Wang Genomics Lab at Columbia University, is an open source, Python-based command line-driven application that semi-automatically classifies variants based on annotation data. wInterVar, the server-based version of InterVar, has a web-based GUI interface and a more limited feature set (i.e. supports semi-automated classification of exonic variants, does not support indels, and contains an additional passive rules calculator interface if evidence codes are already know)[27]. The tool is highly interoperable and designed to be used as a component of an analytic pipeline. The tool supports variant annotation management and classification decision support. The tool accepts files in the following three formats:

1) Custom Tab-Delimited File (TSV) with pre-annotated variants;
2) Unannotated Variant Call Format (VCF);
3) ANNOVAR input format files.

TSV files are accepted without additional processing, and VCF and ANNOVAR files are first automatically annotated using the ANNOVAR annotation tool[29], also developed by the Wang Genomics Lab, prior to processing. The tool then predicts the assignment of 18 of 28 evidence criteria from the ACMG/ACMP Guidelines, including PVS1, PS1, PS4, PM1, PM2, PM4, PM5, PP2, PP3, PP5, BA1, BS1, BS2, BP1, BP3, BP4, BP6, and BP7 (See **Appendix A** for evidence code definitions). The user then reviews and either confirms or overrides these predicted codes. The remaining 10 evidence codes (PS2, PS3, PM3, PM6, PP1, PP4, BS3, BS4, BP2, and PB5) must be assigned manually within the application or by uploading an evidence file. The tool then returns a detailed, structured CSV file containing predicted pathogenicity classification and supporting evidence codes, which can be extracted for export into downstream systems such as ClinVar. The tool is intended for a single user and does not support role-based collaboration or consensus-making. The tool does not perform variant knowledge base management. In conclusion, InterVar is a highly interoperable command line tool that could easily be incorporated into existing analytics pipelines to perform semi-automated pathogenicity classifications. See **Figure 2.2** for screenshots of wInterVar, the server-based version of InterVar (command line-based InterVar is not pictured).

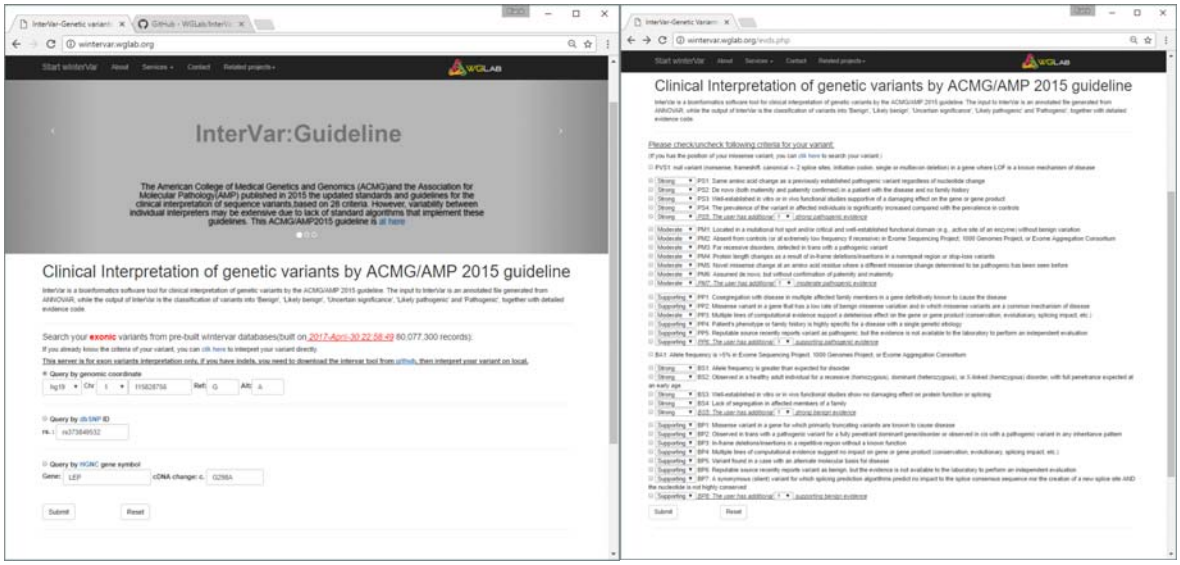

Figure 2.2. Screenshot: wInterVar.

Left: wInterVar server-based tool. Right: Passive rules calculator interface.

2.1.3 Clinical Genome Resource (ClinGen) Pathogenicity Calculator[18,28]

One of ClinGen's primary missions is to develop tools that facilitate genomic knowledge sharing and curation. ClinGen has developed two tools which are relevant to variant interpretation: the ClinGen Pathogenicity Calculator[18,28], and the ClinGen Variant Curation Tool[30]. We will discuss the Pathogenicity Calculator in this section, and the Variant Curation Tool in the next section.

The ClinGen Pathogenicity Calculator, developed at Baylor College of Medicine for ClinGen, is a web-based, freely available, open source, decision support tool that implements the ACMG/AMP Guidelines using formal reasoning rules. This tool was used in the CSER Variant Bake-off study as part of the consensus process. The tool is integrated with the ClinGen suite of web applications, including the ClinGen Allele Registry. Users must register for a free account and log in to the website to access the tool. Users may interact with the application using the web-based interactive, multi-layered GUI or programmatically via the API, which allows for possible integration with existing analytics pipelines. The tool pulls in variant annotations from a variety of sources, but does not support importing of annotation data from a laboratory's pipeline. The tool provides variant classification decision support. Users input evidence code assignments by interacting with a dynamic "evidence document" web-based interface via the mouse, or programmatically using the API. The user may enter additional evidence as free text comments or Universal Resource

Locator (URL) links, as well as change the strength of certain evidence criteria if their professional judgment diverges from the guidelines. The tool then automatically assigns the appropriate classification and documents the evidence that supports that assignment. The tool's rules engine supports multiple interpretation criteria in anticipation of additional guidelines that may be introduced in the future. The tool supports consensus-making, capture of variant knowledge, and knowledge sharing by supporting the storage and sharing of variant interpretations and evidence within a defined collaborating group. This tool can be used to facilitate consensus by displaying multiple independent assessments of the same variant, but it does not provide additional support or structure for the consensus process. The tool does not support role-based collaboration, such as among personnel within the same laboratory. The tool outputs a structured, human-readable variant reports as a dynamic webpage with a static URL, which can be printed as a PDF file and uploaded as supporting evidence to ClinVar, or parsed with a script to extract structured data. The tool does not natively support export of a machine-readable file format such as CSV or XML. In summary, the ClinGen Pathogenicity Calculator is a fully-featured passive rules calculator for the ACMG/AMP Guidelines which supports inter-laboratory collaboration, knowledge-sharing, and consensus-making. See **Figure 2.3** for a screenshot.

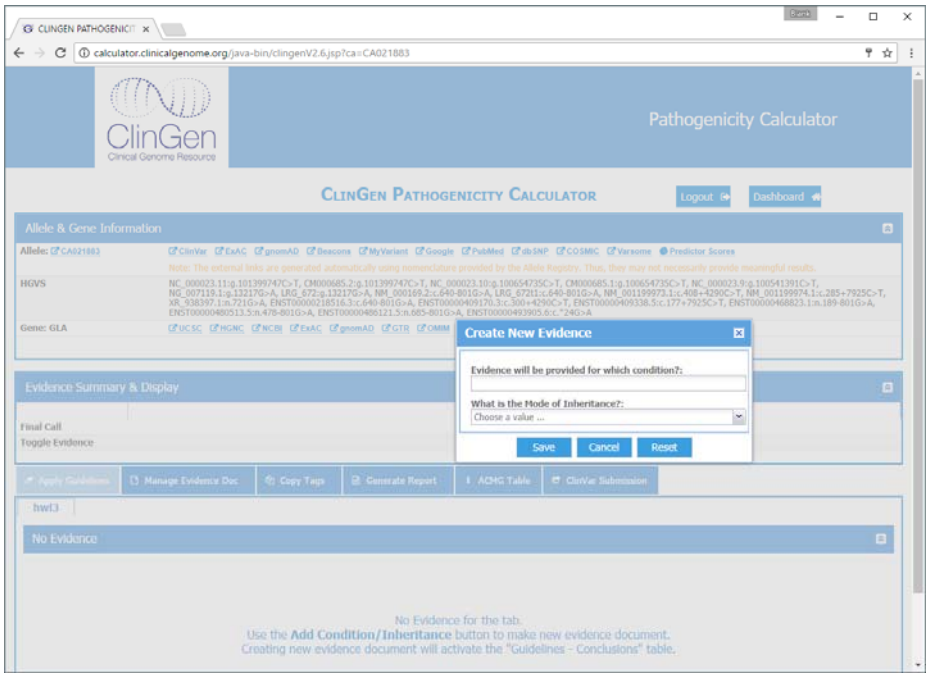

Figure 2.3. Screenshot: ClinGen Pathogenicity Calculator.

2.1.4 Clinical Genome Resource (ClinGen) Variant Curation Tool[30,31]

The ClinGen Variant Curation Tool is a component of the web-based ClinGen Curator Interfaces web application, which was developed by the Cherry Lab at Stanford University to support the work of ClinGen-approved curators in developing evidence-based expert consensus knowledge for variants and genes[30,31]. Access to the tool is limited to ClinGen-approved curators who are members of ClinGen Clinical Domain Working Groups. The tool is not meant to be used by individual laboratories outside the context of ClinGen knowledge base curation (e.g. performing variant interpretation to report individual genomic test results). The Variant Curation Tool pulls relevant evidence data from web-based sources, such as population data or results of prediction algorithms, into one interface. ClinGen-approved curators then use the interface to enter a curated interpretation. Users interact with the application using the web-based interactive GUI. The tool provides guidelines-based decision support for variant classification. Users input evidence code assignments by clicking on each evidence code and selecting "Met", "Not Met", or "Not Evaluated". The tool then calculates the pathogenicity classification. If the user manually changes the classification, the user must enter a rationale to justify over-riding the guidelines. The tool supports consensus-making by supporting the work of the ClinGen-approved curator who is reviewing existing evidence. The tool does not provide additional support or structure for the consensus process. The tool does not support role-based collaboration, such as among personnel within the same laboratory. Interpretations are saved in the ClinGen knowledge base for knowledge-sharing and public use. The tool does not support the management of a separate knowledge base (e.g. an individual laboratory's internal knowledge base). The tool does not support export of a machine-readable file to facilitate data export to sources other than the ClinGen knowledge base. Therefore, the tool cannot be incorporated directly into an individual laboratory's variant interpretation pipeline. In summary, the ClinGen Variant Curation Tool is a fully-feature passive rules calculator for the ACMG/AMP guidelines that supports evidence-based variant knowledge curation exclusively for the ClinGen knowledgebase. See **Figure 2.4** for a screenshot.

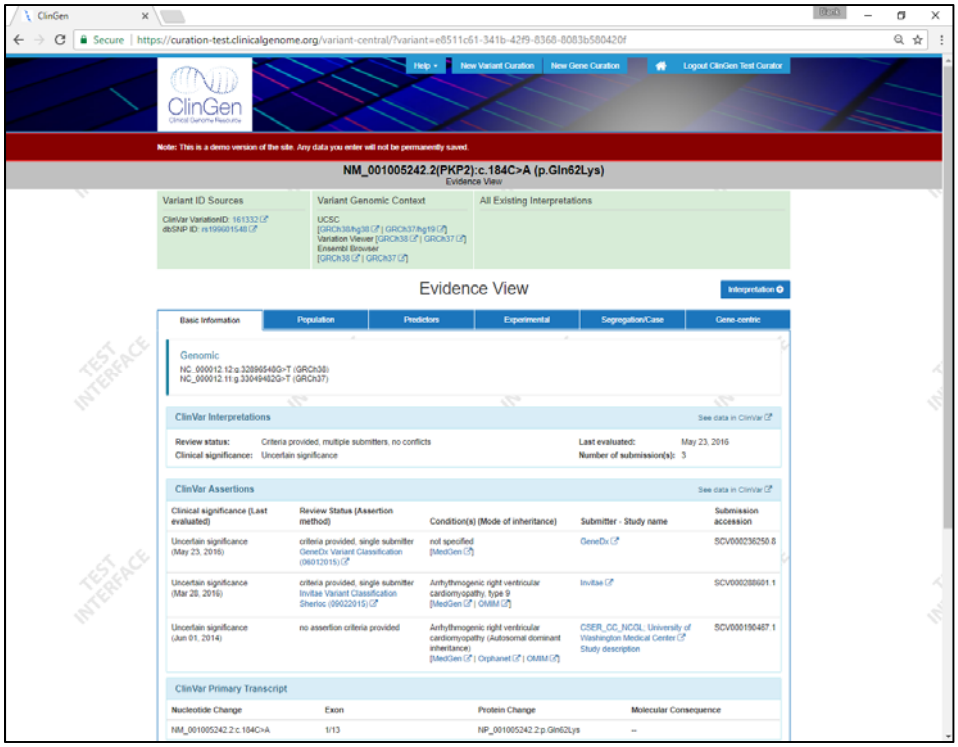

Figure 2.4. Screenshot: ClinGen Curator Interfaces.

2.2 GAPS AND OPEN QUESTIONS

The four existing tools offer contrasting approaches to supporting variant interpretation (See **Table 2.2**). Each tool provided variant classification decision support. GVIT offers a simple and focused web-based rules calculator. InterVar offers an interoperable, semi-automatically pathogenicity classifications engine. ClinGen Pathogenicity Calculator offers a jack-of-all-trades application featuring a detailed rules calculator, integration with ClinGen resources, collaboration functionality to support data-sharing and interpretation comparisons between collaborators, and the ability to generate a PDF evidence document for submission to ClinVar. The ClinGen Variant Curation Tool offers a full-featured variant interpretation interface for ClinGen Knowledge Base curators. Three of the tools provide data export functionality that can facilitate knowledge-sharing. While there is no way to export data directly from the ClinGen Variant Curation Tool, all data are saved into the openly available ClinGen Knowledge Base. Each tool offers a different set of features and have different use cases.

There were also gaps in the functionality offered by the existing tools. Most of the tools offer some evidence management functionality, including InterVar, ClinGen Pathogenicity Calculator, and ClinGen Variant Curation Tool. Rarely supported functionality includes role-based collaboration, consensus process, and variant knowledge management. The ClinGen Variant Curation Tool was developed specifically to support expert curated interpretations for the ClinGen Knowledge Base, which is a collaborative effort. The ClinGen Pathogenicity Calculator provide limited inter-laboratory collaboration functionality. However, none of the tools provides direct support for routine, role-based collaboration occurring between various personnel and experts consultants collaborating in variant interpretation within single sequencing laboratories. While the ClinGen Pathogenicity Calculator provides a way to save and share variant data within a team, none of the tools provides support for managing a laboratory's internal variant knowledgebase. Each tool supports open knowledge sharing within the clinical genomics community. However, the ClinGen Variant Curation Tool can only be used by ClinGen-approved curators to manage the ClinGen knowledge base, and cannot be used by individual laboratories for clinical variant interpretation.

These gaps and open questions led us to propose the design for MEdicine Gene Annotation (MEGA), a computerized tool supporting team-based collaboration, knowledge sharing, and consensus-making in variant interpretation.

