Development and Roll-out of a Variant Database for Use by Newborn Screening Programs

Jaime E. Hale, MS
Clinical Data Scientist
New England Newborn Screening Program
University of Massachusetts Medical School

April 10, 2019
APHL Newborn Screening and Genetic Testing Symposium
Chicago, IL
Posters Presented By Our Group

Variant Interpretation at the Program Level: A Review, with Updates to Previously-Shared Processes
Anna Maria Comeau, PhD
New England Newborn Screening Program, University of Massachusetts Medical School, Worcester, MA, University of Massachusetts Medical School, Department of Pediatrics, Worcester, MA.

1. BACKGROUND
We developed crawling variant interpretation at a technical and functional level to ensure compliance with ACM guidelines. We began with a manual process and developed a software tool that handles this task automatically, streamlining the workflow.

2. ADVANCES
The automated process is a valuable asset in the development of a new database. This database, coupled with the automation, allows the creation of a new database, which can be used to share data with others. The software also facilitates data extraction and storage.

3. ADVANTAGES from Advancement
- Automation reduces the workload for users and increases accuracy.
- Interoperability with existing systems is enhanced.

4. INTERESTED?
UMass Medical School

Newborn Screening Knowledge Framework

ABSTRACT
As per American College of Medical Genetics and Genomics (ACMG) guidelines, all newborns are screened for a variety of conditions. This process involves analyzing genetic markers, assessing variants, and interpreting results. The knowledge framework presented here outlines the steps involved in the interpretation process, from data acquisition to reporting. The framework is designed to ensure consistency and accuracy in the interpretation of genetic data.

Modern Software Architecture

University of Massachusetts Medical School
umassmed.edu
Variant Interpretation at the Program Level: A Review, with Updates to Previously-Shared Processes
Anne Marie Comeau, PhD
Jaime E. Hale, MS, Binod Kumar, PhD, Inderneel Sahai, MD, Mahesh Vangala, MS, Keith Pelletier, BA, Jomol Mathew, PhD
1 New England Newborn Screening Program, University of Massachusetts Medical School, Worcester, MA; 2 University of Massachusetts Medical School, Department of Pediatrics, Worcester, MA; 3 Data Sciences & Technology, University of Massachusetts Medical School, Worcester, MA

1. BACKGROUND
We developed streamlined variant interpretation protocols, and standardized interpretation instruments to ensure compliance with ACMG guidelines. We began with a manual process and developed an RDMS to track variants, document our reviews, and serve as the source from which we provide clinical reports of variants. The RDMS knowledge base also facilitates data extraction and ongoing curation.

2. ADVANCES
We report success in automating our processes in parallel with the development of a modular graph database. Both enhance speed and consistency. The modular graph database provides more functionality and flexibility than the RDMS and enhances the feasibility of sharing any number of processes with others.

3. ADVANTAGES from Advances:
- PLANNED FOR INTERFACING WITH NATIONAL EFFORTS
- MODULAR – USE WHAT YOU WANT

4. INTERESTED?
GIVE US A CALL (774-455-4600)
Three-Fold Purpose for a Local Database

1. Develop an internal tracking system of all variants ever detected

2. Documentation of variant interpretation details

3. Provide a foundation from which we can generate clinical reports to healthcare providers of infants tested
Data Sources...disease databases, EGL, gnoMAD, ClinVar, etc.
### Database Field Design: Variants Table

<table>
<thead>
<tr>
<th>Field</th>
<th>DSdb</th>
<th>IDUA</th>
<th>GAA</th>
<th>ABDC1</th>
<th>EmVClass</th>
<th>gnoMAD</th>
<th>ClinVar</th>
<th>dbSNP</th>
<th>UCSC</th>
<th>Polyphen</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nucleotide Change</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Amino Acid Change/Protein Change/Consequence</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Alias Listings</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>RS Number</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Chromosome Position</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Cytogenetic Location</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Locus</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Variant Type/Annotation</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Molecular Consequence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Classification/Effect/Phenotype</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Date First Reported</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Links to Literature/Citations</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Other Submissions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Allele Count</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Allele Numbers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Number of Homozygotes</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Allele Frequency</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Conservation Across Species</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Date Last Reviewed/Updated</td>
<td>X</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Data Warning</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>
# ACMG Guidelines for Sequence Variations

