



Proactive Approaches and Automated Aspects of Variant Classification

Karla R. Bowles, PhD, FACMG
Senior Laboratory Director



Lab Accuracy

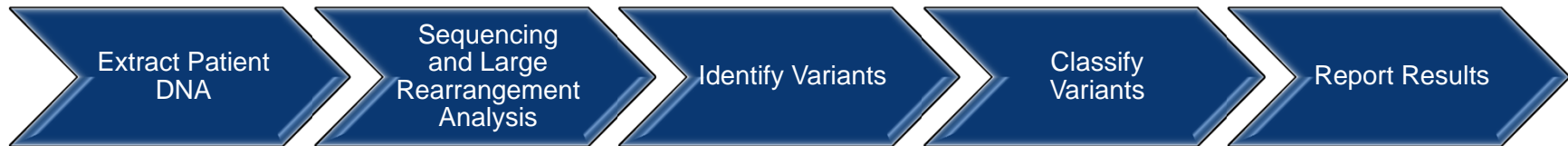


Can the lab find all of the variants (DNA changes)?

Does the lab classify variants (DNA changes) correctly?



Variant Classification Process



- Process may be manageable for a few novel variants/day, but what about 100 novel variants/day?
 - Hire more Lab Directors/Genetic Counselors
 - Expensive to maintain
 - Limited number of people review each variant – Error prone
 - Outsource to a third party or use off-the-shelf software
 - Expensive to maintain
 - Who is responsible for accuracy and how is it ensured?
 - Develop and validate automation tools
 - Expensive to develop
 - Lab controls accuracy
 - Cost-effective over time



Our Classification Process

Computer Automated

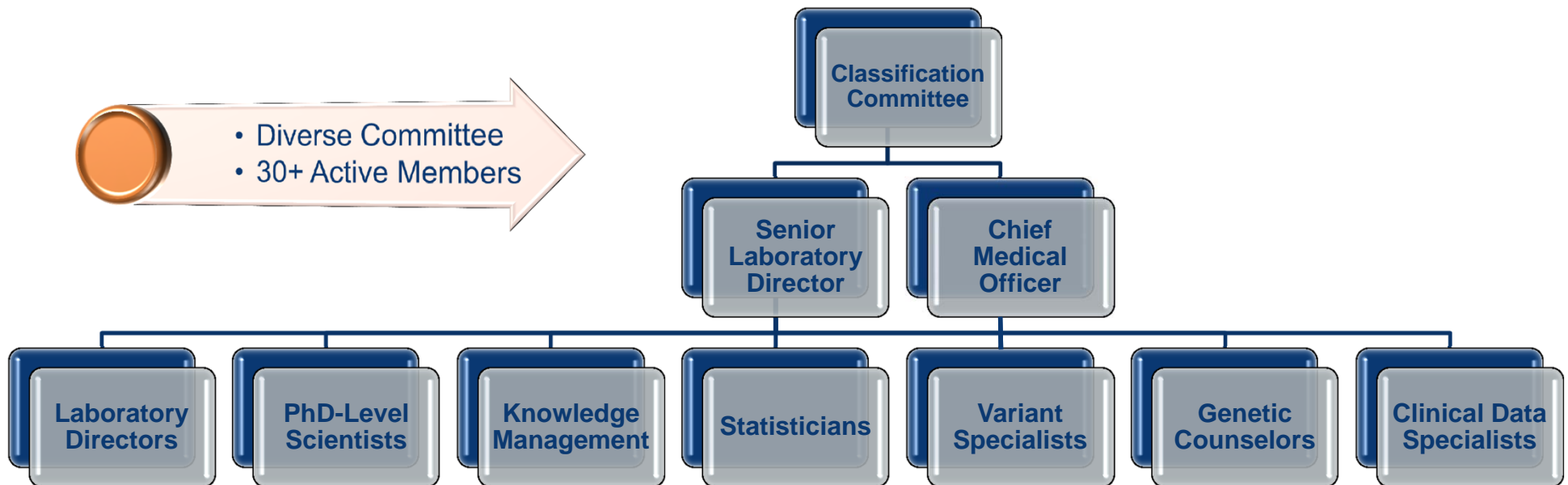
Computer Assisted

Mostly Manual



Our Classification Committee

- 
- Diverse Committee
 - 30+ Active Members





VITA (Variant Information Tracking Application)

Presents important sequence information

Allelic/nearby
variants

Functional
domains

Exon/intron
locations

Splicing
analysis

Population
data

Complex
regions

Automated Computer Analysis



VITA Classification Program

Presents important sequence information

Allelic/nearby
variants

Functional
domains

Exon/intron
locations

Splicing
analysis

Population data

Complex
regions

Bins variants into classification categories

Assigns default classification using predefined/validated classification criteria