Chapter 3. SPECIFIC AIMS & INNOVATION

Our overall goal is to design a novel computerized tool supporting team-based collaboration, knowledge sharing, and consensus making in variant interpretation.

3.1 AIM 1: INITIAL DESIGN DEVELOPMENT

Aim 1. Use literature review and informal expert input to characterize current variant interpretation workflow, in order to design an initial workflow, graphical user interface (GUI), and technology architecture for a computerized tool supporting *collaborative* variant interpretation. ***Aim 1.A.*** Use literature review and informal expert input to understand and document typical workflow for variant interpretation; then develop a proposed workflow to support *collaborative* variant interpretation. ***Aim 1.B.*** Use workflows developed in *Aim 1.A* and informal expert input to develop an initial GUI and technology architecture to support *collaborative* variant interpretation.

3.2 AIM 2: USER-CENTERED DESIGN DEVELOPMENT

Aim 2. Use user-centered design methodology to a) further characterize the typical workflow for variant interpretation, b) define user requirements and desired functionality, and b) to refine the design for a workflow, GUI, and technology architecture supporting *collaborative* variant interpretation. ***Aim 2.A.*** Use semi-structured interviews with intended users (i.e. clinical molecular geneticists) to a) characterize existing variant interpretation workflow and b) define user requirements and desired functionality for a computerized tool supporting *collaborative* variant interpretation. ***Aim 2.B.*** Use design feedback sessions with intended end users to obtain user feedback on the proposed *collaborative* workflow and GUI developed in *Aims 1.A and 1.B*. ***Aim 2.C.*** Use feedback obtained from *Aims 2.A and 2.B* to develop a revised workflow, GUI, and technology architecture supporting role-based, *collaborative* variant interpretation.

3.3 INNOVATION

This research program seeks to accelerate the adoption of precision genomic medicine by supporting and innovating upon the existing variant interpretation workflow. Whereas other

similar tools have focused on implementing decision support for guidelines-based determination of pathogenicity classification, our proposed tool focuses on supporting team-based collaboration, knowledge sharing, and consensus making. Additionally, we used a user-centered design approach, which is rarely reported in genomic, bioinformatics and translational informatics, in order to design a tool that reflects the needs of actual end users[32–38].

Chapter 4. METHODS

We engaged in a two-part process in conducting this research: initial design development and user-centered design development.

4.1 INITIAL DESIGN (AIM 1)

4.1.1 Project Requirements Definition and Initial Workflow Design (Aim 1.A)

Our initial goal was to design an openly available tool that supports collaborative variant interpretation within a large, multi-institutional setting, similar to the CSER Variant Bake-off scenario. Our initial steps were to develop an understanding of the content domain, problem scope, and project requirements[39]. While the content domain is genomics, issues relating to collaboration, knowledge-sharing, and consensus-making cross domain boundaries. We sought to understand:

1) Typical workflow and information requirements for variant interpretation;
2) The role of collaboration in variant interpretation and analogous scenarios from other domains; and
3) The role of consensus-making in variant interpretation and analogous scenarios from other domains.

We then pursued the following strategies to achieve the above objectives:

1) Review of published literature;
2) Obtaining informal input from genomics subject matter experts (SMEs); and
3) Obtaining informal input from non-genomics SMEs.

For the purposes of our study, we defined SMEs as geneticists with experience in collaborative genomic variant annotation and interpretation, without regard to indication- or disease-related expertise. The University of Washington (UW) participates in two major National Human Genome Research Institute (NHGRI)-sponsored collaborative multi-institutional genomics medicine consortia: Electronic Medical Records and Genomics (eMERGE) and CSER. UW also has multiple clinical and research laboratories involved in performing, interpreting, and/or developing genomic tests. Therefore, many SMEs at UW have significant experiences in collaborative variant interpretation within multi-institutional consortia and other settings.

The primary investigator is not a genomics SME, so we utilized literature review, observation of genomics research meetings, and informal input from SMEs to develop an understanding of the genomics content domain. The primary investigator observed and participated regularly in a genomics research laboratory's group meetings at the University of Washington, as well as observed and participated in national workgroup teleconference meetings within the eMERGE and CSER consortia. As the design concept began to emerge, the primary investigator presented early stages of the design concept to various audiences in different settings, including genomic research group meetings, eMERGE consortia workgroup meetings, the University of Washington Precision Medicine Interest Group, and presentations to non-genomics, academic audiences. A deeper understanding of the genomics domain was gained from interactions with SMEs, and diverse inputs and perspectives were obtained from SMEs and non-SMEs alone regarding the role of collaboration, consensus, and information sharing in science and medicine. These above steps provided us with a broad understanding of the content domain, problem scope, and project requirements, which we then used to model a generalized, typical workflow for genomic variant interpretation, as well as a proposed workflow to support collaborative variant interpretation.

4.1.2 Graphical User Interface and Technology Architecture Design (Aim 1.B)

In the initial design phase, we evaluated an existing prototype tool developed to capture variant annotations and classifications for a genomics research laboratory at the University of Washington. While this prototype tool was not widely adopted for use, the tool provided valuable information that informed our initial design process. We analyzed the tool for the following elements:

1) End user information requirements;
2) Factors affecting adoption;
3) Technology architecture; and
4) Visual design;

The existing prototype was adopted as the starting point of the current design process, and its GUI and technology architecture were adopted and revised based on the results of the requirements definition (from *Section 4.1.1*), the proposed workflow (from *Section 4.1.1*), and the existing tool evaluation (from *Section 4.1.2*). Prototypes, mockups, and workflow diagrams

were created using a REDCap[40] test environment, LucidChart[41], Microsoft Visio Professional 2016[42], and Microsoft PowerPoint Professional 2016[43].

4.2 User-Centered Design (Aim 2)

4.2.1 User Interviews & Design Feedback Sessions (Aims 2.A. & 2.B.)

After creating the initial workflow, GUI, and technology architecture designs, we utilized user-centered design methodology to refine the design. We selected user-centered design methodology in order to design the tool to best fit the needs of the intended end users, their tasks, and their environments[32]. This phase of the study was determined to be human subjects research that qualifies for exempt status by the University of Washington Human Subjects Division (HSD) Institutional Review Board (IRB).

4.2.2 Participant Recruitment and Enrollment

For the user-centered design process, we sought participants who would be most representative of the intended end users of the tool. Our inclusion and exclusion criteria were as follows:

- Inclusion Criteria:
    - Interprets genetic or genomic tests as part of normal work responsibilities (clinical or research)
    - Affiliated with the University of Washington
    - Age 20 or older
    - Fluent in written and spoken English
    - Consents to audiotaped interview
- Exclusion Criterion:
    - Declines participation in study

We identified individuals affiliated with the University of Washington with experience in genetic variant interpretation as potential participants. The primary investigator sent prospective participants a brief email explaining the purpose of the study with a link to the study enrollment website. Prospective participants received up to one additional reminder email prior to enrollment. Interested individuals completed an integrated electronic consent form and brief screening

questionnaire on the study enrollment website, which was hosted on the University of Washington's Research Electronic Data Capture (REDCap) platform[40]. The screening survey contained 13 questions regarding the participant's basic demographic information, work roles, relevant experience, and contact information (See **Figure 4.1**). The primary investigator then reviewed screening questionnaire responses to ensure study eligibility and contacted the participant to complete enrollment.

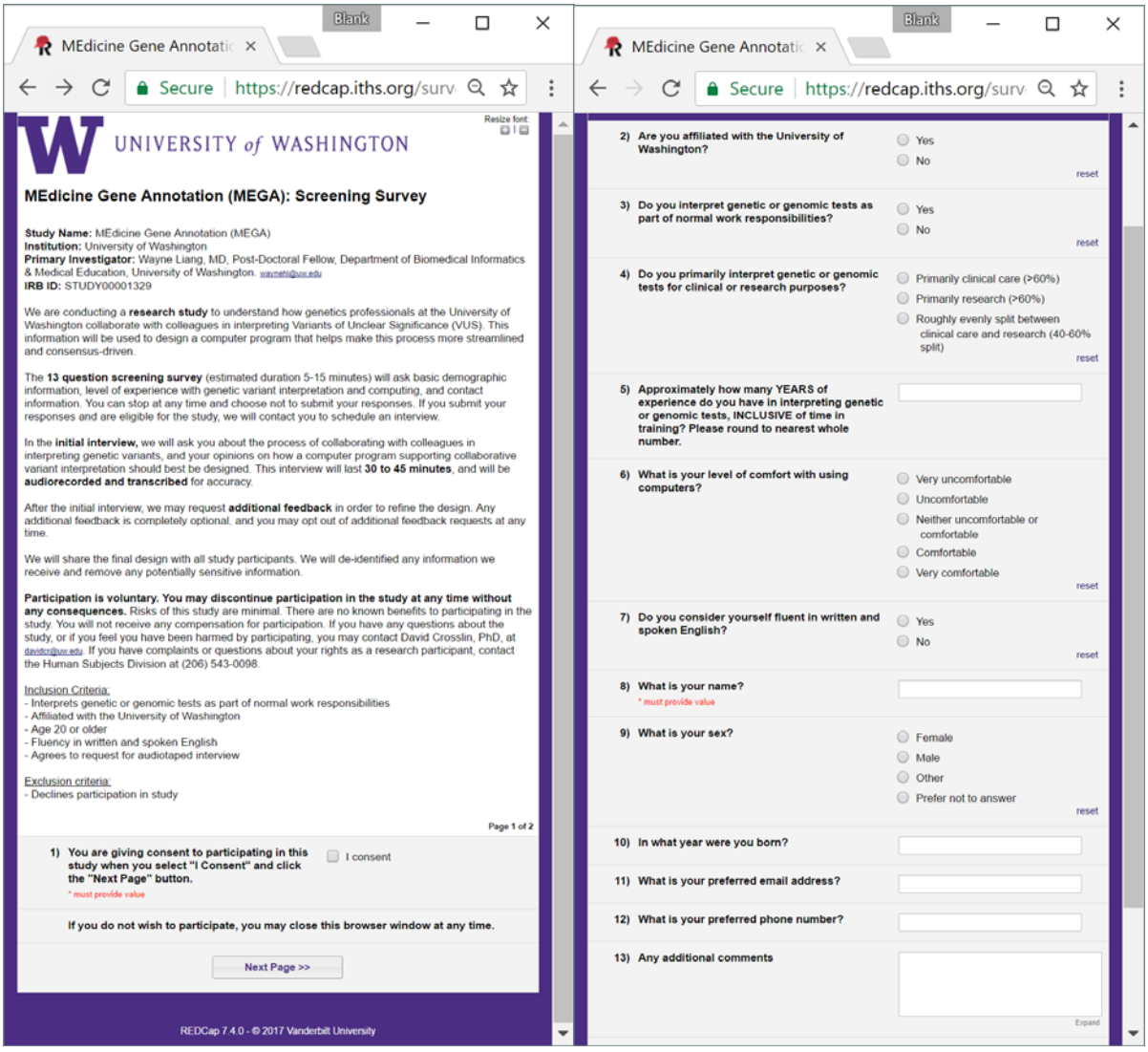

Figure 4.1. Screenshot: study enrollment website.

Due to time and cost limitations of conducting the study, we set our goal enrollment at 4-8 participants, which is within the common accepted range for a user-centered, problem discovery

study[44]. We continued active recruitment until we reached goal enrollment, and closed enrollment after receiving no responses from outstanding recruitment emails.

4.2.3 Participant Interview and Design Feedback Session

The primary investigator conducted a one-on-one semi-structured interview and design feedback session with each participant. We elected to conduct the interview and design feedback session during the same encounter in order to minimize time requirements for busy study participants, and due to time and cost limitations for conducting this study. The interview and design feedback session was conducted in the participant's workspace and lasted less than one hour. The session was audio-recorded and the primary investigator recorded handwritten field notes.

The purpose of the interview was to gather participant data on the following topics:

1. Demographic information;
2. Work roles, experiences and expertise relating to genetic and genomic test interpretation;
3. The processes and tools that the participant uses to perform variant interpretation;
4. The nature and role of collaboration in variant interpretation;
5. Specific user needs and desired functionality for a tool supporting collaborative variant interpretation;

The primary investigator developed a brief interview guide containing prompts for the above topics in advance of the session (See **Table 4.3** for interview guide questions). The primary investigator then used this guide to structure the interview, with specific probing questions developed organically during the interview. Data richness was valued over quantity. The interview session concluded when either sufficient data saturation or the time limit was reached[45].

Following the interview, we immediately proceeded to the design feedback session. The primary investigator described the proposed functionality and workflow of the tool, provided paper printouts of the proposed Delphi-method based workflow and GUI design, and elicited feedback regarding the tool's functionality, workflow, and visual design.

Table 4.3. Semi-Structured Interview Guide

1. **Verify screening survey responses (e.g. demographics, work experience, contact information)**
2. **Ask participant to describe work role, expertise, and experiences directly related to genetic and genomic test interpretation.**
   E.g. "Tell me about your role at the University of Washington as it relates to genomic test interpretation."
3. **Ask participant to describe in detail the process of variant interpretation in his or her lab.**
   E.g. "Can you walk through the process of variant interpretation in your laboratory, from specimen collection to return of results?"
4. **Ask participants to describe the role of inter- or intra-institutional collaboration and consensus making in variant interpretation.**
   E.g. "Can you tell me about individuals who are involved in variant interpretation and their roles?"
5. **Ask participants for design suggestions for a computer program supporting variant interpretation.**
   E.g. "If you had a computer program, what features would you want?"
6. **Review current design concepts and obtain open-ended feedback.**
7. **Ask participants whether they are interested in future contact.**

4.2.4 Transcription & Thematic Coding

Audio-recordings from the sessions were transcribed by a professional transcription service (FlatWorld Solutions Pvt. Ltd.). The transcripts were reviewed and edited for accuracy by the primary investigator. Any identifying or potentially sensitive information were removed by the primary investigator.

We used an analysis method based on grounded theory (GT) methodology in analyzing the session transcripts[46–48]. A grounded theory-based methodology was selected as GT provides a flexible methodological framework for constructing thematic concepts from analysis of interview data, and has been previously used to evaluate a novel clinical genomics informatics tool[49]. The primary investigator reviewed interview field notes immediately following each interview to identify emerging themes. Codes were developed and refined using constant comparison[50]. These codes were then used to perform line-by-line selective coding of interview transcripts using ATLAS.ti versions 7 and 8 for Windows[51]. The coded transcripts were analyzed until we reached a sufficient understanding of participant work processes, user needs, and design requirements. We limited our use of grounded theory methodology to the extent directly applicable to the user-centered design

process; we did not attempt to develop a substantive theory for collaborative variant interpretation, which is outside the scope of this research.

While we considered using multiple coders, we ultimately elected to use a single coder, the primary investigator, for the following reasons:

1. Data collection, coding, and analysis by a single analyst is appropriate and often preferred for grounded theory methodology[46];
2. Additional coders may add richness to the data analysis in certain scenarios, but are of marginal benefit in this study given the limited scope of analysis and the small number of transcripts (n=4);
3. Time and cost limitations of the study.

4.2.5 Design Refinement (Aim 2.C.)

Analysis of participant data and transcripts from the interview and design feedback sessions informed the process of refining the initial design. Participants provided additional feedback and other data via email throughout the design process until a final proposed design was reached.