Adapted from Richards et al. Genet Med. 2015 May;17(5):405-24

## Evidence of Pathogenicity

<table>
<thead>
<tr>
<th>Category</th>
<th>Evidence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PVS1</td>
<td>Null variant (non-disease)</td>
<td>Beware of genetic consequences</td>
</tr>
<tr>
<td>PS1</td>
<td>Same amino acid change</td>
<td>Beware of changes</td>
</tr>
<tr>
<td>PS2</td>
<td>De novo (both maternal and paternal)</td>
<td>Beware of changes</td>
</tr>
<tr>
<td>PS3</td>
<td>Well-established</td>
<td>Functional studies</td>
</tr>
<tr>
<td>PS4</td>
<td>The prevalence of the variant</td>
<td>Beware of changes</td>
</tr>
</tbody>
</table>

## Strong (PS1-PS4)

<table>
<thead>
<tr>
<th>Category</th>
<th>Evidence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PM1</td>
<td>Located in a mutation</td>
<td></td>
</tr>
<tr>
<td>PM2</td>
<td>Absent from controls</td>
<td>Population data</td>
</tr>
<tr>
<td>PM3</td>
<td>For recessive disorders</td>
<td>Requires testing for both parents</td>
</tr>
<tr>
<td>PM4</td>
<td>Protein length change</td>
<td></td>
</tr>
<tr>
<td>PM5</td>
<td>Novel missense change</td>
<td>Beware of changes</td>
</tr>
<tr>
<td>PM6</td>
<td>Assumed de novo</td>
<td></td>
</tr>
</tbody>
</table>

## Moderate (PM1-PM6)

<table>
<thead>
<tr>
<th>Category</th>
<th>Evidence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PP1</td>
<td>Co-segregation with disease</td>
<td></td>
</tr>
<tr>
<td>PP2</td>
<td>Missense variant in a patient</td>
<td></td>
</tr>
<tr>
<td>PP3</td>
<td>Multiple lines of evidence</td>
<td></td>
</tr>
<tr>
<td>PP4</td>
<td>Patient’s phenotype</td>
<td></td>
</tr>
<tr>
<td>PP5</td>
<td>Reputable source reported</td>
<td></td>
</tr>
</tbody>
</table>

## Supporting (PP1-PP5)

<table>
<thead>
<tr>
<th>Category</th>
<th>Evidence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BP1</td>
<td>Missense variant</td>
<td></td>
</tr>
<tr>
<td>BP2</td>
<td>Observed in translocation pattern</td>
<td></td>
</tr>
<tr>
<td>BP3</td>
<td>In-frame deletion</td>
<td></td>
</tr>
<tr>
<td>BP4</td>
<td>Multiple lines of evidence</td>
<td></td>
</tr>
<tr>
<td>BP5</td>
<td>Variant found in a patient</td>
<td></td>
</tr>
<tr>
<td>BP6</td>
<td>Reputable source reported</td>
<td></td>
</tr>
<tr>
<td>BP7</td>
<td>A synonymous change found in an unrelated individual</td>
<td></td>
</tr>
</tbody>
</table>

## Evidence of Benign Impact

<table>
<thead>
<tr>
<th>Category</th>
<th>Evidence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BA1</td>
<td>Allele frequency</td>
<td></td>
</tr>
<tr>
<td>BS1</td>
<td>Allele frequency</td>
<td></td>
</tr>
<tr>
<td>BS2</td>
<td>Observed in a heterozygote at an expected frequency at an equal frequency in controls</td>
<td></td>
</tr>
<tr>
<td>BS3</td>
<td>Well-established</td>
<td>Natural history</td>
</tr>
<tr>
<td>BS4</td>
<td>Lack of segregation</td>
<td>Caveat: The presence of the variant contributing to the phenotype</td>
</tr>
</tbody>
</table>