Presents results for human review

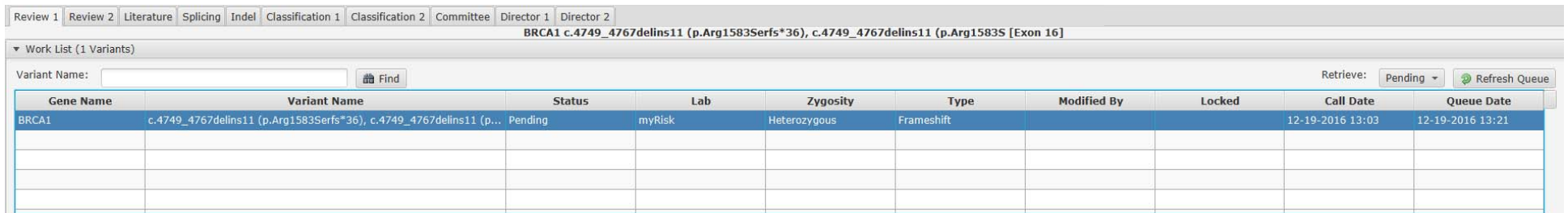
PhD Literature/Expert
Review

RNA Splicing Analysis

Structural Analysis

Data Verification

VITA Classification Program



The screenshot displays the VITA Classification Program interface. At the top, there are navigation tabs: Review 1, Review 2, Literature, Splicing, Indel, Classification 1, Classification 2, Committee, Director 1, and Director 2. Below these tabs, the current variant is identified as "BRCA1 c.4749_4767delins11 (p.Arg1583Serfs*36), c.4749_4767delins11 (p.Arg1583S [Exon 16])". A "Work List (1 Variants)" section contains a search bar for "Variant Name:" with a "Find" button and a "Retrieve:" dropdown menu set to "Pending", along with a "Refresh Queue" button. The main data is presented in a table with the following columns: Gene Name, Variant Name, Status, Lab, Zygosity, Type, Modified By, Locked, Call Date, and Queue Date. The table contains one row of data for the BRCA1 variant.

Gene Name	Variant Name	Status	Lab	Zygosity	Type	Modified By	Locked	Call Date	Queue Date
BRCA1	c.4749_4767delins11 (p.Arg1583Serfs*36), c.4749_4767delins11 (p...	Pending	myRisk	Heterozygous	Frameshift			12-19-2016 13:03	12-19-2016 13:21

- Well Controlled Program Ensures Quality
 - Queue system ensures all analyses are performed before classification committee review and final classification
 - Computer enforces classification verification by multiple individuals
 - Computer alerts users of unexpected classifications
 - Computer provides auditable trail of all data and review notes



Identify and Evaluate Literature

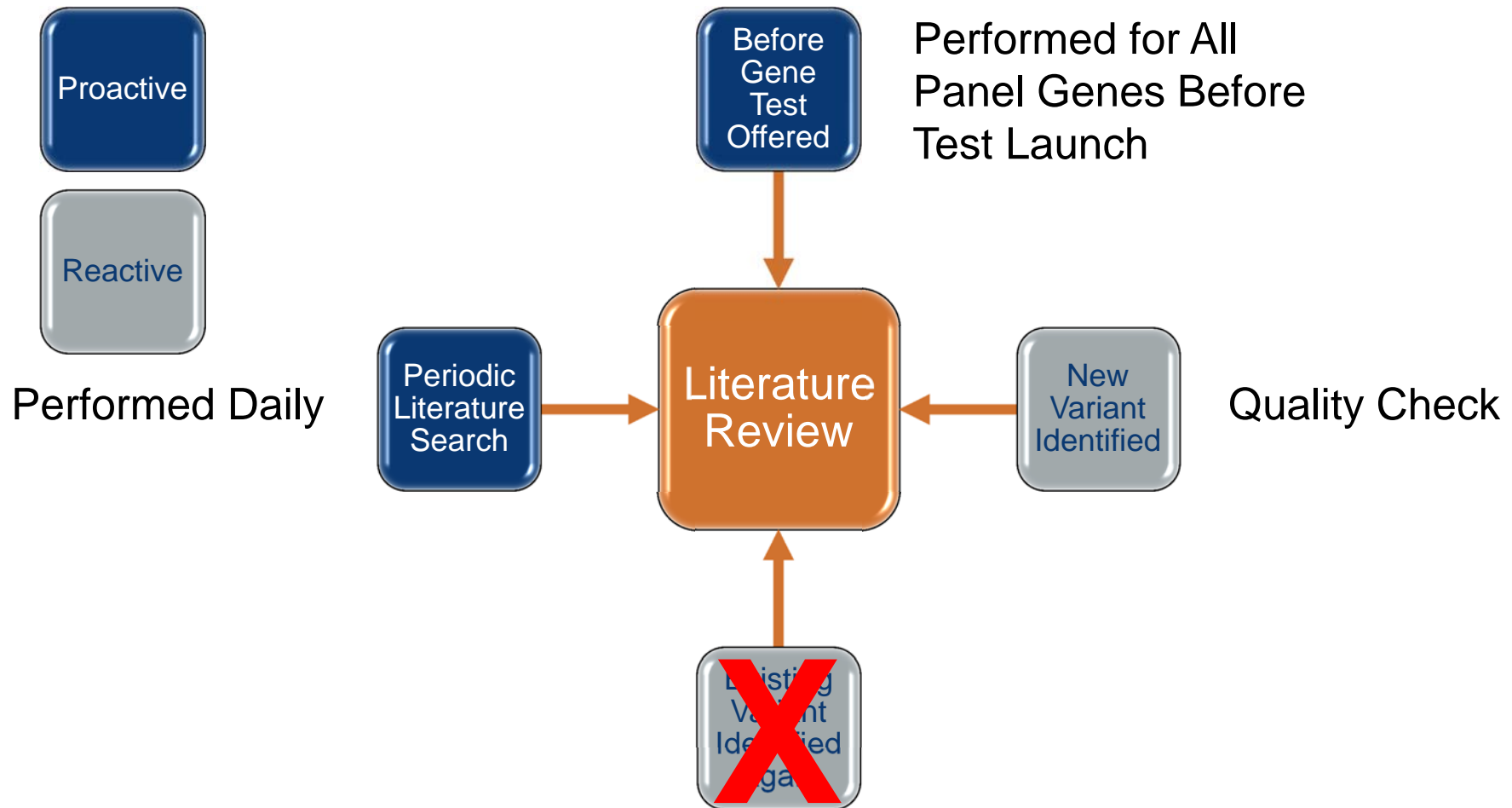
- Critical questions:
 - When and how often should we evaluate the literature?
 - What tools should we use?
 - PubMed
 - Google
 - Third party software
 - Laboratory-developed tools