Chapter 5. RESULTS

5.1 INITIAL DESIGN RESULTS

5.1.1 Typical Workflow for Next Generation Sequencing Test Interpretation (Aim 1.A)

We based our initial sequencing test interpretation workflow design on examples from the literature. Tarczy-Hornoch et al. (2013) described a generalized typical workflow diagram for somatic and tumor sequencing based on a survey of six major academic centers within the Clinical Sequencing Exploratory Research (CSER) consortia (See **Figure 5.1**)[52].

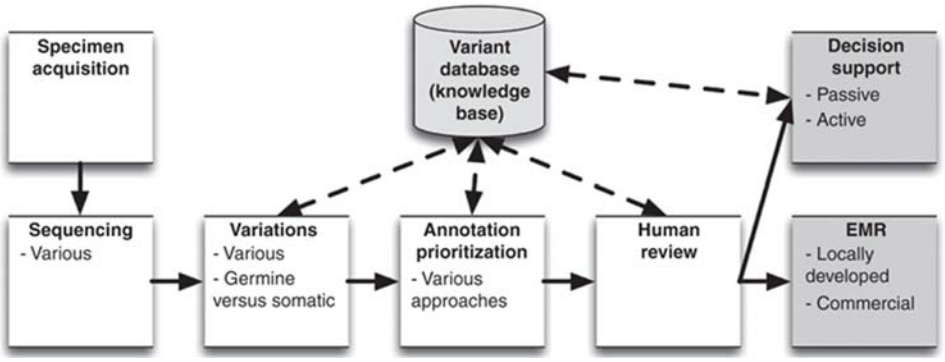

Figure 5.1. Typical Next Generation Sequencing interpretation workflow at Clinical Sequencing Exploratory Research (CSER) sites.
From Tarczy-Hornoch et al. (2013) Genet Med, Figure 1[52].

Good et al. (2014) published another example of a typical workflow for tumor sequencing interpretation (See **Figure 5.2**)[53]:

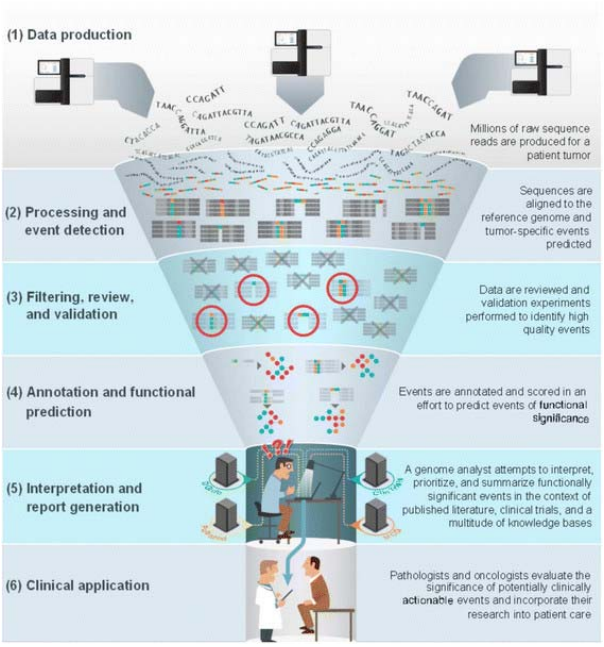

Figure 5.2. Typical cancer genomics Next Generation Sequencing interpretation workflow.

From Good et al. (2014) Genome Biology, Figure 1[53].

At the eMERGE Consortium Meeting 2014, Rehm presented the Partners Healthcare workflow for variant interpretation (See **Figure 5.3**)[23]:

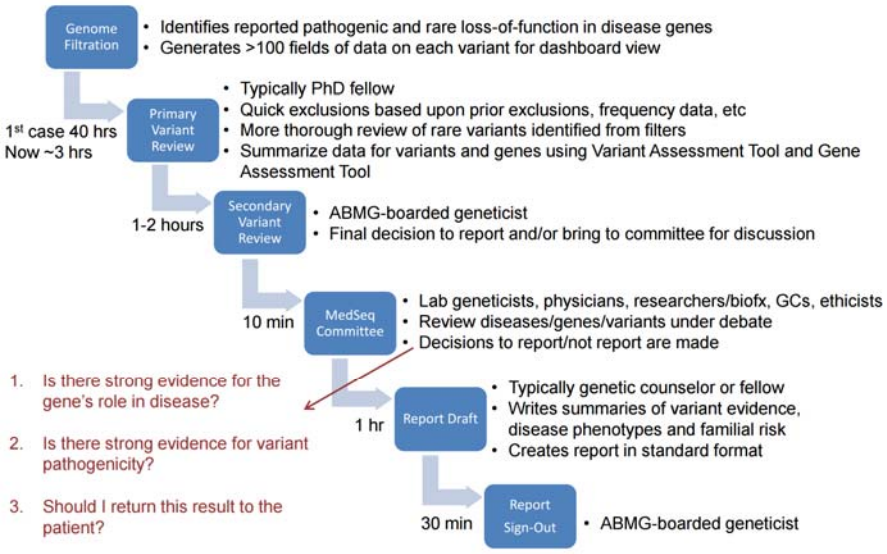

Figure 5.3. Next Generation Sequencing interpretation workflow at Partners Healthcare.

From Rehm (2014) eMERGE Consortium Meeting[23]

These three published workflows are remarkably similar and complementary to each other. *Tarczy-Hornoch et al. (2013)*[52] starts at specimen acquisition, emphasizes informatics-related elements (e.g. variant knowledge bases, electronic medical record (EMR) integration, and decision support), and de-emphasizes details relating to sequence analysis. *Good et al. (2014)*[53] starts at a later step of genomic sequence data production, provides more details relating to sequence analysis, and skip over the role of clinical IT systems. Finally, *Rehm (2014)*[23] provides rich detail on the interpretation workflow, including the roles of laboratory personnel, from variant annotation to return of results.

Based on these published workflows, as well as additional informal input from SMEs, we created the following generalized workflow diagram (See **Figure 5.4**). **Table 5.4** compares our workflow diagram with the three published examples.

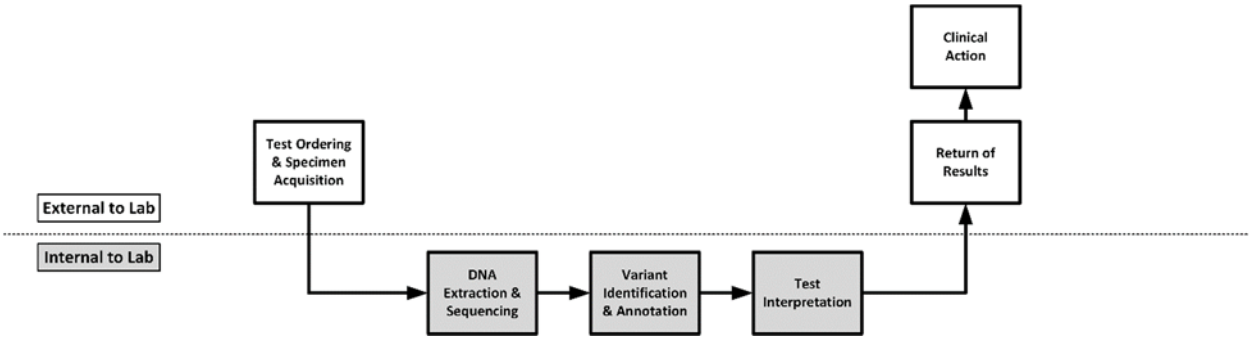

Figure 5.4. Typical Next Generation Sequencing interpretation workflow.
Boxes represent discrete processes in the workflow. Arrows indicate stepwise progression through the workflow. The horizontal line divides processes external (above line) and internal (below line) to the laboratory.

Table 5.4. Comparison of Next Generation Sequencing (NGS) Interpretation Workflow Models
For ease of comparison, numbering was added to the Tarczy-Hornoch et al. (2013)[52] workflow, and one non-linear entity ("Variant database") was excluded. Original numbering was retained for the Good et al. (2014)[53] workflow.

Our Workflow	**Tarczy-Hornoch et al. (2013)**[52]	**Good et al. (2014)**[53]	**Rehm (2014)**[23]
1. Test Ordering & Specimen Acquisition	1. Specimen Acquisition	(Not included)	(Not included)
2. DNA Extraction & Sequencing	2. Sequencing	1. Data Production	(Not included)
3. Variant Identification & Annotation	3. Variations	2. Processing and Event Detection	(Not included)
	4. Annotation Prioritization	3. Filtering, Review, and Validation	1. Genome Filtration
		4. Annotation and Functional Prediction	
4. Test Interpretation	5. Human Review	5. Interpretation and Report Generation	2. Primary Variant Review 3. Secondary Variant Review 4. MedSeq Committee 5. Draft Report
5. Return of Results	6. EMR		6. Report Sign-Out
6. Clinical Action	7. Decision Support	6. Clinical Application	(not included)

5.1.2 Proposed Workflow for Collaborative Variant Interpretation (Aim 1.A)

Next, we focused on developing a proposed workflow supporting collaboration in variant interpretation. We based our workflow design on the CSER Variant Bake-off study as our use case (See *Section 1.3*)[7]. This study demonstrated that the use of a consensus process among cross-institutional collaborators in variant interpretation increases concordance in interpretation results. While the CSER Bake-off study demonstrated a level of inter-laboratory collaboration that is unusual in everyday practice, we posit that a tool supporting this use case may be applicable for

smaller scale inter- and intra-laboratory collaborations in variant interpretation. Therefore, our initial collaborative workflow design was based on supporting the CSER Bake-off workflow.

We posit that a modified Delphi Method may be applicable to consensus making in variant interpretation. The Delphi Method is a well-established structured communication process designed to collect and consolidate the opinions of multiple experts in order to facilitate consensus making, group problem solving, forecasting, and decision-making[54,55]. This method was initially developed in 1967 to support informed intuitive judgement in long-rang forecasting by experts, and variants of the process has since been applied to a wide range of domains and problems involving multiple expert opinions. Through a series of questionnaires, the opinions of individual experts are obtained, merged, and returned to participants for review and feedback. Progression through multiple rounds has been observed to lead to a convergence of opinions. Collaborative variant interpretation has similarities to other consensus processes involving multiple experts. Given the Delphi Method's flexible and broad applicability for various consensus processes, we used the Delphi Method as a starting point in creating a proposed workflow supporting collaborative variant interpretation.

In our proposed Delphi Method-based workflow, individual geneticists served as experts and variant interpretations represented expert opinions. Our proposed workflow consists of three Delphi Method rounds. In Rounds 1, the tool solicits and captures variant interpretations from individual geneticists, automatically synthesizes a group interpretation, and returns group results with concordance statistics to all participants. Variant interpretation entries consist of a Likert scale of classifications (P, LP, VUS, LB, and B), a self-assessed confidence score, and a free text box to enter classification rationale. In order to avoid biasing interpretations, the participants would be blinded to each other's identities and ground results are returned in anonymized form. Round 2 is identical to Round 1. In Round 3, participants receive group results, are unblinded to each other's identities, then participate in a collaborative consensus meeting supported by our tool to discuss discordant findings and finalize the group consensus classification. The consensus classification and any minority opinions are captured by the tool and saved in an internal variant knowledge base. See **Figure 5.5** for the Delphi Method-based consensus process and **Figure 5.6** for the Delphi Method-based collaborative workflow.

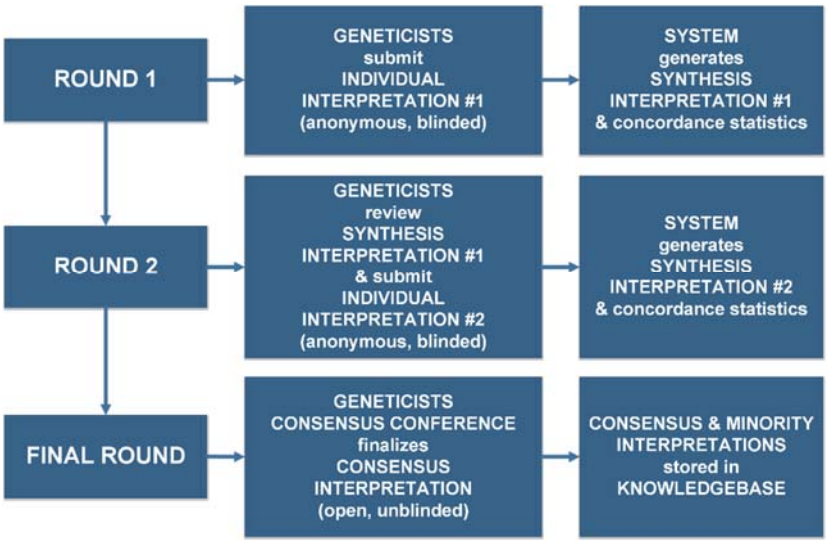

Figure 5.5. Delphi Method-based consensus process for variant interpretation.

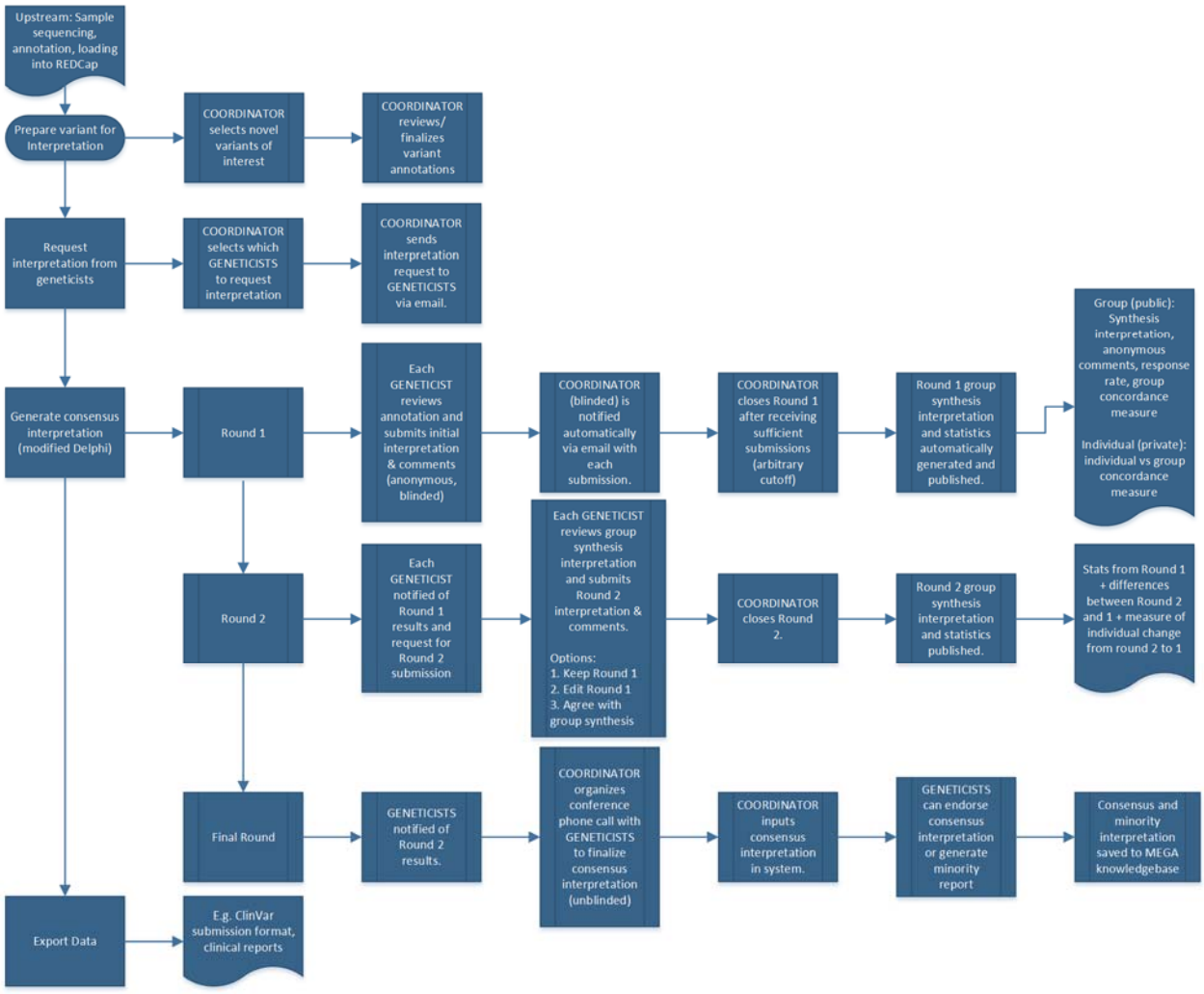

Figure 5.6. Proposed Delphi Method-based variant interpretation workflow.