**New England Newborn Screening Program**
**NENSP Variant Database Crow’s Foot Diagram**

### BabyList

<table>
<thead>
<tr>
<th>PK</th>
<th>BABY_SYSKEY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Candidate Disorder
- Diagnosis
- Date of Diagnosis
- Center Making Diagnosis
- Specialist Making Diagnosis
- Comments about FU or DX

### SpecimenFindingsAndContacts

<table>
<thead>
<tr>
<th>PK</th>
<th>ID</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FK</th>
<th>Guthrie Lab ID</th>
</tr>
</thead>
</table>

- Seq Run number
- Date Seq Result
- Date Seq Result ready for Reporting

<table>
<thead>
<tr>
<th>FK</th>
<th>Variant SyskeyA through AZ</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>FK</th>
<th>FK_BABY (BABY_SYSKEY)</th>
</tr>
</thead>
</table>

- Follow Up Comment
- Reported To
- Reported Date

### Variants

<table>
<thead>
<tr>
<th>PK</th>
<th>Variant Syskey</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- GENE
- c Change to report
- p change to report
- Date NENSP last reviewed varaint
- Current ClinSign Assigned by NENSP
- Date Current ClinSign Assigned by NENSP
- Comment for Report
- ACMG related fields (36 individual fields)
- Disease-Specific database related fields (5 individual fields)
- ClinVar related fields (7 individual fields)
- EMV (EGL) related fields (5 individual fields)
- gnomAD related fields (9 individual fields)
- ExAC related fields (7 individual fields)
- UCSC related fields (4 individual fields)
- dBSNP related fields (6 individual fields)
- Polyphen related fields (5 individual fields)
- PMID related fields (12 individual fields)
- ClinSign change related fields (7 individual fields)

*Babylist table and SpecimenFindingsAndContacts table are linkable to NENSP LIMS via BABY_SYSKEY and Guthrie Lab ID fields for baby and specimen demographic information.*
### Variants Table

<table>
<thead>
<tr>
<th>Variant Syn#</th>
<th>GEN#</th>
<th>c. Change to</th>
<th>p. Change to</th>
<th>DataEntry in VariantWorld</th>
<th>2nd Reader</th>
<th>Date of Init</th>
<th>RS#</th>
<th>RS# Source</th>
<th>Date NENSP</th>
<th>Current Clin</th>
<th>Date Current</th>
<th>NENSP Staff</th>
<th>Comment for Report</th>
<th>Pseudodef</th>
<th>CRIM Negati</th>
<th>Early Onset</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 66 GAA</td>
<td><em>N366A</em></td>
<td>Anne</td>
<td>Jaime</td>
<td>10/20/2017 r23800273</td>
<td>gnoMAD</td>
<td>10/4/2017 Benign</td>
<td>10/4/2017 Anne, Jaime, B observed in 1% of general pop</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 38 GAA</td>
<td><em>T310T</em></td>
<td>Anne</td>
<td>Jaime</td>
<td>10/19/2017 r23800273</td>
<td>gnoMAD</td>
<td>10/4/2017 Benign</td>
<td>10/4/2017 Anne, Jaime, B observed in 1% of general pop</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 27 GAA</td>
<td><em>G309A</em></td>
<td>Anne</td>
<td>Jaime</td>
<td>10/18/2017 r23800273</td>
<td>gnoMAD</td>
<td>10/4/2017 Benign</td>
<td>10/4/2017 Anne, Jaime, B observed in 1% of general pop</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Variants Table