Literature Review

- Literature lists are generated daily by an automated algorithm that includes:
 - Searches by multiple gene names
 - Searches alternative nomenclatures (i.e., HGVS vs. BIC)
- The Algorithm:
 - Removes redundant citations
 - Provides URLs to publications
 - Highlights search terms found in each reference
 - Sorts by most relevant citation
- Process and algorithm tested and validated to ensure identification of relevant literature

Proactive vs. Reactive Literature Review





Identify and Evaluate Literature

- Knowledge Management – PhD Scientists
 - Review all literature
 - Create a written summary
 - Alert Laboratory Directors and subject experts if significant literature is identified
- Subject Matter Experts – PhD Scientists
 - mRNA splicing analysis – mRNA splicing experts
 - Structural analysis – Structural Biologists
 - Functional analyses –Biochemists
 - Segregation analysis –Geneticists
 - Statistical analysis –Statisticians
 - Population data – Statisticians and Population Geneticists



Variant Reclassification

Computer Automated Process

Computer Assisted Process

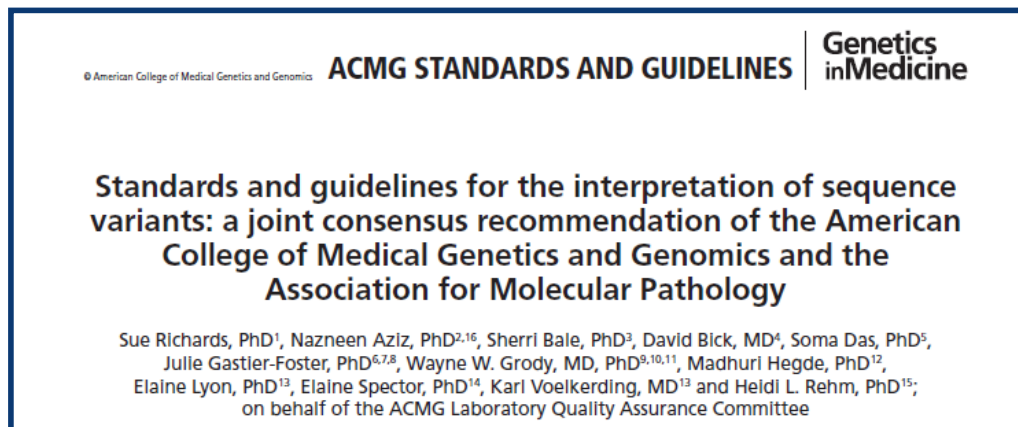
Mostly Manual Process





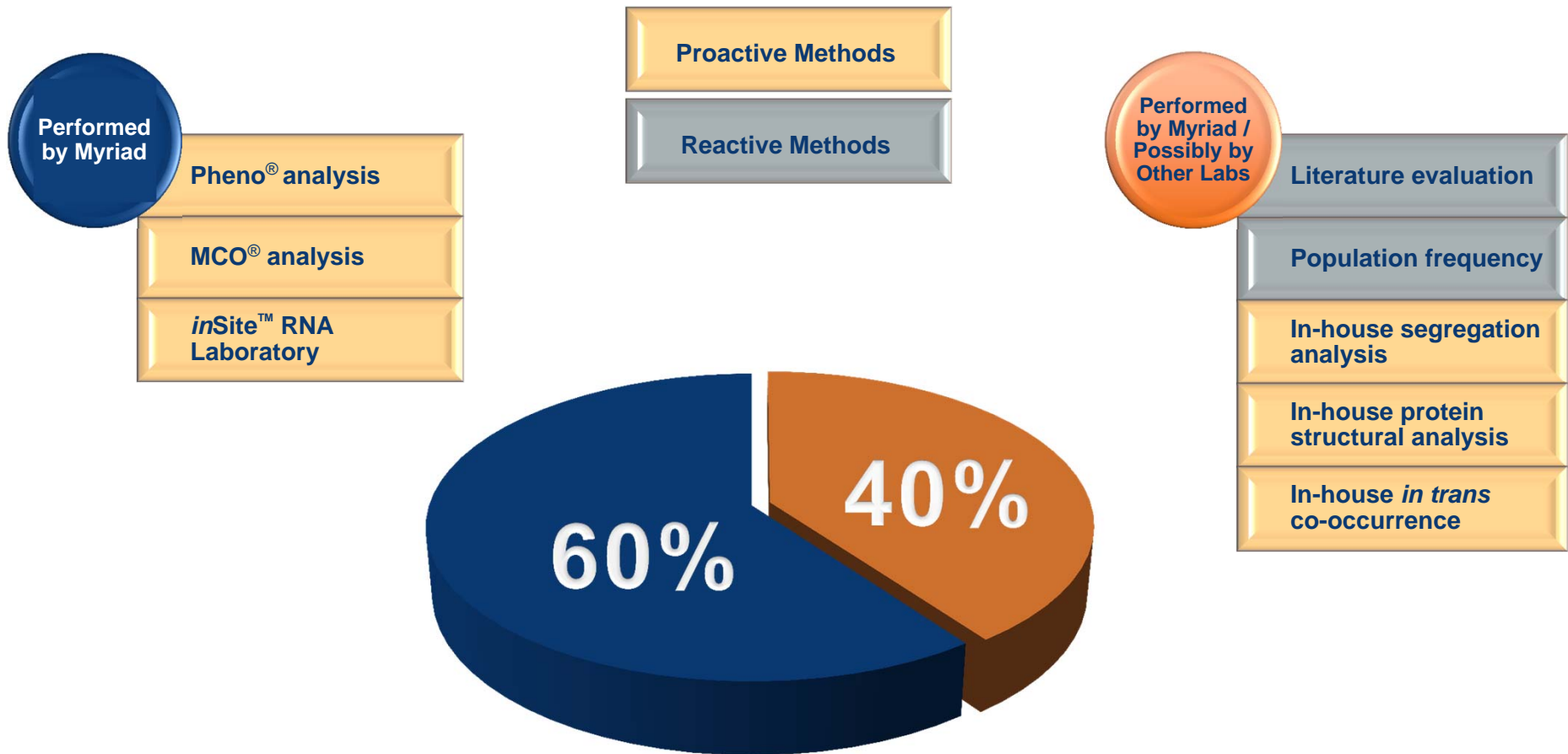
Variant Reclassification

- Variant reclassification has been historically “reactive”



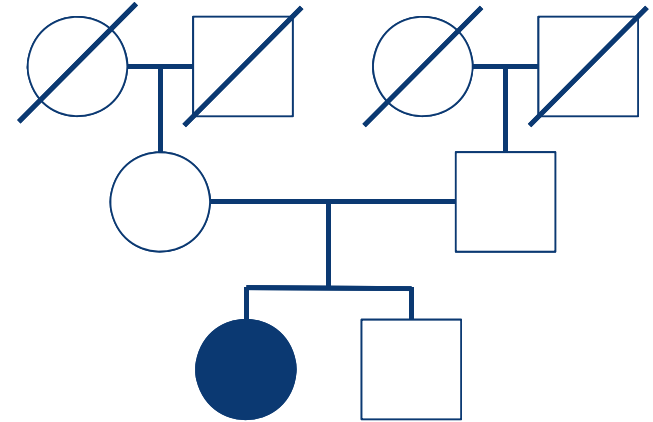
- ACMG classification guidelines are reactive
 - Don't generate novel data
 - Recommend how to analyze data you already have
- “Proactive” reclassification is critical – Generate our own data
 - Most VUS will never be reclassified if we wait for data to come to us