5.1.3 Initial Graphical User Interface Design (Aim 1.B)

We utilized an existing prototype tool supporting capture of variant interpretation as the starting point of our initial GUI design. The existing tool is a customized data entry form within the REDCap project management webpage (See **Section 5.1.4** for additional details of the technology architecture)[40]. The form displays editable text boxes containing both annotation data, and empty text boxes to capture variant interpretation submissions. Geneticists interact with the form by logging into the institutional REDCap project management website, selecting the appropriate project, selecting the form, entering interpretation data, and submitting the form. See **Figure 5.7** for a screenshot of the existing GUI.

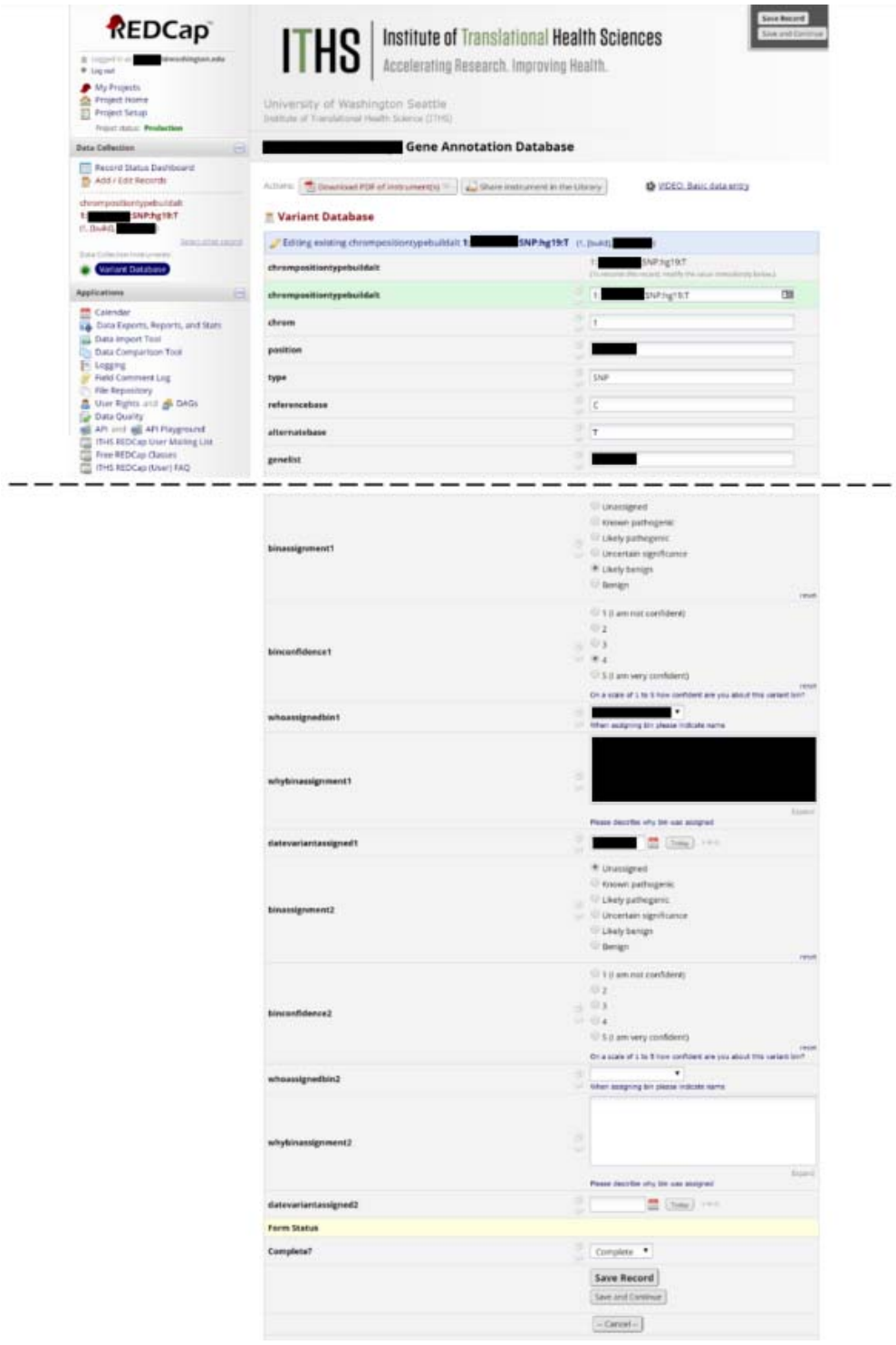

Figure 5.7. Screenshot: GUI of existing REDCap form-based prototype tool.

The dotted line indicates visual elements which were collapsed due to space constraints.

The black boxes indicate redacted fields.

We obtained informal input from SMEs familiar with this tool. Users indicated that the user interface provided adequate data required for variant interpretation. Users provided positive feedback on the tools' ability to captured structured annotation and variant information. Users also indicated high acceptance of a web-based interface. However, users felt that the existing tool had significant design limitations that affected its usability and usefulness. The interface was described as not aesthetically pleasing and "clunky". Variant annotation data were displayed as a flat, vertical list of survey response fields, resulting in a long webpage requiring excessive scrolling. Each data element was labeled using the associated database variable name code and was not human-readable. Annotation data were displayed as editable data fields, which introduces the potential for editing errors. Users provided negative feedback on the multiple steps required to access the form and on poor integration with existing workflow. Users desired easy access to the tool and increased integration with their existing workflow.

In the next iteration of our GUI design, we sought to incorporate informal feedback while making only minor modification to the existing technology infrastructure. In this iteration, we implemented additional functionality native to REDCap, including data piping within user surveys. The resulting GUI is a customized REDCap-generated web-based survey. At top, annotation data are pulled in automatically and displayed as non-editable text, and the lower portion of the survey contains data fields to capture interpretation data entered by the geneticist. The data piping syntax is blended with a limited set of HTML syntax to introduce additional elements such as descriptive text and basic formatting See **Figure 5.8** for a screenshot of this GUI design iteration.

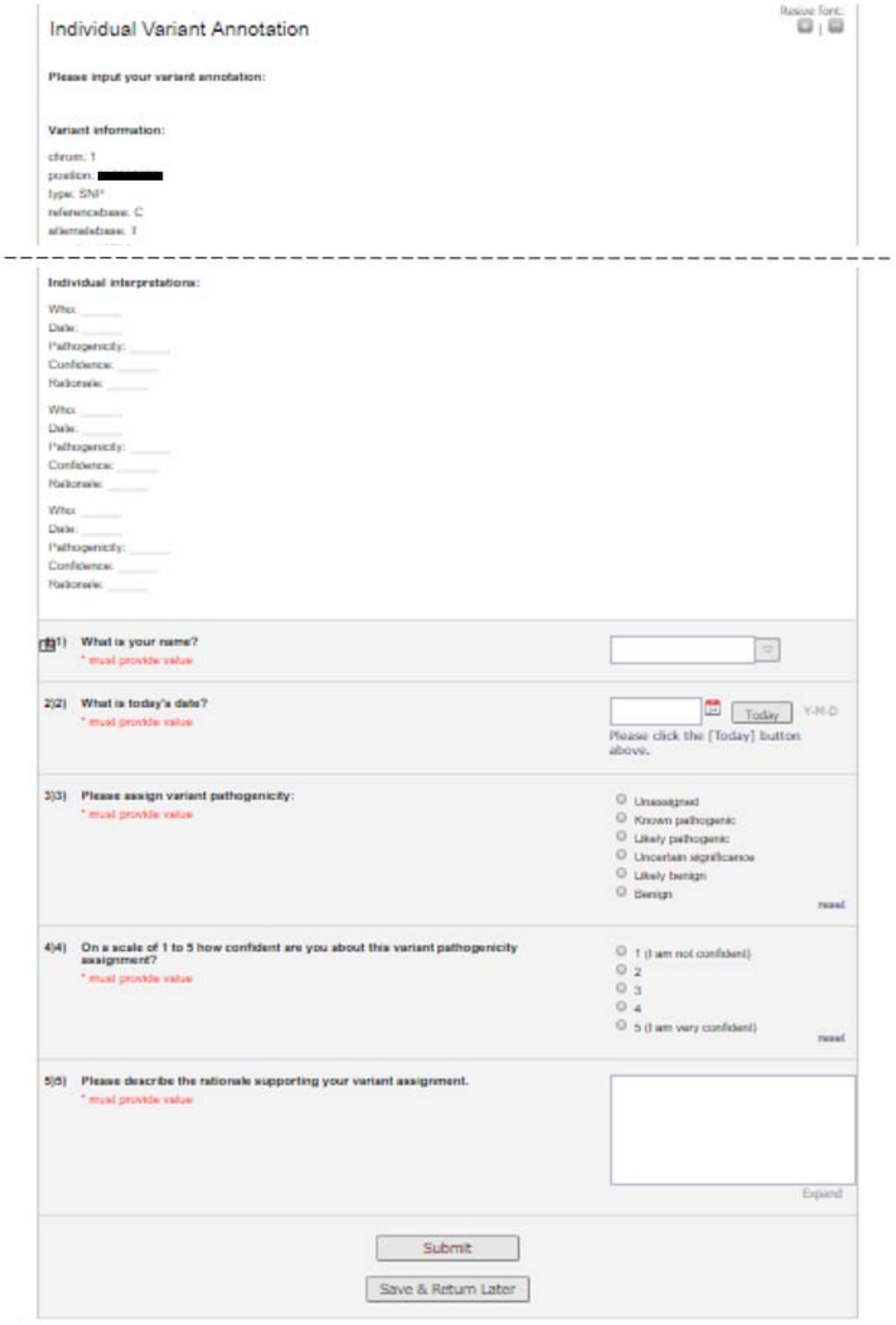

Figure 5.8. Screenshot: GUI design iteration using extended REDCap functionality.

The dotted line indicates visual elements that were collapsed due to space constraints. The black boxes indicate redacted elements of this example screenshot.

This GUI design iteration provides several advantages over the existing tool. REDCap's survey functionality allows the system to request interpretations from select experts by sending an automatically generated email message with a customized link to that individual's email address. The survey system can be configured such that the user does not need to authenticate with a username and password in order to submit an interpretation. This eliminates the need for the user to log into the REDCap website, navigate to the specific project and form, and interact with the project management page of the REDCap project. This format also provides additional security for existing data, because it eliminates the possibility of the user accessing or modifying other variant entries, or unintentionally changing the annotations of an existing variant. Basic HTML syntax support allows for more flexibility in how data are displayed, including the ability to include additional text (e.g. informative labels, instructions).

The REDCap survey-based design is an improvement from the initial design. However, functional limitations of the survey tool make this design infeasible. Due to limited HTML support, data elements must be presented as a long vertical list. While the HTML syntax allows for embedding certain multimedia objects, it does not support embedding external webpages or inline frames to allow for advance, interactive features. Additionally, data elements cannot be piped into HTML syntax, thereby preventing linking to context-specific external content using a unique variant identifier (e.g. links to external variant databases). Finally, REDCap database architecture is flat, without native support for linking multiple tables using external keys. Additionally, the survey functionality does not natively support allowing multiple individuals submit survey responses for the same data entry. This database structure limits the ability of using the native survey tool to obtain opinions from multiple experts (i.e. survey respondents) for the same variant, or separating variant data from interpretation data.

For the following design iteration, we explored the use of custom hooks, plugins, and the REDCap API. This allows us to write completely custom software, while still being able to take advantage of REDCap as the backend database. We are then able to simulate a multi-table architecture by creating multiple REDCap projects controlled programmatically using the API, in order to separate entries by record type. The resulting design iteration borrows from the previous iteration's layout,

with annotation data presented at top and data entry fields at bottom. The upper portion including the ability to include multimedia objects, such as data visualizations. Data can be organized into different tabs by category, and interpretation data by multiple submitters can be presented in the same interface. Inline frames allow for scrolling of annotation data (upper portion) while preserving the user's access to the interpretation data entry fields (lower portion). We felt that this design was adequate to move forward for user feedback in the user centered-design phase of the study, and therefore concluded the initial GUI design work at this point. See **Figure 5.9** for a mockup of this interface.

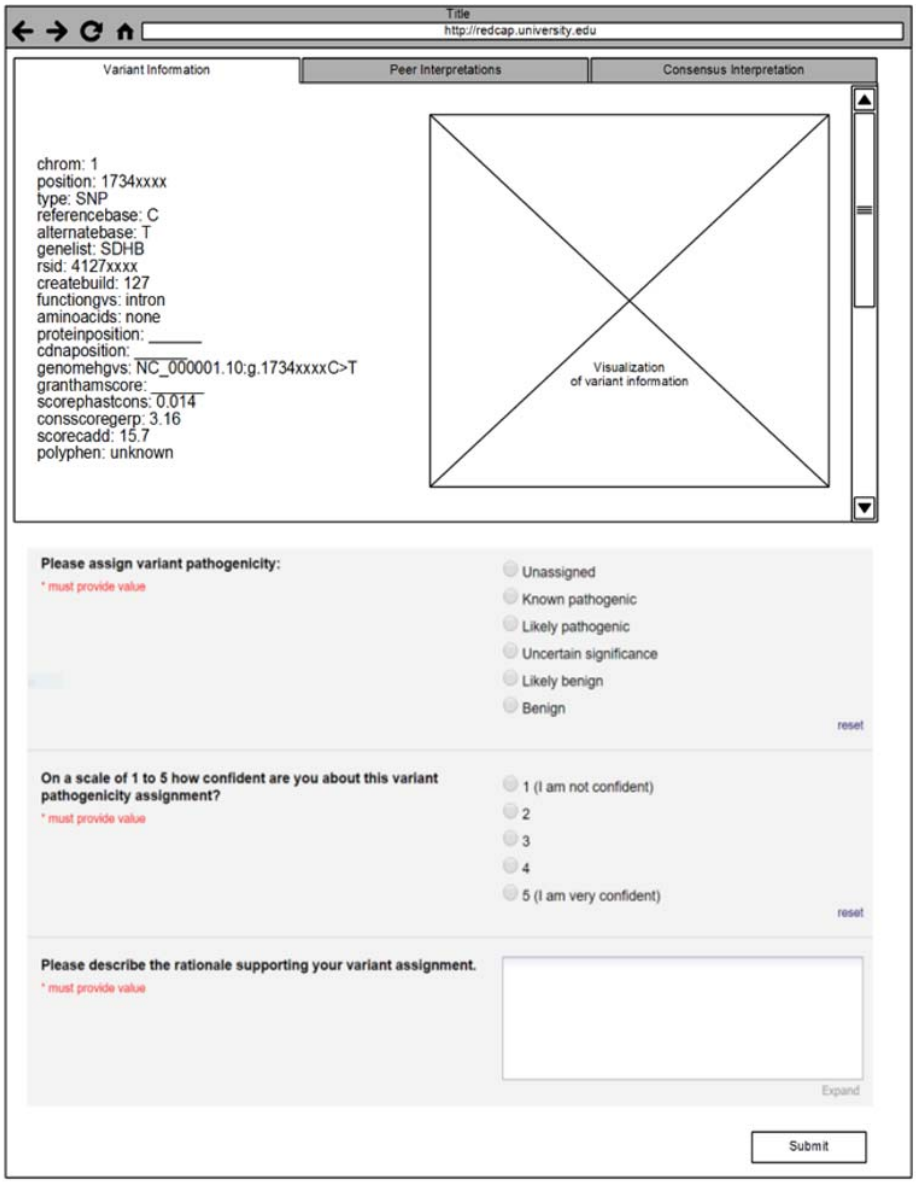

Figure 5.9. Mockup: GUI design iteration using REDCap API access.