<table>
<thead>
<tr>
<th>Variant Loca</th>
<th>Variant Type</th>
<th>Molecular Cc</th>
<th>ClinSign Cor</th>
<th>DiseaseDB E</th>
<th>DiseaseDB D</th>
<th>DiseaseDB C</th>
<th>DiseaseDB B</th>
<th>DiseaseDB A</th>
<th>ClinVar Clinical Significance</th>
<th>ClinVar Comment</th>
<th>ClinVar Date</th>
<th>ClinVar Check</th>
<th>ClinVar Vari</th>
<th>ClinVar Mech</th>
<th>EnvClass C1</th>
<th>EnvClass C2</th>
<th>EnvVar</th>
</tr>
</thead>
<tbody>
<tr>
<td>exon 20</td>
<td>Single Nucleot</td>
<td>3 prime UTR vs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>likely benign</td>
<td>likely benign b</td>
<td>6/14/2016</td>
<td>10/4/2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>exon 20 non co</td>
<td>Single Nucleot</td>
<td>3 prime UTR vs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>non pathogenic</td>
<td>non pathogenic</td>
<td>5/31/2016</td>
<td>10/6/2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon 8</td>
<td>Single Nucleot</td>
<td>Missense</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>pathogenic</td>
<td>pathogenic by</td>
<td>2/11/2016</td>
<td>11/20/2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon 7</td>
<td>Single Nucleot</td>
<td>Missense</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>pathogenic</td>
<td>pathogenic by</td>
<td>11/30/2012</td>
<td>11/30/2012</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon 6</td>
<td>Single Nucleot</td>
<td>Missense</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>benign</td>
<td>benign by 4 ex</td>
<td>2/13/2017</td>
<td>5/18/2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon 5</td>
<td>Single Nucleot</td>
<td>Missense</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>benign</td>
<td>by one no</td>
<td>10/6/2016</td>
<td>1/11/2018</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon 4</td>
<td>Single Nucleot</td>
<td>Missense</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>pathogenic</td>
<td>pathogenic by</td>
<td>7/11/2017</td>
<td>7/11/2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon 3</td>
<td>Single Nucleot</td>
<td>Missense</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>benign</td>
<td>benign by 4 ex</td>
<td>6/14/2016</td>
<td>10/4/2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intron 8</td>
<td>Deletion</td>
<td>Intron</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>presumably no</td>
<td>presumably no</td>
<td>5/31/2016</td>
<td>11/11/2018</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intron 7</td>
<td>Single Nucleot</td>
<td>Missense</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>benign</td>
<td>benign by 4 ex</td>
<td>6/14/2016</td>
<td>9/20/2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intron 6</td>
<td>Single Nucleot</td>
<td>Missense</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>benign</td>
<td>benign by 3 inc</td>
<td>6/14/2016</td>
<td>9/25/2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intron 5</td>
<td>Single Nucleot</td>
<td>Missense</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>benign</td>
<td>benign by 3 inc</td>
<td>6/14/2016</td>
<td>9/25/2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intron 4</td>
<td>Single Nucleot</td>
<td>Missense</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>benign</td>
<td>benign by 3 inc</td>
<td>6/14/2016</td>
<td>9/25/2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intron 3</td>
<td>Single Nucleot</td>
<td>Missense</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>benign</td>
<td>benign by 3 inc</td>
<td>6/14/2016</td>
<td>9/25/2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intron 2</td>
<td>Single Nucleot</td>
<td>Missense</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>benign</td>
<td>benign by 3 inc</td>
<td>6/14/2016</td>
<td>9/25/2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intron 1</td>
<td>Single Nucleot</td>
<td>Missense</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>benign</td>
<td>benign by 3 inc</td>
<td>6/14/2016</td>
<td>9/25/2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon 10</td>
<td>Single Nucleot</td>
<td>Syntonic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>benign</td>
<td>benign by 4 ex</td>
<td>6/14/2016</td>
<td>10/4/2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon 9</td>
<td>Single Nucleot</td>
<td>Synomnous</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>benign</td>
<td>benign by 3 inc</td>
<td>6/14/2016</td>
<td>10/4/2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon 8</td>
<td>Single Nucleot</td>
<td>Synomnous</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>benign</td>
<td>benign by 3 inc</td>
<td>6/14/2016</td>
<td>10/4/2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon 7</td>
<td>Single Nucleot</td>
<td>Synomnous</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>benign</td>
<td>benign by 3 inc</td>
<td>6/14/2016</td>
<td>10/4/2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon 6</td>
<td>Single Nucleot</td>
<td>Synomnous</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>benign</td>
<td>benign by 3 inc</td>
<td>6/14/2016</td>
<td>10/4/2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon 5</td>
<td>Single Nucleot</td>
<td>Synomnous</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>benign</td>
<td>benign by 3 inc</td>
<td>6/14/2016</td>
<td>10/4/2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon 4</td>
<td>Single Nucleot</td>
<td>Synomnous</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>benign</td>
<td>benign by 3 inc</td>
<td>6/14/2016</td>
<td>10/4/2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon 3</td>
<td>Single Nucleot</td>
<td>Synomnous</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>benign</td>
<td>benign by 3 inc</td>
<td>6/14/2016</td>
<td>10/4/2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon 2</td>
<td>Single Nucleot</td>
<td>Synomnous</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>benign</td>
<td>benign by 3 inc</td>
<td>6/14/2016</td>
<td>10/4/2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon 1</td>
<td>Single Nucleot</td>
<td>Synomnous</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>benign</td>
<td>benign by 3 inc</td>
<td>6/14/2016</td>
<td>10/4/2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Variant Entry Form Part 1