Reclassification Techniques



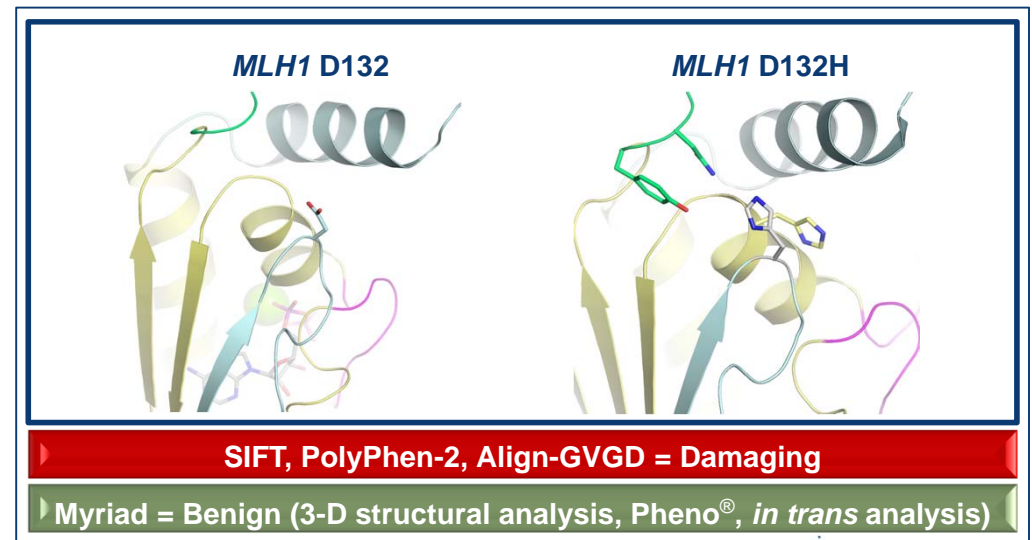
In-House Segregation Analysis

- Classical approach to reclassification
- Limited utility for cancer genetics
 - High phenocopy rate
 - Reduced penetrance for many genes
 - Smaller American family sizes
- Our approach
 - Proactively target variants
 - “Close” to being reclassified
 - Other lines of evidence available
 - Proactively reach out to families and offer free family analysis
 - Store data in custom pedigree program
 - Perform statistical analysis customized to small pedigrees



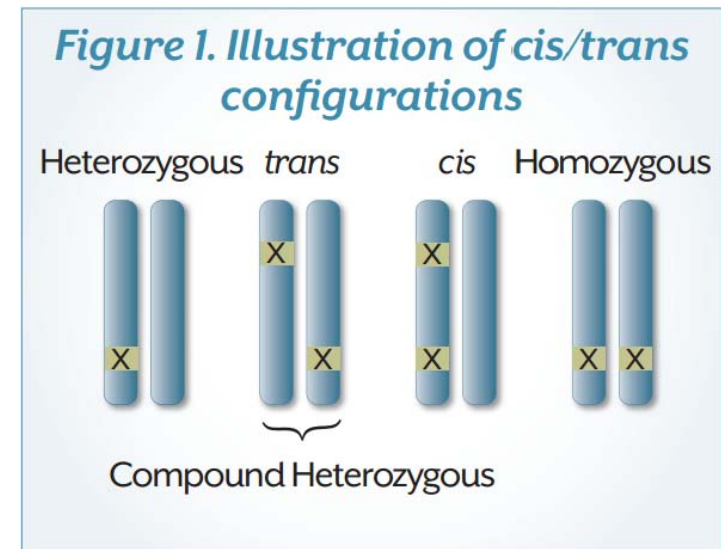
In-House Protein Structural Analysis

- Diagnostic labs generally can't generate crystal structures
- Crystal structures are publicly available – Protein Data Bank (PDB), Protein Data Bank, Europe (PDBe), etc.
- Knowledgeable Structural Biologist required
 - Validate current crystal structures before use
 - Develop additional structures
 - *MLH1* N-terminus – Wu H, Zeng H, Lam R, Tempel W, Kerr ID, Min J (2015). *Acta Crystallogr F Struct Biol Commun* 71, p981-5.
 - Map variants



In-House *In Trans* Co-occurrence Analysis

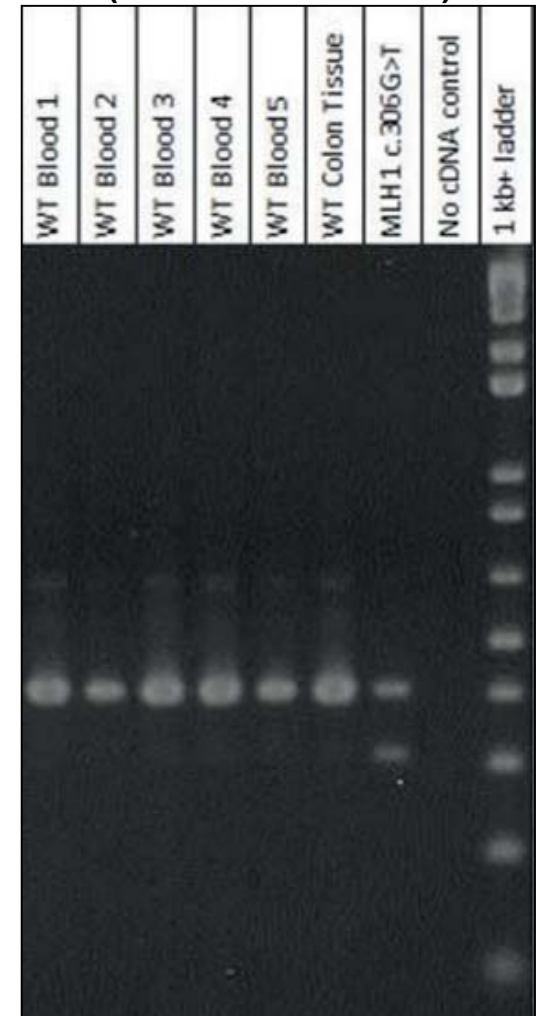
- For many genes, pathogenic variant homozygosity or compound heterozygosity is either lethal or results in a severe phenotype
- Can be used to downgrade variant classification
- We proactively determine phase
 - Offer free family analysis
 - Construct and validate haplotypes
 - Computer determines haplotypes
 - Family testing not required
 - Computer immediately informs classification committee of co-occurrences



