Phenotype Discussion
Summary

Josh Denny / Marylyn Ritchie
Key points

• eMERGE has pioneered use of phenotyping in the EHR, and is a model for other networks repurposing EHRs

• In general phenotype creation is still hard though has accelerated some (e1=14, e2=29, e3=27)

• PheKB has 154 mostly rule-based phenotypes, 75 have (already) been attributed to eMERGE

• Use of common data models and phenotype languages/models (OMOP, FHIR) should accelerate phenotype translation across sites

• Machine learning represents an opportunity to accelerated some but still require gold standard to train and portability has not been as robustly demonstrated as rule-based algorithms
Phenotypes in PheKB now

<table>
<thead>
<tr>
<th></th>
<th>Public (n = 44)</th>
<th>Non-public (n = 110)</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICD-9 or -10 codes</td>
<td>39</td>
<td>73</td>
<td>73%</td>
</tr>
<tr>
<td>Medications</td>
<td>31</td>
<td>51</td>
<td>53%</td>
</tr>
<tr>
<td>CPT codes</td>
<td>23</td>
<td>44</td>
<td>44%</td>
</tr>
<tr>
<td>NLP</td>
<td>28</td>
<td>36</td>
<td>42%</td>
</tr>
<tr>
<td>Laboratory test results</td>
<td>21</td>
<td>37</td>
<td>38%</td>
</tr>
<tr>
<td>Vital signs</td>
<td>5</td>
<td>14</td>
<td>12%</td>
</tr>
</tbody>
</table>

GWAS discovery in innovative phenotypes: MACE on Statins identified locus independent of ΔLDL

Clustering phenotypes: phenotype risk scores & mining for human phenotype ontology

\[
PRS_i = \sum_{j=1}^{k} \begin{cases} 1 \log \frac{n_{total}}{n_j} \
0 \end{cases}, \text{ where } j= \text{ Mendelian gene phenotypes}
\]
Key points - 2

• There is a tradeoff: complicated phenotypes that take more time vs. simpler algorithms we can extend. Where is the greatest value for eMERGE?

• Multimodal phenotypes and use of text records/NLP are a hallmark of many eMERGE phenotypes

• Long history of phenotype innovation - pharmacogenomic, longitudinal phenotypes, OCR, portable NLP modules, KNIME, deeper phenotyping, PheWAS, phenotype risk scores

• Where does sequencing take us for EHR utility?