Variants detected by NENSP

Variant Syskey: 30  GENE: GAA
Date of Initial Detection: 10/11/2017
Data Entry in VariantWorld: Anne

Change to report: 2238G>C  p. change to report: Trp745Cys

Do not enter c. or p. Only enter nucleotide change or protein change using the 3 letter amino acid abbreviations.

NENSP Current Assignment of Clinical Significance

Current ClinSign Assigned by NENSP: Pathogenic
Date Current ClinSign Assigned by NENSP: 10/20/2017
NENSP last reviewed variant: 10/20/2017
NENSP staff who determined clinical significance: Anne, Jaime, Binod, Neela

Comment for Report: *Text in “Comment for Report” box will appear directly on the report*

NENSP Assignment of ACMG Criteria

Number ACMG PV1: 1  ACMG PV1
Number of ACMG PS1-PS4: 2  ACMG PS1  ACMG PS2  ACMG PS3  ACMG PS4
Number of ACMG PM1-PM6: 1  ACMG PM1  ACMG PM2  ACMG PM3  ACMG PM4  ACMG PM5
Number of ACMG PP1-PPS: 2  ACMG PP1  ACMG PP2  ACMG PP3  ACMG PP4  ACMG PP5
Number ACMG BA1: 1  ACMG BA1
Number ACMG BS1-BS4: 1  ACMG BS1  ACMG BS2  ACMG BS3  ACMG BS4
Number ACMG BP1-BP7: 1  ACMG BP1  ACMG BP2  ACMG BP3  ACMG BP4  ACMG BP5  ACMG BP6  ACMG BP7

ACMG comments:

Has the ClinSign Ever Changed? If yes, see below for details on previous classifications

Variant Summary

RS #: rs1800312
RS# Source: gnomAD
Variant Location: exon 16
Variant Type: Single Nucleotide Variant
Molecular Consequence: Missense

- Associated with Pseudodeficiency
- Associated with CRIM Negative status
- Associated with ONLY Early Onset Disease
- Associated with ONLY Late Onset Disease