In-House RNA Splicing Analysis

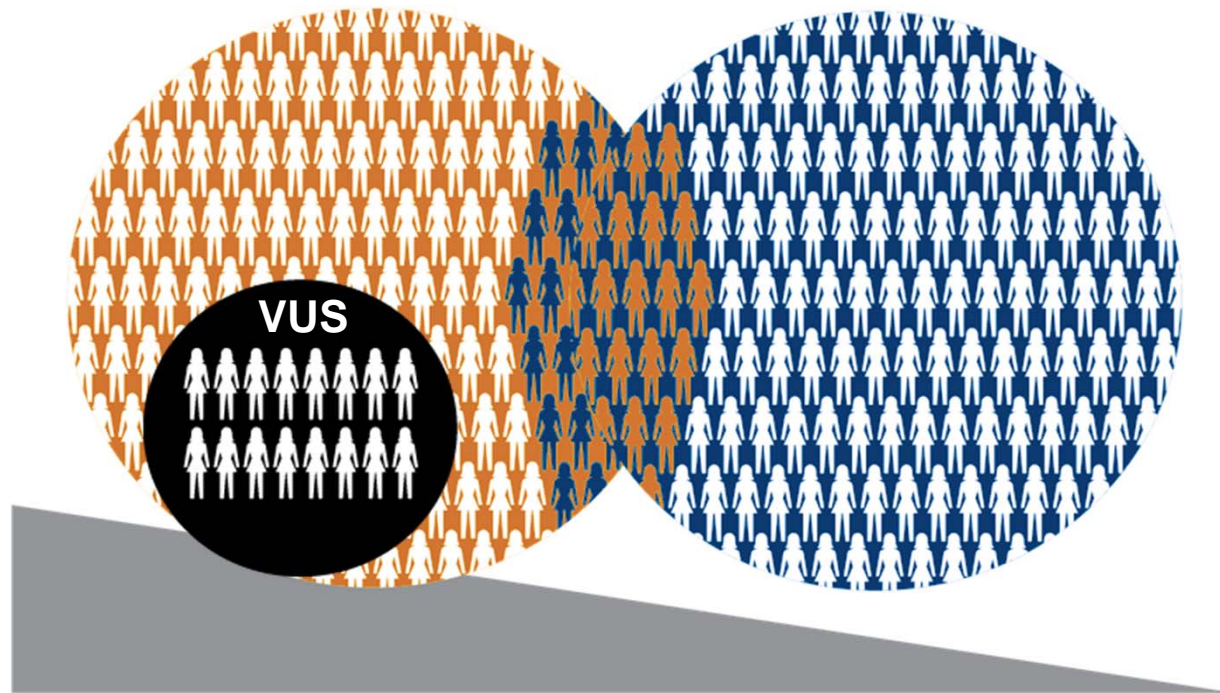
- *In silico* RNA splicing algorithms
 - Flag potential variant-associated splicing defects
 - Often inaccurate
 - Cannot determine full vs. partial splicing defects
 - Additional data required
- Our approach
 - Identify variants with a high likelihood of disrupting RNA splicing
 - Offer free RNA testing to patients
 - Use results to upgrade variants