5.1.4 Initial Technology Architecture Design (Aim 1.B)

We utilized the existing prototype discussed in *Section 5.1.3* as the starting point of the technology architecture design. The existing prototype was developed as a REDCap project with a custom form. REDCap is a robust research database and management tool that has been successfully used to manage large volumes of data in a secure, HIPAA-compliant manner. REDCap has been widely adopted at academic centers and is actively being developed, which makes it an attractive platform for building opening available tools for broad dissemination. The existing tool uses custom R Statistical Software[56] scripts interacting with the REDCap API to import variant annotation from VCF files into the REDCap project database. A unique variant code name is generated computationally using the method described in Crosslin et al. (2015)[57]. A custom form within the REDCap project management web interface is used to capture interpretation data (See **Figure 5.7**). Variant data is stored in a flat REDCap database; each row corresponds to a variant, and each column corresponds to an annotation or interpretation data element. Data can be extracted from REDCap by R scripts interacting with the API. R scripts can then be used to transform the data in order to generate reports for downstream systems such as a clinical IT system or ClinVar. See **Figure 5.10** for an illustration of the technology architecture.

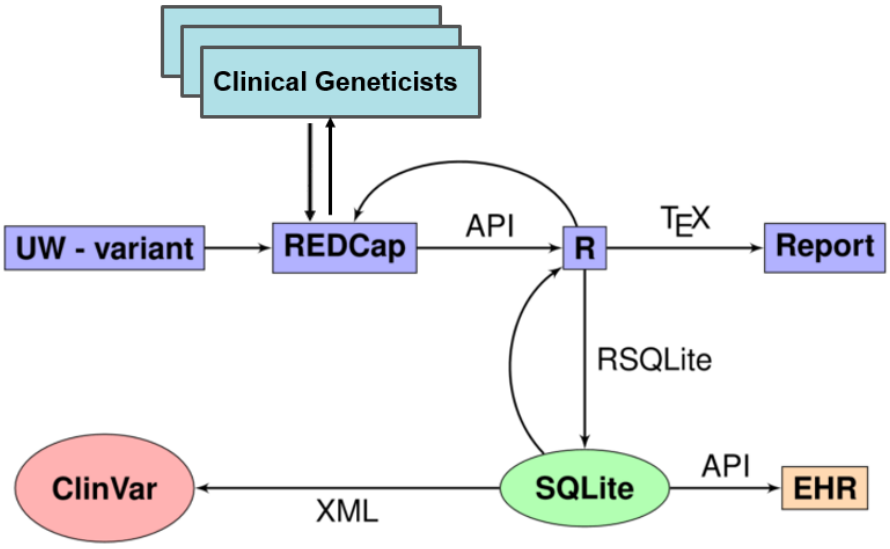

Figure 5.10. Technology architecture of the existing prototype tool.

For our proposed design, we elected to retain the existing technology architecture with minor modifications. We retained REDCap as the database backend, along with the R scripts to interact

with the REDCap API for data import, manipulation, and export. However, instead of using only REDCap native functionality, we plan to use advanced REDCap features such as hooks, plugins, and/or API to create custom software that can interact with REDCap as the backend database.

5.2 User-Centered Design Results

5.2.1 Participant Interviews and User-Centered Design Feedback Sessions (Aim 2)

We invited six SMEs via email to participate in the user-centered design phase of the research. Four participants provided consent and enrolled successfully (response rate = 67%). Each participant has a doctoral degree, is a University of Washington faculty member, and performs genetic test interpretation as a part of clinical and/or research responsibilities. See Table 5.5 for additional participant characteristics.

Table 5.5. User-Centered Design Participants. (N=4)

Female	1 (25%)
Age in years (average)	55 (Range 40s to 70s)
Years of experience in genetic test interpretation (average)	14 (Range 8 to 20)
Primary role (multiple selections possible)	
Laboratory-based	3 (75%)
Direct patient care	2 (50%)
Genetic interpretation for clinical or research purpose?	
Primarily research (>60%)	1 (25%)
Roughly evenly split (40-60%)	2 (50%)
Primarily clinical (>60%)	1 (25%)
Comfort with using computers	
Very comfortable	1 (25%)
Comfortable	3 (75)
Neither uncomfortable or comfortable	0 (0%)
Uncomfortable	0 (0%)
Very uncomfortable	0 (0%)

All four participants successfully completed the interview and design feedback session. Total interview duration ranged from 42 minutes 29 seconds to 51 minutes 43 seconds. Transcribed

word count ranged from 6,347 to 7,471. We developed 20 codes from a thematic analysis of interview field notes, and then used these codes to annotate interview transcripts. Each quotation may be associated with more than one code. See **Table 5.6** for code definitions and **Figure 5.11** for coded concept frequencies.

Table 5.6. Code Names, Groups, Definitions, and Occurrences (n) in Interview Transcripts

Codes by Code Groups	Definitions	n
Design Feedback		
feedback-visual	Specific feedback on the initial design's visual elements (e.g. graphical user interface).	1
feedback-workflow	Specific feedback on the initial design's workflow.	23
Workflow and Genomic Science		
workflow-gen-clinicalhistory	Phenotype and clinical history in genomic test interpretation.	9
workflow-gen-guidelines	Genomic test interpretation guidelines.	23
workflow-gen-knowledge	State of current genomic knowledge.	11
workflow-gen-uncertainty	Uncertainty in genomic test interpretation.	14
Workflow and People		
workflow-people-collaboration	Collaboration in variant interpretation.	24
workflow-people-consensus	Consensus-making in variant interpretation.	10
workflow-people-roles	People's functional roles in variant interpretation.	20
Workflow and Process		
workflow-process-override/exit	The ability to over-ride or exit a process.	2
workflow-process-usecases	Use cases for workflow processes.	2
workflow-process-variantselection	The process of filtering and selecting variants.	15
Workflow and Resources		
workflow-resources-experts	Human experts as a knowledge resource.	14
workflow-resources-info	Non-human information resources.	21
workflow-resources-material	"Material" resources (e.g. money, people, time, physical space).	6
workflow-resources-varknowledgemgmt	Managing variant knowledge.	17
Workflow and Technology		
workflow-tech-custom/standard	Customization versus standardization in pipeline technology.	7
workflow-tech-filetypes	File types and formats.	7
workflow-tech-interop/integrate	Technological interoperability and integration.	13
workflow-tech-pipeline	Variant analytical pipeline technology.	13

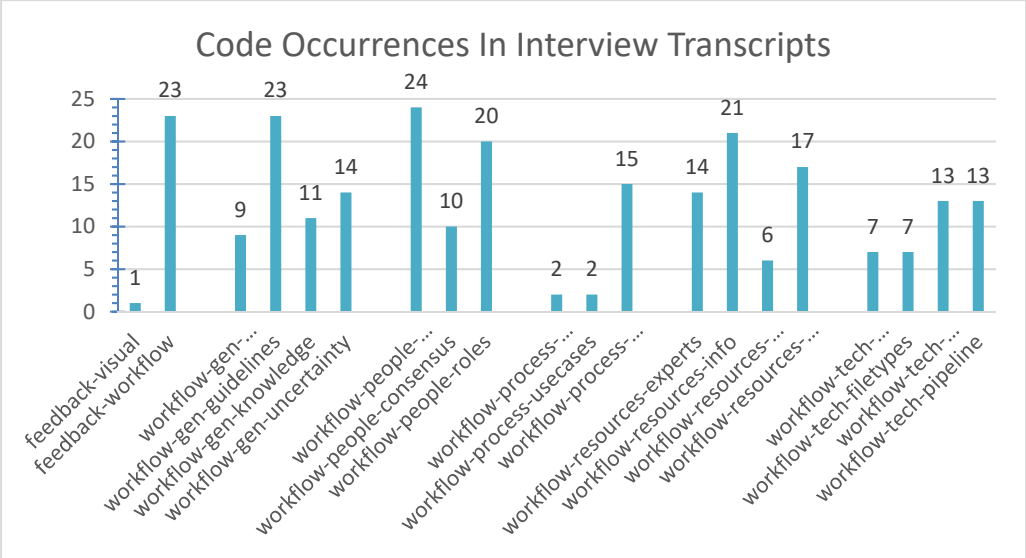

Figure 5.11. Code occurrences in interview transcripts.

5.2.2 Variant Interpretation Workflow with Collaborative Roles (Aims 2.A)

The four participants had different roles and work environments, providing diversity and richness to our interview data. Each participant leads distinct clinical and/or research programs and have active collaborations with each other. Participants A and B are medical directors for two different sequencing laboratories, with completely independent analytic and interpretation pipelines. Participant A's laboratory performs whole exome sequencing, while Participant B's laboratory performs panel sequencing. Participant C is both a clinician and a laboratory medical director. His laboratory utilizes Participant A's pipeline to perform DNA extraction, sequencing, variant identification, and annotation, but his laboratory handles the interpretation. Participant D is a clinician who also serves as an expert consultant for both Participants A and B's laboratories in interpreting variants for clinical and research purposes. All four participants collaborate frequently with internal and external collaborators.

We used participant interview data to confirm our original workflow model (See **Figure 5.4**), then extended it to include collaborative roles (See **Figure 5.12**). In our model, the clinician and patient decide to order the test, and the patient provides the specimen. A laboratory intake personnel (in this case, a laboratory-based genetic counselor), obtains all of the phenotypic information from the chart, referring provider, and potentially the patient. The laboratory bioinformatics technician

performs the sequencing, variant identification, and annotation. The bioinformatics technician and genetic counselor may perform another review. Finally, the entire case package, which includes both phenotypic and annotation information, goes to a Senior Molecular Geneticist. In both participant A and B's laboratories, at least 2 Senior Molecular Geneticists always reviews each case, with one assigned to be the report author. A portion of the cases is discussed at a regularly scheduled laboratory consensus conference. External experts may be consulted to provide relevant expertise (not pictured). The report author then finalizes the interpretation, writes the reports, signs it, and sends it to the ordering provider. The ordering provider then receives the report and determine what clinical actions, if any, to take.

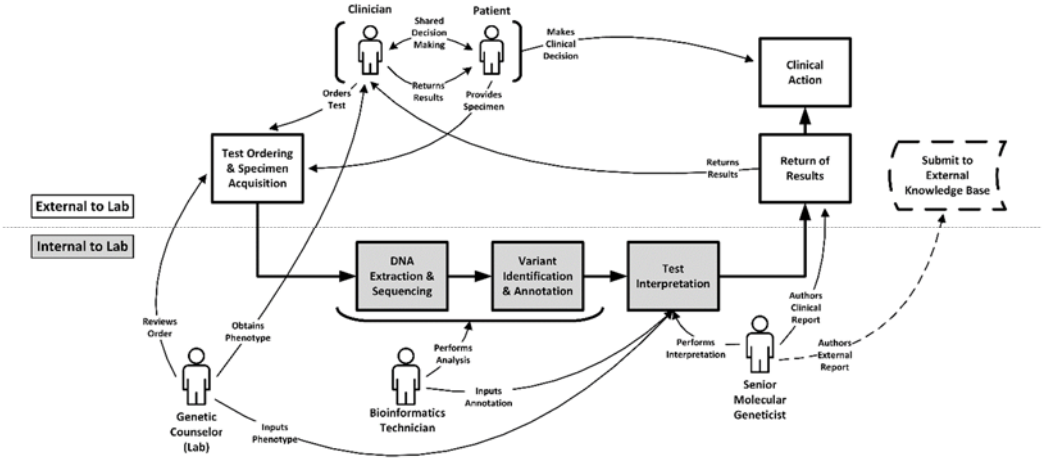

Figure 5.12. Variant interpretation workflow with collaborative roles.

The human figures represent discrete roles performed by one or more individuals. Curved arrows represent actions performed by each role. Dashed arrow and cylinder represent an optional process.

5.2.3 Annotation File Format and Interpretation Report Authoring

Participant A's analytic pipeline outputs an annotation file in a machine and human-readable Microsoft Excel spreadsheet file. We were not able to obtain a sample annotation file for artifact analysis, but Participants A, C, and D were able to describe this file in detail. The layout is a data dense, unformatted table of annotation data. This file is often printed out and reviewed manually on paper, and notes are written by hand in the margins. Reviewers collaborate by comparing their printed annotation file with handwritten notes. After the review is concluded, the report author manually enters interpretation information and relevant supportive evidence into a laboratory

report authoring software. Data elements cannot be imported from the annotation file into the report authoring software, so data transfer is performed manually using copy and paste. Then, the report is entered into the laboratory reporting system.

Participant B's analytic pipeline also outputs an annotation file in a machine and human-readable Microsoft Excel spreadsheet file. Participant B described the file in detail and also provided a sample annotation file to the design team for artifact analysis, without permission to distribute or reproduce. The file contains 11 sheets, with each sheet containing a data dense, unformatted table of raw annotation data generated by one component of the analytic pipeline. The file is reviewed electronically and notations are electronically added to the file. The secondary reviewers will email their annotate file to the designated report author. After the review is concluded, the report author manually enters interpretation information into a Microsoft Word document and may copy and pasting data elements from the annotation file as supportive evidence. The author may search an electronic shared folder containing old, interesting cases to copy and paste previously written interpretation text for a previously encountered variant. The report is then entered into the laboratory reporting software.

5.2.4 User Needs & Desired Functionality (Aims 2.A)

Through participant interviews, we uncovered user needs and desired functionality relating to five areas: genomic science, people, processes, resources, and technology. Our findings are summarized in **Table 5.7**.

Table 5.7. Desired Functionality for a Tool Supporting Variant Interpretation

Genome Science (Content Area)

- Support capture and display of phenotypic information.
- Support decision support based on the ACMG/AMP Guidelines.
- Support searching and retrieving external evidence sources (e.g. PubMed, ClinVar).
- Support request and capture of external expert opinion.
- Support storage and retrieval of variant knowledge.
- Support sharing of variant knowledge.