Variant Comment:
Variant Entry Form Part 1

**Variants detected by NENSP**

- **Variant Syskey:**
- **GENE:** GAA
- **Date of Initial Detection:** 10/11/2017
- **Data Entry in VariantWorld:** Anne
- **2nd Reader of DataEntry in VariantWorld:** Jaime

**NENSP Current Assignment of Clinical Significance**

- **Current ClinSign Assigned by NENSP:** Pathogenic
- **Date Current ClinSign Assigned by NENSP:** 10/20/2017
- **Date NENSP last reviewed variant:** 10/20/2017
- **NENSP staff who determined clinical significance:** Anne, Jaime, Binod, Neela

**Variant Summary**

- **RS #:** rs1800312
- **RS# Source:** gnoMAD
- **Variant Location:** exon 16
- **Variant Type:** Single Nucleotide Variant
- **Molecular Consequence:** Missense

- **Variant Comment:**

**ACMG Criteria**

- **Number ACMG PV1:**
- **Number of ACMG PS1-PS4:**
- **Number of ACMG PM1-PM6:**
- **Number of ACMG PP1-PP3:**

**ACMG comments:**

- **Has the ClinSign Ever Changed?**

*Text in "Comment for Report" box will appear directly on the report*

New England Newborn Screening Program
Variant Entry Form Part 2

Data from consulted databases

DiseaseDB date last updated: 5/31/2016

DiseaseDB Effect: potentially mild

The field DiseaseDB Effect is listed in the Disease DBs as follows:
Pompe DB: Effect  Pompe DiseaseDB
ALD DB: Remark  ALD DiseaseDB
MPS1: ptype  MPS1 DiseaseDB

DiseaseDB Comment:
Variant Entry Form Part 2

Data from consulted databases

- DiseaseDB
- EmVClass
- gnoMAD
- ExAC
- ClinVar
- dbSNP
- UCSC
- Polyphen

- Reported ClinSign varies between DBs

EmVClass Classification: Pathogenic
EmVClass date last updated: 6/15/2017

EmVClass Comment:

- EmVClass checked NOT LISTED
**Variant Entry Form Part 2**

**Data from consulted databases**

- **gnoMAD Allele Frequency:** 0.00001992
- **gnoMAD allele freq warning:**
- **gnoMAD comment:**
- **gnoMAD PolyPhen Prediction:** Possibly Damaging
- **gnoMAD SIFT Prediction:** Deleterious
## PMID Citations consulted [Link to PubMed](#)

<table>
<thead>
<tr>
<th>PMID</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>18458862</td>
<td>homo juvenile onset not found in 100 healthy</td>
</tr>
<tr>
<td>21757382</td>
<td>the isolated p.W746C mutation displayed 8% of normal GAA activity in transfected fibroblast using</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Specimen Variant Entry Form

Enter Parameter Value

Guthrie# in format AA0101201234

|   |

OK Cancel
Specimen Variant Entry Form

- Baby’s demographic data from our current LIMS
- Biochemical NBS results
- Eventual Diagnosis

New England Newborn Screening Program

University of Massachusetts Medical School umassmed.edu
Report to Healthcare Provider
Report Attachment of Sequencing Results for Gene Associated with MPS1

<table>
<thead>
<tr>
<th>Name of Gene Sequenced:</th>
<th>IDUA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name of Variant(s) Detected</td>
<td>Reference Gene ID: NG_008103.1</td>
</tr>
<tr>
<td>c.386-2A&gt;G</td>
<td></td>
</tr>
<tr>
<td>c.246C&gt;G p.(His82Gln)</td>
<td></td>
</tr>
<tr>
<td>c.352C&gt;T p.(Leu118Leu)</td>
<td>Benign</td>
</tr>
<tr>
<td>c.352C&gt;T p.(Leu118Leu)</td>
<td>Benign</td>
</tr>
</tbody>
</table>

There are no other variants detected. Please see additional information below.
## Example Queries