***MLH1* c.306G>T**
(Last Base of Exon 3)



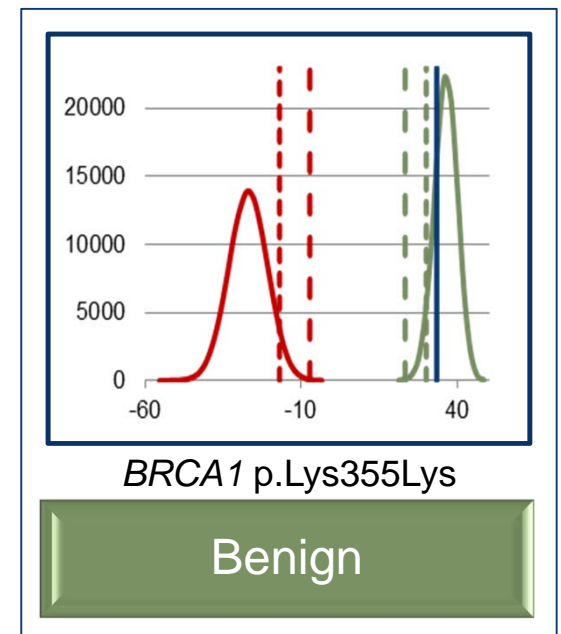
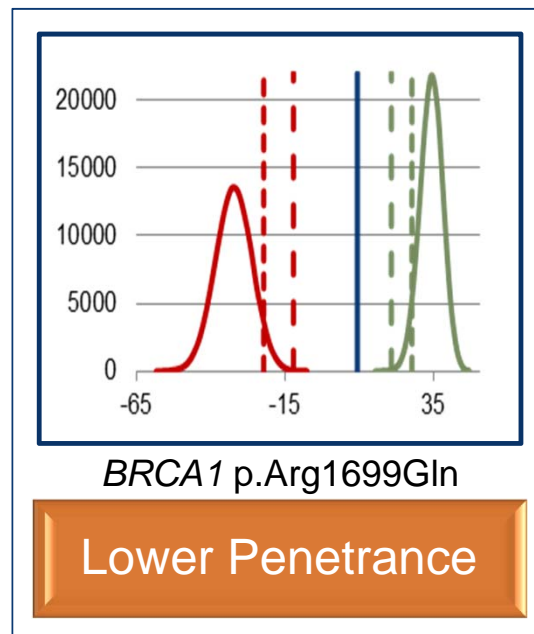
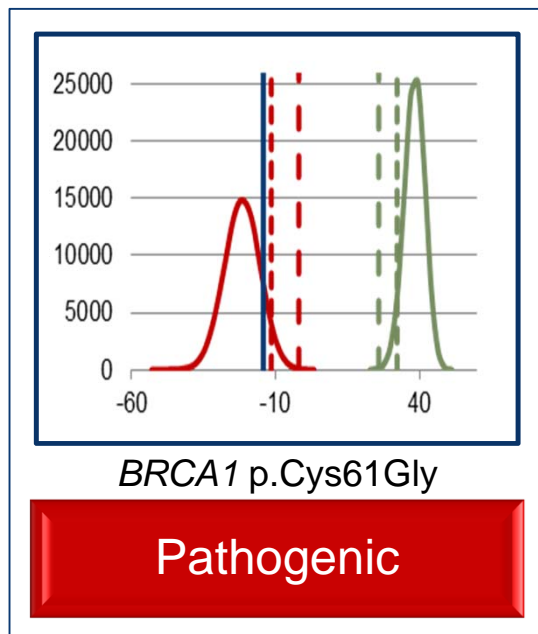


Pheno[®] Analysis



Severity of Personal and Family History of Cancer

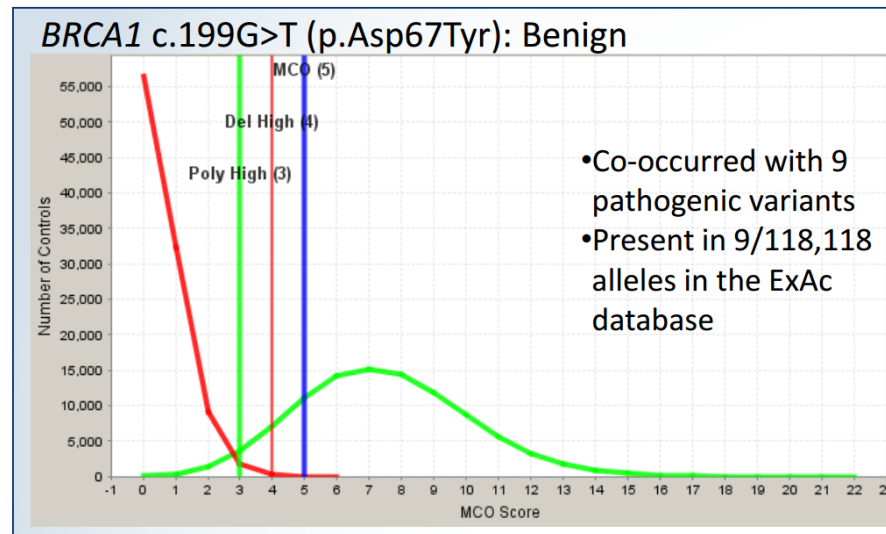
Pheno[®] Analysis



- Pruss D *et al.* Development and validation of a new algorithm for the reclassification of genetic variants identified in the *BRCA1* and *BRCA2* genes. *Breast Cancer Res Treat.* 2014;147(1):119-32.
- Morris B *et al.* Classification of genetic variants in genes associated with Lynch syndrome using a clinical history weighting algorithm. *BMC Genet.* 2016 Jul 1;17(1):99.