People

- Support collaboration among laboratory personnel and ad-hoc external collaborators based on roles.

Process

- Integrate with existing laboratory workflow.
- Support the existing intra-laboratory consensus process.
- Provide users flexibility and control to over-ride or exit from processes.
- Support report authoring.

Resources

- Increase efficiency in the use of time, labor, and other material resources in variant interpretation.
- Support data reuse to decrease copy-and-pasting and other low level clerical work.

Technology

- Support automated assignment of predicted ACMG/AMP 2015 evidence codes.
- Integrate with existing pipeline technology.
- Use standards-based technologies where available and appropriate.
- Support customization.

5.2.5 Participant Feedback on Delphi Method-based Proposed Workflow (Aim 2.B)

We solicited and obtained feedback on the Delphi Method-based proposed workflow (See **Figure 5.5)** during the one-on-one design feedback session.

Participants were very enthusiastic about having a tool that could capture variant knowledge in an internal knowledge base, and being able to export that data to downstream sources (e.g. clinical information systems, ClinVar). Participants would like such a system to be searchable and provide analytics data such as allele frequencies. Participants requested that an ACMG/AMP calculator be included in the program, and would ideally like the program to semi-automate the classification

process. Participants liked the idea of a tool that can pull together different reviewers' interpretations in one place, to facilitate collaboration and consensus making. Participants reported frustration with having to copy and paste or manually type annotation data into interpretation reports, and would welcome a tool that helps with automating information flow at this step.

Participants generally had negative opinions about applying the Delphi Method process to variant interpretation. Participants felt that a formal consensus process might be useful in multi-laboratory collaboration scenarios like the CSER Variant Bake-off study, but may be excessive and unnecessary for routine cases performed in individual laboratories. All participants were concerned about the time and additional work required to go through three Delphi Rounds. Participants requested that the Delphi Method component be significantly streamlined or removed. Participants also were concerned about the reliability of a machine-generated synthesis interpretation, and felt that concordance statistics may not provide meaningful information with few reviewers.

Participants had mixed feelings about being blinded to the identity of other reviewers. Participants suggested that blinding could potentially reduce bias in variant interpretation, but felt that the effect of social pressure is likely low, and in a single laboratory, attempts at blinding would not be very effective.

5.2.6 Participant Feedback on Initial GUI Designs (Aim 2.B)

We solicited and obtained feedback on the initial GUI design (See **Figure 5.9**) during the one-on-one design feedback session.

Overall, participants provided positive feedback on the visual layout. They felt that the visual design was clean, visually appealing, logical and easy to understand. Participants felt comfortable interacting with a web-based interface to perform variant interpretation. Participants agreed with visually separating the display of annotation data from the interpretation entry interface. Overall, the visual design was acceptable.

Participants provided negative feedback on specific functionality elements that directly affect the visual design. Multiple participants felt that having a separate question asking users to rate their level of confidence was unnecessary, since the pathogenicity classifications already include levels of certainty. One participant suggested that instead of having the reviewer enter supporting evidence for their classification opinion in a free text box, the reviewer should be able to select the appropriate evidence from the annotations and have that be automatically recorded.

Participants provided mixed feedback on how they would like annotation data presented. One participant strongly favored having data presented in a data-dense tabular format, similar to the current annotation spreadsheets in use, and requested that visualizations be eliminated as they would just take up room on the screen. Others participants felt that having visualizations of annotation information could be helpful, but were unsure what those visualizations may be.

5.2.7 User-Centered Design Conclusions

In conclusion, the user-centered design process generated rich data on existing workflows, as well as feedback on initial designs. User feedback on the initial designs focused much more heavily on the tool's workflow and functionality than on the visual layout, which was of secondary concern to the users. Participants desired a tool that saves time and work, provides decision support for the ACMG/AMP guidelines, captures variant knowledge, supports collaboration, supports data export, and is integrated into existing workflow. Participants were open to processes supporting consensus-making, but felt that the Delphi Method was excessive and unnecessary. Users felt that having a separate confidence rating system was unnecessary. Visually, users agreed with having separate panels for displaying annotation data and entering interpretation data. Otherwise, the visual design was acceptable.

Data gathered from the user-centered design process were used to iteratively revise designs to better support the needs of the end users.

5.3 FINAL DESIGN RESULTS

5.3.1 Final Workflow Design (Aims 2.C)

In this final iteration, MEGA is designed to be a tool that supports role-base collaboration, variant classification decision support, variant knowledge management, and knowledge-sharing.

MEGA is organized around three data types: 1. Patients; 2. Cases; and 3. Variants.

1. Patient files contain phenotypic and other clinical information necessary for case interpretation.
2. Case files correspond to the ordered test and contains test data and reports. Multiple cases can be associated with the same patient.
3. Variant files correspond to a specific sequence variant. Each variant file contains annotation data pulled from external and internal sources, such as local allele frequencies, and can be associated with multiple cases and patients.

User access privileges within MEGA are based on user groups. Users can be assigned to multiple user groups. The default user groups within MEGA include: 1. Administrators; 2. Editors; 3. Consultants; and 4. Directors.

1. Administrators can manage administrative tasks, such as user access and program preferences, but do not have access to clinical or variant data. This group may include informational technology support specialists and other laboratory personnel designated by the medical director.
2. Editors can browse all databases, edit annotation data, submit interpretation opinions, assign tasks, and edit draft reports. This group includes laboratory personnel involved in generating data and preparing a case for interpretation, such as case intake specialists and bioinformatics technicians.
3. Consultants can view annotation data, submit individual interpretations, and edit draft reports. However, consultants cannot browser databases, finalize consensus interpretations, or author finalized reports. Consultants are eligible to be assigned as a Reviewer for a case, but not as a Report Author. This group includes external expert collaborators who are invited to provide expert opinion on an ad hoc basis.

4. Directors have all privileges, including finalize consensus interpretations and authoring reports. Directors can be assigned as a Reviewer or a Report Author for a case. This group includes only senior-level personnel such as the medical director, or associate or assistant medical directors.

MEGA integrates with and supports the existing variant interpretation workflow. In the example workflow (**Figure 5.13**), the laboratory intake specialist, who belongs to the Editor user group, enters the patient's phenotypic information into the associated patient file within MEGA. The intake specialist assigns a Director-level user as the case's "Report Author" (required), and assigns additional Consultant- or Director-level users as "Reviewers" (optional). These task assignments can be changed at any time. The bioinformatics technician, who belongs to the Editor user group, performs the sequence analysis using the existing pipeline, then imports the annotations into the associated case file within MEGA. The bioinformatics technician, or another Editor, reviews the case file, runs the classification prediction function, drafts variant interpretations and the report (optional), attaches any questions for the reviewers (optional), then flags the case as ready for review. Review requests are sent via email to the Report Author and Reviewers, who then can log into MEGA to review the case and submit individual interpretations. The Report author is able to see everyone else's work, and pulls their work into their interpretation text and report. The Report Author can close the review process at any time by submitting finalized variant interpretations and finalizing the report, which are saved into the MEGA knowledge base. MEGA can then export human and machine-readable reports into downstream systems, including clinical IT systems and ClinVar.

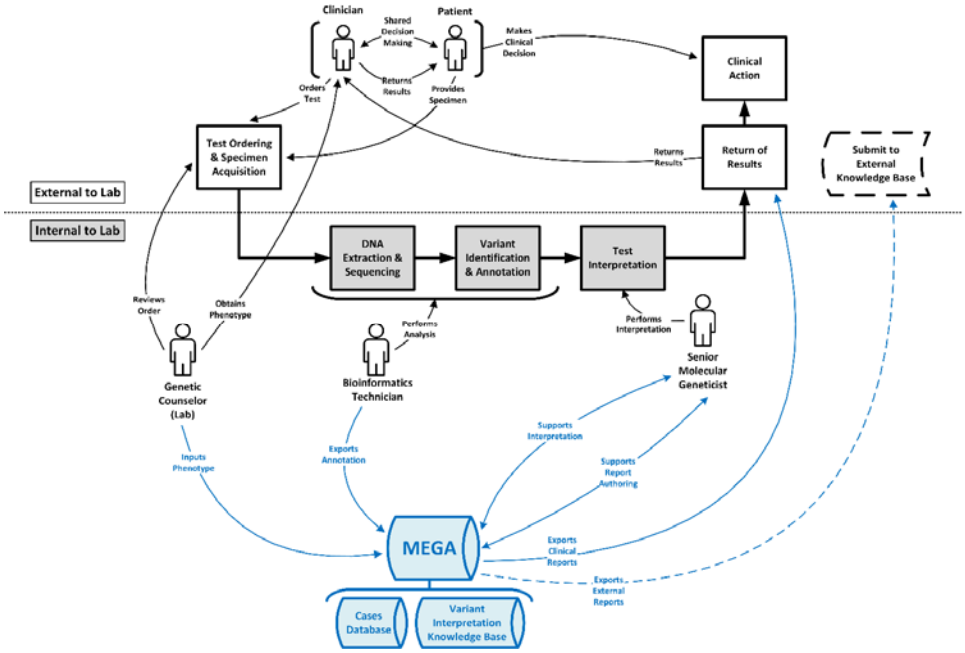

Figure 5.13. Variant interpretation workflow with MEGA integration.

Blue elements indicate MEGA components and its interactions with the existing workflow.

A frequently requested feature is decision support implementing the ACMP/AMP interpretation guidelines. Specifically, participant requested a system that 1) automatically assigns evidence codes, and 2) automatically calculates the pathogenicity interpretation. Instead of developing our own rules engine, we have elected to integrate InterVar (See *Section 2.1.2*) as MEGA's rules and prediction engine. MEGA will pass the output of the sequencing pipeline into InterVar, and ingest the predicted evidence codes and classifications returned by InterVar. Advantages of integrating InterVar include not duplicating prior work, supporting adoption of an open source tool, and freeing up resources to focus on other features. Disadvantages of this approach include not having complete control of MEGA's rules engine, limited capabilities (currently InterVar is only able to predict 18 of 28 evidence codes), and concerns about the tool's trustworthiness. We plan to utilize InterVar to predict evidence codes and classifications, which then will need to be confirmed by the Report Author prior to submission. We believe the benefits of time saved by using InterVar outweigh the costs of having to review predicted evidence codes for accuracy. If InterVar does not demonstrate adequate performance during testing, we will remove the evidence code prediction feature and replace it with a pathogenicity calculator.

5.3.2 Final GUI Design (Aims 2.C)

Like prior iterations, the final GUI design is a web-based interface. The layout retains the visuals separation of annotation data and the interpretation interface into separate panels. However, the layout has changed significantly from the prior iteration to support workflow and functionality changes based on user feedback (See **Figure 5.9**). The interface for variant interpretation is divided vertically into three sections: 1) Header; 2) Interpretation Panel; and 3) Evidence Browser. We will discuss each section separately from top to bottom. See **Figure 5.14** for the GUI layout and **Figure 5.15** for a mockup of the final GUI Design.

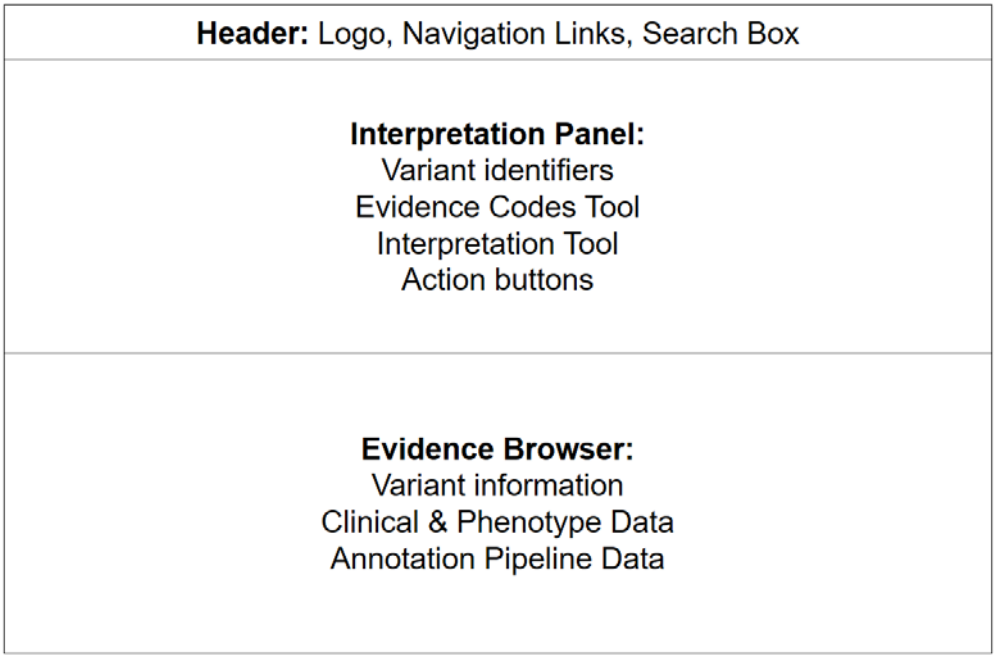

Figure 5.14. Final design: variant interpretation GUI layout.

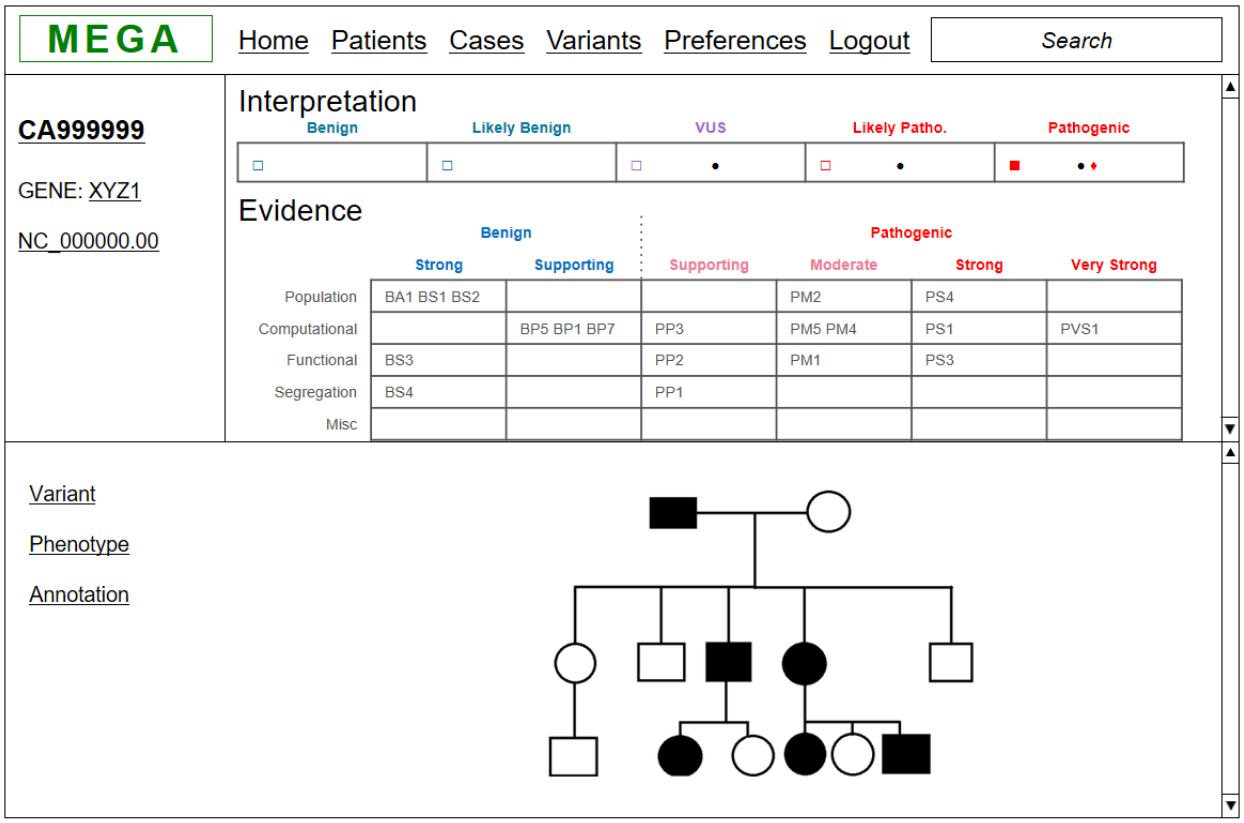

Figure 5.15. Final design: variant interpretation GUI mockup

The uppermost section is the Header, which contains the MEGA tool name and logo, top level navigation links, and a universal search box. The logo orients the user to the context, the navigation links allow the user to access other tool functions, and the search box allows the user to search text found in patient, case, and variant files. This header is present unchanged throughout all pages of the tool to provide consistency in top-level navigation.

