What do public databases really tell us about classifications of variants in BRCA1 and BRCA2?

Stephen E. Lincoln\(^1\), Shan Yang\(^1\), Melissa S. Cline\(^2\), Can Zhang\(^2\), Yuya Kobayashi\(^1\), Scott Topper\(^1\), David Haussler\(^2\), Benedict Paten\(^2\), and Robert L. Nussbaum\(^1,3\)

\(^1\)Invitae, San Francisco, CA; \(^2\)University of California, Santa Cruz; \(^3\)University of California, San Francisco

Methods

Classifying which genetic variants observed in a patient are pathogenic (and thus potentially actionable) and which are not is complex and requires expert judgment by laboratory directors who must weigh various lines of evidence in determining classifications. For this reason variant classifications sometimes disagree among laboratories, even when contemporaneous. Our study focused on the commonly tested BRCA1 and BRCA2 genes, in which germline pathogenic variants substantially increase the risk of breast, ovarian, and other cancers, and where finding a pathogenic variant in a patient may change clinical management decisions significantly. We observed 99.8% concordance between discordant classifications have been suggested as a reason that public databases “should be precluded from clinical use.” Although disagreements are real, this proposition is inconsistent with our experience. In a prior study\(^2\) of 975 patients, we observed 99.8% concordance between BRCA1/2 reports, (a) that were produced using publicly available data, the literature, and the most recent guidelines, and (b) classifications from Myriad, who had used their proprietary database and slightly different criteria. Moreover, to the extent that discordances do exist in public databases, they enable

- Interlaboratory quality control,
- Building consensus among experts for specific variants, and
- Improving procedures for variant classification

Results

We compared classifications on the basis of whether they would change management—i.e., pathogenic and likely pathogenic were considered discordant from benign, likely benign, or VUS (variant of uncertain significance). On a per-variant basis, we found high concordance: only 30 variants (1.5%) showed discordance between any two submitters. The largest class of discordant variants were rare missense changes (18/30), which composed almost half (45.5%) of our data set, although the vast majority of such variants (98.0%) had concordant classifications among all submitters. Although numerous, rare missense variants are, as a class, infrequently observed in patients (6.4% prevalence). Other discordances were in canonical RNA splice sites (5/30) or an intron (2/30) or were in-frame deletions (2/30). We contacted the directors of submitting laboratories to help resolve discordances.

### Methods

1. BRCA1/2 data were extracted from the May 2016 ClinVar XML release.
2. Analysis was limited to submissions that met objective criteria, namely
   - They came from established clinical laboratories,
   - They had at least 200 BRCA1/2 classifications in ClinVar,
   - They were mostly recent (>50% were <5 years old)
3. Data integration was improved by standardizing variant nomenclature. Data were subjected to quality control, and clearly erroneous records removed.
4. Comparisons considered only differences that would significantly change management decisions under current guidelines.
5. We also collected a sequential series of 30,000 patients tested for BRCA1/2 to measure the prevalence of variants by type.

### ClinVar submitter

<table>
<thead>
<tr>
<th>ClinVar submitter</th>
<th>Classified variants</th>
<th>Comparable variants</th>
<th>Most recent classification</th>
<th>Evidence provided</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ambry</td>
<td>2792</td>
<td>1613</td>
<td>Feb 2015</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>SCRP/Myriad</td>
<td>2327</td>
<td>1531</td>
<td>Dec 2015</td>
<td>No</td>
<td>Likely benign variants generally unreported.</td>
</tr>
<tr>
<td>Invitae</td>
<td>1998</td>
<td>1367</td>
<td>Mar 2016</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>GeneDx</td>
<td>1216</td>
<td>957</td>
<td>Oct 2015</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Counsyl</td>
<td>272</td>
<td>256</td>
<td>Feb 2015</td>
<td>No</td>
<td>No variants of uncertain significance submitted</td>
</tr>
<tr>
<td>CHEO</td>
<td>257</td>
<td>220</td>
<td>N/A</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>Emory</td>
<td>203</td>
<td>183</td>
<td>June 2015</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>5124</strong></td>
<td><strong>2006</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Legend:** Comparable variants are those with two or more submitters. There are known biases in the data (noted here). Some submitters had not provided updates for more than one year (bold), and some submitters excluded certain classes of variants (see notes). Most submitters unfortunately do not provide classification evidence for specific variants. Copy number variants were also only reported by only a few submitters (not shown).

Clinical prevalence is the fraction of patients (out of 30,000) with any such variants. AF is allele frequency.

Given that the variant types with discordance are rare and low prevalence, the per-patient discordance is estimated to be substantially higher than the observed per-variant discordance.

### Conclusion

Although our analysis shows that clinically significant disagreements in BRCA1/2 variant classification are infrequent, discordances are of course important to patients and clinicians. We believe it is essential for the genetics community to resolve these differences collaboratively, as is standard practice in other areas of medicine, in order to deliver the best possible patient care. The open sharing of data through ClinVar affords us the opportunity to do this on a global scale.