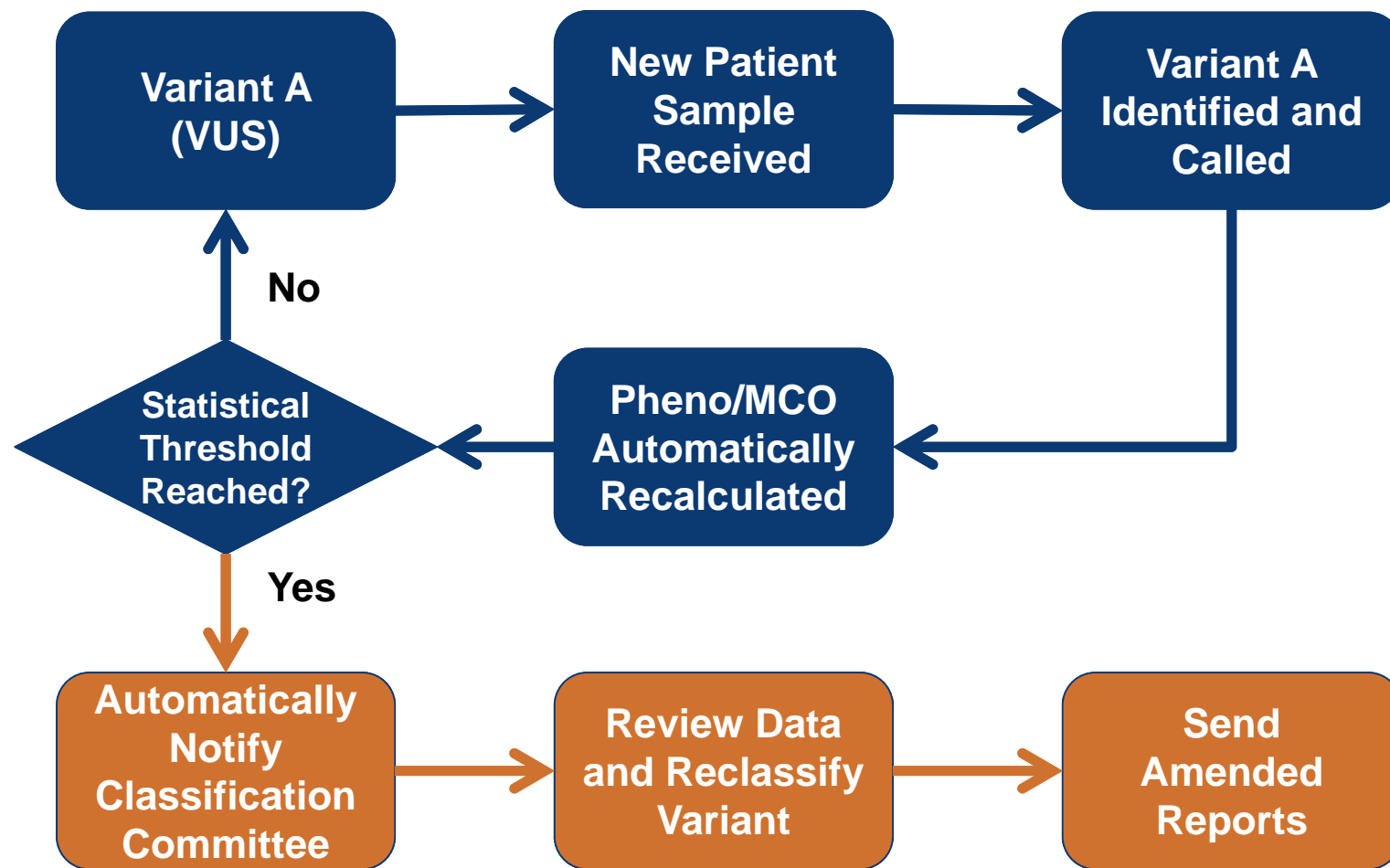
Mutation Co-occurrence (MCO)[®] Analysis

- It is highly unlikely, but not impossible, to carry 2 pathogenic mutations
 - In same gene (*in cis* or *in trans*), or
 - In 2 different genes in the same pathway
 - Example: *BRCA1* and *BRCA2*
- MCO analysis measures the statistical significance of a variant co-occurring multiple times with one or more pathogenic mutations



Coffee B *et al.* Utilization of Mutation Co-occurrence (MCO) Analysis as Evidence for Benign and Likely Benign Variant Classification. Presented at ACMGG Annual Meeting, March 2015.

Pheno[®] and MCO[®] Automation – 24/7





Summary

- Larger gene panels require a robust approach to variant classification and reclassification
- ACMG classification guidelines
 - Address how to evaluate data already obtained
 - Do not address how to generate novel variant data
- Laboratories should proactively develop novel classification technologies and offer these to patients/families undergoing testing as part of standard-of-care
- Novel and proactive technologies will advance the science of variant classification, resulting in more definitive test results and improved patient outcomes