The middle section is the Interpretation Panel. It remains static and viewable even while the user scrolls through data in the bottom panel, but is collapsible if desire (e,g. if the bottom panel needs to expand). It contains basic variant identifiers, the Evidence Codes Tool, the Interpretation Tool, and action buttons. Basic variant identifiers include ClinGen Allele Registry ID, gene name, and HGVS nomenclature. Below the variant identifies is the Evidence Codes Tool, which is an interface to interact with ACMG/AMP Guidelines evidence codes. The layout of the Evidence Codes Tool is based on Table 5 of *Richards et al. (2015)*[21], with rows indicating evidence type, and columns indicating evidence strength (See **Figure 6.4** for the original table). Columns are

color coded with 3 colors: blue (benign, strong; benign, supporting); pink (pathogenic, supporting; pathogenic, moderate); and red (pathogenic, strong; pathogenic, very strong). In order to reduce textual clutter within each cell, we display only the evidence code name (e.g. “BA1”), with descriptive text displayed only on mouse hover-over. If a cell contains multiple evidence codes, the cell is divided vertically into symmetric portions equal to the number of evidence codes (maximum 3). If a cell contains no evidence codes, then none is displayed. Each code box contains indicators for whether this code was predicted by the rules engine, indicators for how many other reviewers selected that evidence code (user identities displayed on hover-over), and a check box for the current user to select that evidence code. The evidence code box will change colors if it provides supporting evidence for a specific pathogenicity assignment. Below the Evidence Codes Tool is the Interpretation Tool, which is an interface to calculate and assign interpretations. There are 5 color-coded columns: blue (LB; B); light purple (VUS); and red (LP; P). Indicators represent interpretation assignments predicted by the rules engine, as well as those submitted by other reviewers. A check box allows the user to select their interpretation. At the bottom of the Interpretation Tool is a comment box for users to include additional information (optional). Below the Interpretation Tool are action buttons, which include “Save Opinion,” “Finalize Interpretation”, and “Cancel”. “Save Opinion” allows Reviewers to submit their individual interpretation opinions for review, “Finalize” allows the Report Author to finalize the interpretation information, and “Cancel” allows the user to not save any changes.

The bottom section is the Evidence Browser. This section is where data needed for variant interpretation are displayed, such as variant information, phenotypic data, and annotation pipeline data. Variant data are pulled from multiple internal and external sources. Phenotypic data includes data previously entered by the intake specialist or another Editor. Annotation data are displayed in a form that resembles the spreadsheet currently in use. Given pipelines are highly customized with heterogeneous data sources and formats, and providers are accustomed to viewing annotation data outputs in a density-dense spreadsheet format, our initial plan is to display the input annotation file in a web-based spreadsheet viewer.

After submitting their interpretation, Report Author and Reviewers are taken to the separate Report Editor interface, where human readable interpretation text snippets will be automatically generated

for each classified variant, along with its supporting evidence. These interpretation text snippets are automatically saved in draft form and can be edited. Editors and Reviewers can optionally save interpretation text snippets drafts to assist the Report Author. The Report Author will be able to edit and finalize interpretation snippets, then drag and drop these snippets into a template report form to expedite report authoring. Any finalized interpretation snippets will be saved in the snippets database for future reuse, and can be submitted to ClinVar or other public variant repository.

The goal of the initial GUI design process (**Figure 5.9**) was to develop a starting layout to generate feedback through the user-centered design process. Therefore, it is not surprising that the GUI underwent significant revisions based on user feedback. The current GUI design (**Figure 5.15**) retains some elements of the initial GUI design, while reflecting other significant changes. The current layout retains the use of separate panels for Evidence Review and Interpretation Panel. The Interpretation Panel has been moved to the top to promote a sense of visual consistency while the user browses through heterogeneous data types in the Evidence Review panel. The Interpretation Panel interface was significantly revised based on user feedback on workflow and functionality. The variant pathogenicity assignment question was replaced by a tabular interface to support guidelines-based classification decision support. The assignment confidence question was removed based on user feedback. The assignment rationale free text box was de-emphasized in favor of capturing structured evidence based on user-assigned evidence codes. The Evidence Review panel was also revised based on user feedback. The visualization box was removed, and variant annotation data is displayed in a tabular, spreadsheet format to mimic existing artifacts. Tabs supporting the Delphi Method-based consensus process was removed and replaced with tabs to view other types of data, such as pedigree diagrams. Overall, the current interface blends the initial design work by the research team with the user-centered design process, to better serve the needs of end users.

Other interfaces in development (not shown) include Login Page, Patient Browser, Case Browser, Variant Browser, Task Checklist, Phenotype Editor, Report Author, Search Results, User Management Page, and User Profile.

Chapter 6. DISCUSSION

We report on a user-centered design for a novel tool that supports role-based collaboration, consensus making, and knowledge sharing in variant interpretation. The primary purpose of our research is to address the issue of inter-evaluator discordance in variant interpretation, a long-standing problem in genomic medicine that hinders the effective implementation of precision genomic medicine. While the publication of the ACMG/AMP Guidelines is a major milestone, there is still much work to be done to standardize and improve the reliability of variant interpretation. ClinGen provides significant leadership in coordinating efforts to improve the quality of variant knowledge bases such as ClinVar, and producing standards-based tools to support variant interpretation. The ClinGen Pathogenicity Calculator and the Variant Curation Tool are examples of excellent tools produced by the consortium. Other groups have built tools to support guidelines implementation, such as InterVar and Genetic Variant Interpretation Tool. Building upon prior work, we propose a design for an openly available tool that supports role-based collaboration, consensus-making, knowledge base management, and knowledge sharing for variant interpretation.

We used user-centered design methodology in the design of our tool. These techniques are not commonly used in genetics and genomics, and should be further encouraged. Genomics sciences have become a computationally-intensive field, but genomic analysis tools are not easily usable. Our experience demonstrated that user-centered design methodology can uncover extremely rich data regarding user requirements and desired functionality that can significantly change the direction of the design process. In our case, the proposed Delphi Method-based workflow, a core component of our initial design, was eliminated based on user feedback, and the project's focus pivoted to streamlining role-based collaborative processes as well as data flows. In this case, the primary designer is not a SME, and therefore relied heavily upon user input to understand existing processes. However, even designers who also are SMEs stand to benefit from using user-centered design methodology to elicit other perspectives from within the same field.

In the design process of this tool, we encountered open questions that will require further exploration. Our tool can support linking of patient and variant files to support linked pedigrees

and co-segregation studies. However, Ethical, Legal, and Social Implication (ELSI) considerations, such as patient consent, patient privacy, risk of unintentional disclosure of sensitive information, and others will need to be explored and addressed. Additionally, another area that warrants further exploration is the development of emerging standards for clinical genomic data, which have not matured to the point of widespread adoption. There is a need for standards for annotation files and structured genomic test report documents. The Clinical-Grade Variant File Specifications Workgroup, established by the US Centers for Disease Control and HL7 Clinical Genomics Workgroup in 2012, is defining a VCFclin standard to represent clinical-grade NGS based sequence variants and haplotype results; however, work remains in the early stages, and VCF files remain the most common accepted standard[58]. The HL7 Clinical Genomics Pedigree Model Release 1 is an emerging data standard for transmitting pedigree information, which should be considered for use in future implementations of a pedigree module in MEGA[59]. Additionally, the Human Phenotype Ontology (HPO), an emerging standardized vocabulary for phenotypic findings in active development, should be considered for use in future implementations of the Phenotype module[60].

We endeavor to develop an openly available, open-source tool that supports collaboration, consensus making, and knowledge-sharing in variant interpretation, for the benefit of the genomics community. We believe our design is mature enough to move forward into building a working prototype. Future direction include performing additional design feedback sessions to further develop the design, performing usability testing with a working prototype, developing a working Minimally Viable Product (MVP), and implementing within an actual laboratory's workflow.

BIBLIOGRAPHY

1. Collins FS, Varmus H. A New Initiative on Precision Medicine. N Engl J Med. 2015 Feb 26;372(9):793–5.

2. Karczewski KJ, Daneshjou R, Altman RB. Chapter 7: Pharmacogenomics. PLoS Comput Biol. 2012;8(12).

3. Innocenti F, van Schaik RHN. Pharmacogenomics. Second. Innocenti F, van Schaik RHN, editors. Pharmacogenomics. Totowa, NJ: Humana Press; 2013. (Methods in Molecular Biology; vol. 1015).

4. Masys DR, Jarvik GP, Abernethy NF, Anderson NR, Papanicolaou GJ, Paltoo DN, et al. Technical desiderata for the integration of genomic data into Electronic Health Records. J Biomed Inform. 2012;45(3):419–22.

5. Welch BM, Eilbeck K, Del Fiol G, Meyer LJ, Kawamoto K. Technical desiderata for the integration of genomic data with clinical decision support. J Biomed Inform. 2014 Oct;51:3–7.

6. Gallego CJ, Bennette CS, Heagerty P, Comstock B, Horike-Pyne M, Hisama F, et al. Comparative effectiveness of next generation genomic sequencing for disease diagnosis: Design of a randomized controlled trial in patients with colorectal cancer/polyposis syndromes. Contemp Clin Trials. 2014;39(1):1–8.

7. Amendola LM, Jarvik GP, Leo MC, McLaughlin HM, Akkari Y, Amaral MD, et al. Performance of ACMG-AMP Variant-Interpretation Guidelines among Nine Laboratories in the Clinical Sequencing Exploratory Research Consortium. Am J Hum Genet. 2016 Jun 2;98(6):1067–76.

8. Kuderer NM, Burton KA, Blau S, Rose AL, Parker S, Lyman GH, et al. Comparison of 2 Commercially Available Next-Generation Sequencing Platforms in Oncology. JAMA Oncol. 2016 Dec 15;1–3.

9. Amendola LM, Dorschner MO, Robertson PD, Salama JS, Hart R, Shirts BH, et al. Actionable exomic incidental findings in 6503 participants: Challenges of variant classification. Genome Res. 2015;25(3):305–15.

10. Overby CL, Kohane I, Kannry JL, Williams MS, Starren J, Bottinger E, et al. Opportunities for genomic clinical decision support interventions. Genet Med. 2013;15(10):817–23.

11. Landrum MJ, Lee JM, Benson M, Brown G, Chao C, Chitipiralla S, et al. ClinVar: Public archive of interpretations of clinically relevant variants. Nucleic Acids Res. 2016;44(D1):D862–8.

12. Online Mendelian Inheritance in Man (OMIM) [Internet]. Available from: https://omim.org/

13. Stenson PD, Ball E V., Mort M, Phillips AD, Shiel JA, Thomas NST, et al. Human Gene Mutation Database (HGMD): 2003 update. Hum Mutat. 2003 Jun;21(6):577–81.

14. Dorschner MO, Amendola LM, Turner EH, Robertson PD, Shirts BH, Gallego CJ, et al. Actionable, pathogenic incidental findings in 1,000 participants' exomes. Am J Hum Genet. 2013;93(4):631–40.

15. Rehm HL, Berg JS, Brooks LD, Bustamante CD, Evans JP, Landrum MJ, et al. ClinGen — The Clinical Genome Resource. N Engl J Med. 2015 Jun 4;372(23):2235–42.

16. Harrison SM, Riggs ER, Maglott DR, Lee JM, Azzariti DR, Niehaus A, et al. Using ClinVar as a Resource to Support Variant Interpretation. Curr Protoc Hum Genet. 2016;(April):8.16.1-8.16.23.

17. Hunter JE, Irving SA, Biesecker LG, Buchanan A, Jensen B, Lee K, et al. A standardized, evidence-based protocol to assess clinical actionability of genetic disorders associated with genomic variation. Genet Med. 2016 Apr 28;(April).

18. Patel RY, Shah N, Jackson AR, Ghosh R, Pawliczek P, Paithankar S, et al. ClinGen Pathogenicity Calculator: a configurable system for assessing pathogenicity of genetic variants. Genome Med. 2017 Dec 12;9(1):3.

19. ClinGen's Resources & Tools [Internet]. Available from: https://www.clinicalgenome.org/resources-tools/

20. Pepin MG, Murray ML, Bailey S, Leistritz-Kessler D, Schwarze U, Byers PH. The challenge of comprehensive and consistent sequence variant interpretation between clinical laboratories. Genet Med. 2016;18(1):20–4.

21. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015;17(5):405–23.

22. Harrison SM, Dolinsky JS, Knight Johnson AE, Pesaran T, Azzariti DR, Bale S, et al. Clinical laboratories collaborate to resolve differences in variant interpretations submitted to ClinVar. Genet Med. 2017;(January).

23. Rehm H. Genomic Testing: Actionability, Validation, and Standards for Lab Reports. eMERGE Consortium Meeting. 2014 Jan.

24. Kleinberger J, Maloney KA, Pollin TI, Jeng LJB. An openly available online tool for implementing the ACMG/AMP standards and guidelines for the interpretation of sequence variants. Genet Med. 2016 Nov 17;18(11):1165–1165.

25. Genetic Variant Interpretation Tool [Internet]. Available from: http://www.medschool.umaryland.edu/Genetic_Variant_Interpretation_Tool1.html

26. Li Q, Wang K. InterVar: Clinical Interpretation of Genetic Variants by the 2015 ACMG-AMP Guidelines. Am J Hum Genet. 2017 Feb 2;100(2):267–80.

27. wInterVar [Internet]. Available from: http://wintervar.wglab.org/

28. ClinGen Pathogenicity Calculator [Internet]. Available from: http://calculator.clinicalgenome.org/site/cg-calculator

29. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 2010 Sep 1;38(16):e164–e164.