<table>
<thead>
<tr>
<th>Variant Sysk</th>
<th>Date of Initiation</th>
<th>GENE</th>
<th>cChange to report</th>
<th>pChange to report</th>
<th>Reported Variant</th>
<th>Current Clin</th>
<th>Clin Sign</th>
</tr>
</thead>
<tbody>
<tr>
<td>53</td>
<td>9/20/2017</td>
<td>GAA</td>
<td>*3G&gt;A</td>
<td></td>
<td>c.*3G&gt;A</td>
<td>Benign</td>
<td>observer</td>
</tr>
<tr>
<td>58</td>
<td>10/3/2017</td>
<td>GAA</td>
<td>*91G&gt;A</td>
<td></td>
<td>c.*91G&gt;A</td>
<td>Benign</td>
<td>observer</td>
</tr>
<tr>
<td>77</td>
<td>10/18/2017</td>
<td>GAA</td>
<td>1000G&gt;T</td>
<td></td>
<td>c.1000G&gt;T</td>
<td>Uncertain</td>
<td></td>
</tr>
<tr>
<td>107</td>
<td>GAA</td>
<td></td>
<td>1082C&gt;T</td>
<td>Pro361Leu</td>
<td>c.1082C&gt;T p.(Pro361Leu)</td>
<td>Likely Pathoge</td>
<td></td>
</tr>
<tr>
<td>108</td>
<td>GAA</td>
<td></td>
<td>1106T&gt;C</td>
<td>Leu369Pro</td>
<td>c.1106T&gt;C p.(Leu369Pro)</td>
<td>Likely Pathoge</td>
<td></td>
</tr>
<tr>
<td>105</td>
<td>GAA</td>
<td></td>
<td>1171A&gt;G</td>
<td>Met391Val</td>
<td>c.1171A&gt;G p.(Met391Val)</td>
<td>Uncertain Sign</td>
<td></td>
</tr>
<tr>
<td>199</td>
<td>3/22/2019</td>
<td>GAA</td>
<td>1226A&gt;T</td>
<td>Asp409Val</td>
<td>c.1226A&gt;T p.(Asp409Val)</td>
<td>Uncertain Sign</td>
<td></td>
</tr>
</tbody>
</table>

Record: 1 of 95

University of Massachusetts Medical School

umassmed.edu
Example Queries

~32% of variants entered to date have been benign

~49% of variants entered to date have been VOUS

~32% of variants entered to date have been benign
~49% of variants entered to date have been VOUS

~32% of variants entered to date have been benign
30 of 89 (~34%) VOUS are VOUS due to the absence of any available data.
Example Queries

10 of 89 (~11%) VOUS are VOUS due to conflicting data.

Interpretations reported in consulted sources conflict.

# of ACMG Pathogenic criteria by category

# of ACMG Benign criteria by category
19 of 58 benign variants entered to date have a population frequency of $\geq 50\%$. 
Example Queries

SQL query for list of all PMID Numbers

Totals Query
Example Queries

**SQL query for list of PMID Numbers**

```
SELECT [PMID1] AS PMIDs FROM Variants WHERE [PMID1] Is Not Null UNION ALL
SELECT [PMID2] AS PMIDs FROM Variants WHERE [PMID2] Is Not Null UNION ALL
```

**Totals Query**

![Totals Query Output](image-url)
A cross-sectional single-centre study on the spectrum of Pompe disease, German patients: molecular analysis of the GAA gene, manifestation and genotype-phenotype correlations.


Abstract

BACKGROUND: Pompe disease (Glycogen storage disease type II, GSD II, acid alpha-glucosidase deficiency, acid maltase deficiency, OMIM # 252300) is an autosomal-recessive lysosomal storage disorder due to a deficiency of acid alpha-glucosidase (GAA, acid maltase, EC 3.2.1.20, SwissProt P10253). Clinical manifestations are dominated by progressive weakness of skeletal muscle throughout the clinical spectrum. In addition, the classic infantile form is characterised by hypertrophic cardiomyopathy.