30. ClinGen Curator Interfaces [Internet]. Available from: https://curation-test.clinicalgenome.org/

31. ClinGen | Cherry Lab [Internet]. Available from: https://cherrylab.stanford.edu/projects/clingen

32. Usability.gov. User-Centered Design Basics [Internet]. [cited 2017 May 26]. Available from: https://www.usability.gov/what-and-why/user-centered-design.html

33. Cutting E, Banchero M, Beitelshees AL, Cimino JJ, Fiol G Del, Gurses AP, et al. User-centered design of multi-gene sequencing panel reports for clinicians. J Biomed Inform. 2016 Oct;63:1–10.

34. Bolchini D, Finkelstein A, Perrone V, Nagl S. Better bioinformatics through usability analysis. Bioinformatics. 2009;25(3):406–12.

35. Pavelin K, Cham JA, de Matos P, Brooksbank C, Cameron G, Steinbeck C. Bioinformatics meets user-centred design: A perspective. PLoS Comput Biol. 2012;8(7):1–4.

36. de Matos P, Cham JA, Cao H, Alcántara R, Rowland F, Lopez R, et al. The Enzyme Portal: a case study in applying user-centred design methods in bioinformatics. BMC Bioinformatics. 2013;14(1):103.

37. Néron B, Ménager H, Maufrais C, Joly N, Maupetit J, Letort S, et al. Mobyle: A new full web bioinformatics framework. Bioinformatics. 2009;25(22):3005–11.

38. Javahery H, Seffah A, Radhakrishnan T. Beyond power: making bioinformatics tools user-centered. Commun ACM. 2004 Nov 1;47(11):58.

39. Kulak D, Guiney E. The Facade Iteration & The Filled Interation. In: Use Cases: Requirements in Context. First Edit. 2000.

40. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)-A metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform. 2009;42(2):377–81.

41. Lucidchart.com [Internet]. Available from: https://www.lucidchart.com/

42. Microsoft Visio Professional 2016. Redmond, WA: Microsoft Corporation; 2015.

43. Microsoft Powerpoint Professional 2016. Redmon: Microsoft Corporation; 2015.

44. Macefield R. How To Specify the Participant Group Size for Usability Studies: A Practitioner ’s Guide. J Usability Stud. 2009;5(1):34–45.

45. Fusch PI, Ness LR. Are we there yet? Data saturation in qualitative research. Qual Rep. 2015;20(9):1408–16.

46. Glaser BG, Strauss AL. The Discovery of Grounded Theory: Strategies for Qualitative Research. Sociology Press; 1967.

47. Walker D, Myrick F. Grounded Theory: An Exploration of Process and Procedure. Qual Health Res. 2006;16(4):547–59.

48. Charmaz K. Grounded Theory. In: The SAGE Encyclopedia of Social Science Research Methods. 2455 Teller Road, Thousand Oaks California 91320 United States of America: Sage Publications, Inc.; 2004. p. 49–56.

49. Rasmussen L V., Overby CL, Connolly J, Chute CG, Denny JC, Freimuth R, et al. Practical considerations for implementing genomic information resources. Appl Clin Inform. 2016;7(3):870–82.

50. Parry KW. Constant Comparison. In: The SAGE Encyclopedia of Qualitative Research Methods. 2004.

51. ATLAS.ti 8 Windows. Berlin, Germany: ATLAS.ti Scientific Software Development GmbH; 2016.

52. Tarczy-Hornoch P, Amendola L, Aronson SJ, Garraway L, Gray S, Grundmeier RW, et al. A survey of informatics approaches to whole-exome and whole-genome clinical reporting in the electronic health record. Genet Med. 2013 Oct 26;15(10):824–32.

53. Good BM, Ainscough BJ, McMichael JF, Su AI, Griffith OL. Organizing knowledge to enable personalization of medicine in cancer. Genome Biol. 2014;15(8):438.

54. Helmer O. Analysis of the future: The Delphi Method. Rand Corporation. 1967. p. 1–11.

55. Skulmoski G, Hartman F, Krahn J. The Delphi Method for Graduate Research. J Inf Technol Educ Res. 2007;6(1):1–21.

56. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Australia: R Foundation for Statistical Computing; 2017.

57. Crosslin DR, Robertson PD, Carrell DS, Gordon AS, Hanna DS, Burt A, et al. Prospective participant selection and ranking to maximize actionable pharmacogenetic variants and

discovery in the eMERGE Network. Genome Med. 2015 Dec 3;7(1):67.

58. VCFClin.org. Clinical-Grade Variant File Specifications Workgroup [Internet]. [cited 2017 Jun 8]. Available from: http://vcfclin.org/variant-file-working-group/

59. Health Level Seven International. HL7 Version 3 Standard: Clinical Genomics; Pedigree, Release 1 [Internet]. [cited 2017 Jun 8]. Available from: http://www.hl7.org/implement/standards/product_brief.cfm?product_id=8

60. Köhler S, Doelken SC, Mungall CJ, Bauer S, Firth H V., Bailleul-Forestier I, et al. The Human Phenotype Ontology project: linking molecular biology and disease through phenotype data. Nucleic Acids Res. 2014 Jan;42(D1):D966–74.

APPENDIX A: ACMG/AMP GUIDELINES

Evidence of pathogenicity	Category
Very strong	PVS1 null variant (nonsense, frameshift, canonical ±1 or 2 splice sites, initiation codon, single or multiexon deletion) in a gene where LOF is a known mechanism of disease Caveats: • Beware of genes where LOF is not a known disease mechanism (e.g., *GFAP*, *MYH7*) • Use caution interpreting LOF variants at the extreme 3′ end of a gene • Use caution with splice variants that are predicted to lead to exon skipping but leave the remainder of the protein intact • Use caution in the presence of multiple transcripts
Strong	PS1 Same amino acid change as a previously established pathogenic variant regardless of nucleotide change Example: Val→Leu caused by either G>C or G>T in the same codon Caveat: Beware of changes that impact splicing rather than at the amino acid/protein level PS2 De novo (both maternity and paternity confirmed) in a patient with the disease and no family history Note: Confirmation of paternity only is insufficient. Egg donation, surrogate motherhood, errors in embryo transfer, and so on, can contribute to nonmaternity. PS3 Well-established in vitro or in vivo functional studies supportive of a damaging effect on the gene or gene product Note: Functional studies that have been validated and shown to be reproducible and robust in a clinical diagnostic laboratory setting are considered the most well established. PS4 The prevalence of the variant in affected individuals is significantly increased compared with the prevalence in controls Note 1: Relative risk or OR, as obtained from case–control studies, is >5.0, and the confidence interval around the estimate of relative risk or OR does not include 1.0. See the article for detailed guidance. Note 2: In instances of very rare variants where case–control studies may not reach statistical significance, the prior observation of the variant in multiple unrelated patients with the same phenotype, and its absence in controls, may be used as moderate level of evidence.
Moderate	PM1 Located in a mutational hot spot and/or critical and well-established functional domain (e.g., active site of an enzyme) without benign variation PM2 Absent from controls (or at extremely low frequency if recessive) (Table 6) in Exome Sequencing Project, 1000 Genomes Project, or Exome Aggregation Consortium Caveat: Population data for insertions/deletions may be poorly called by next-generation sequencing. PM3 For recessive disorders, detected in *trans* with a pathogenic variant Note: This requires testing of parents (or offspring) to determine phase. PM4 Protein length changes as a result of in-frame deletions/insertions in a nonrepeat region or stop-loss variants PM5 Novel missense change at an amino acid residue where a different missense change determined to be pathogenic has been seen before Example: Arg156His is pathogenic; now you observe Arg156Cys Caveat: Beware of changes that impact splicing rather than at the amino acid/protein level. PM6 Assumed de novo, but without confirmation of paternity and maternity
Supporting	PP1 Cosegregation with disease in multiple affected family members in a gene definitively known to cause the disease Note: May be used as stronger evidence with increasing segregation data PP2 Missense variant in a gene that has a low rate of benign missense variation and in which missense variants are a common mechanism of disease PP3 Multiple lines of computational evidence support a deleterious effect on the gene or gene product (conservation, evolutionary, splicing impact, etc.) Caveat: Because many in silico algorithms use the same or very similar input for their predictions, each algorithm should not be counted as an independent criterion. PP3 can be used only once in any evaluation of a variant. PP4 Patient's phenotype or family history is highly specific for a disease with a single genetic etiology PP5 Reputable source recently reports variant as pathogenic, but the evidence is not available to the laboratory to perform an independent evaluation

LOF, loss of function; OR, odds ratio.

Figure 6.1. ACMG/AMP Guidelines: evidence of pathogenicity.

Reproduced without modification from Table 3, Richards et al. (2015)[21].

Evidence of benign impact	Category
Stand-alone	BA1 Allele frequency is >5% in Exome Sequencing Project, 1000 Genomes Project, or Exome Aggregation Consortium
Strong	BS1 Allele frequency is greater than expected for disorder (see **Table 6**)
	BS2 Observed in a healthy adult individual for a recessive (homozygous), dominant (heterozygous), or X-linked (hemizygous) disorder, with full penetrance expected at an early age
	BS3 Well-established in vitro or in vivo functional studies show no damaging effect on protein function or splicing
	BS4 Lack of segregation in affected members of a family
	Caveat: The presence of phenocopies for common phenotypes (i.e., cancer, epilepsy) can mimic lack of segregation among affected individuals. Also, families may have more than one pathogenic variant contributing to an autosomal dominant disorder, further confounding an apparent lack of segregation.
Supporting	BP1 Missense variant in a gene for which primarily truncating variants are known to cause disease
	BP2 Observed in *trans* with a pathogenic variant for a fully penetrant dominant gene/disorder or observed in cis with a pathogenic variant in any inheritance pattern
	BP3 In-frame deletions/insertions in a repetitive region without a known function
	BP4 Multiple lines of computational evidence suggest no impact on gene or gene product (conservation, evolutionary, splicing impact, etc.)
	Caveat: Because many in silico algorithms use the same or very similar input for their predictions, each algorithm cannot be counted as an independent criterion. BP4 can be used only once in any evaluation of a variant.
	BP5 Variant found in a case with an alternate molecular basis for disease
	BP6 Reputable source recently reports variant as benign, but the evidence is not available to the laboratory to perform an independent evaluation
	BP7 A synonymous (silent) variant for which splicing prediction algorithms predict no impact to the splice consensus sequence nor the creation of a new splice site AND the nucleotide is not highly conserved

Figure 6.2. ACMG/AMP Guidelines: evidence of benign impact.

Reproduced without modification from Table 4, Richards et al. (2015)[21].

Pathogenic	(i) 1 Very strong (PVS1) *AND*
	(a) ≥1 Strong (PS1–PS4) *OR*
	(b) ≥2 Moderate (PM1–PM6) *OR*
	(c) 1 Moderate (PM1–PM6) and 1 supporting (PP1–PP5) *OR*
	(d) ≥2 Supporting (PP1–PP5)
	(ii) ≥2 Strong (PS1–PS4) *OR*
	(iii) 1 Strong (PS1–PS4) *AND*
	(a)≥3 Moderate (PM1–PM6) *OR*
	(b)2 Moderate (PM1–PM6) *AND* ≥2 Supporting (PP1–PP5) *OR*
	(c)1 Moderate (PM1–PM6) *AND* ≥4 supporting (PP1–PP5)
Likely pathogenic	(i) 1 Very strong (PVS1) *AND* 1 moderate (PM1–PM6) *OR*
	(ii) 1 Strong (PS1–PS4) *AND* 1–2 moderate (PM1–PM6) *OR*
	(iii) 1 Strong (PS1–PS4) *AND* ≥2 supporting (PP1–PP5) *OR*
	(iv) ≥3 Moderate (PM1–PM6) *OR*
	(v) 2 Moderate (PM1–PM6) *AND* ≥2 supporting (PP1–PP5) *OR*
	(vi) 1 Moderate (PM1–PM6) *AND* ≥4 supporting (PP1–PP5)
Benign	(i) 1 Stand-alone (BA1) *OR*
	(ii) ≥2 Strong (BS1–BS4)
Likely benign	(i) 1 Strong (BS1–BS4) and 1 supporting (BP1–BP7) *OR*
	(ii) ≥2 Supporting (BP1–BP7)
Uncertain significance	(i) Other criteria shown above are not met OR
	(ii) the criteria for benign and pathogenic are contradictory

Figure 6.3. ACMG/AMP Guidelines: rules for classifying sequence variants. Reproduced without modification from Table 5, Richards et al. (2015)[21].

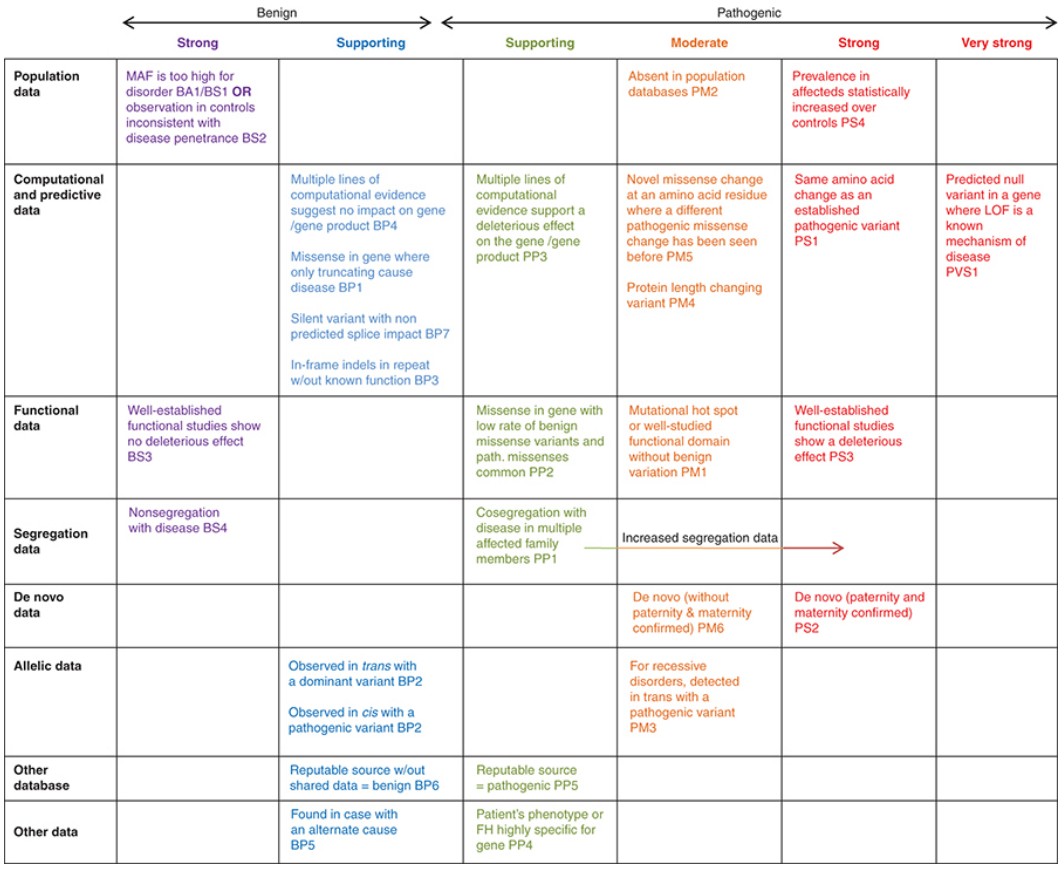

Figure 6.4. ACMG/AMP Guidelines: evidence framework.

Reproduced without modification from Figure 1, Richards et al. (2015)[21].