METHODS: In a cross-sectional single-centre study we clinically assessed 3 patients with classic infantile Pompe disease and 39 patients
**Variant Summary Report**

*New England Newborn Screening Program*

**NENSP Variant Summary Sheet**

**GENE:** GAA  
**Reported Variant:** c.2238G>C p.(Trp746Cys)  
**Interpretation:** Pathogenic

**General Data**

- **RS ID:** rs1800312  
- **Variant Type:** Single Nucleotide Variant  
- **Molecular Consequence:** Missense  
- **Variant Location:** exon 16  
- **Record:** 1 of 1

**ACMG Data**

**Pathogenic**

- **Number PVS1:** 0  
- **Number PS1-PS4:** 0  
- **Number PM1-PM6:** 1  
- **Number PP1-PP5:** 0

**Benign**

- **Number BA1:** 0  
- **Number BS1-BS4:** 0  
- **Number BP1-BP7:** 0

**Key Data from Consulted Databases**

- **DiseaseDB Effect:** potentially mild  
- **EmVClass Classification:** Pathogenic  
- **ClinVar Clinical Significance:** pathogenic  
- **gnomAD Allele Freq:** 0.00003034  
- **dbSNP Allele Freq:** 0.0004  
- **UCSC Conservation:** amino acid conserved across all species  
- **Polyphen Prediction:** probably damaging

**Record:** 1 of 1
ABSTRACT

As per American College of Medical Genetics (ACMG) guidelines, all detected sequence variants are classified and reported as one of the following: pathogenic, likely pathogenic, uncertain significance, likely benign, or benign. For each identified variant, the following types of resources are routinely consulted prior to determination of final classification: disease-specific databases, large clinical outcome databases, large high-throughput sequencing databases, online functional evaluations, and peer-reviewed publications.

Review of specified data from each resource is time-consuming and subject to human error. In addition, conflicting interpretations can result due to inconsistencies between resources. We present here a framework designed to streamline such a data collection process with automation.

Newborn Screening Knowledge Framework

- **Sample Acquisition**
- **Bioinformatics Pipelines**
- **Variant Tracking & Information Aggregation**
- **Knowledge Management**
- **Report Generation**
- **Data Analytics**

**Integartion**

Bioinformatics tools and utilities to help screening programs adapt to and benefit from new forms of sequencing technologies.

**Fact Lookup**

A flexible framework that can work with a screening program’s existing technology systems and infrastructures.

**Change Management**

A highly curated knowledge map that can help screening programs analyze and discover new insights.

**Reports Delivery**

Document generation that makes it easier to create and deliver screening results to clinicians and a central CDS database.

Authors:

- Saurabh Singh, Michael Patinelli
- Joon Hwang, Benedict Caire
- Jane Cooper

New England Newborn Screening Program

UMass Medical School

umassmed.edu
Acknowledgements

Anne Marie Comeau, PhD
Deputy Director, NENSP
Director, NENSP Molecular Laboratory
Professor of Pediatrics, UMass Medical School

Binod Kumar, PhD
Molecular Laboratory Scientist, NENSP

Roger B. Eaton, PhD
Director, NENSP
Associate Professor of Pediatrics, UMass Medical School

Inderneel Sahai, MD
Chief Medical Officer, NENSP
Assistant Professor of Pediatrics, UMass Medical School

Jomol Mathew, PhD
Chief Research Informatics Officer, UMass Medical School
Data Sciences & Technology, UMass Medical School
Assistant Professor of Population and Quantitative Health Sciences, UMass Medical School

Mahesh Vangala, MS
Data Sciences & Technology, UMass Medical School

Keith Pelletier, BA
Data Sciences & Technology, UMass Medical School

CDC Grant Number 1 NU88EH001323-01-00: Newborn Screening New Condition Implementation: Capacity Building and Quality Improvement through Data Harmonization

New England Newborn Screening Program