Analytic Validation Standards for NGS Assays and New Reference Materials Needed

Jared Maguire, Counsyl
Steve Lincoln, Invitae
Lisa Kalman, CDC

Disclosure

- I am an employee of Invitae, a laboratory offering clinical genetic testing services
- I own stock in various biotech companies including Illumina
- My opinions are my own and are not necessarily those of my employer
### Traditional Validation of a Genetic Test

- Focus on clinically relevant variants
- Want a positive control for each reportable variant
  - Patient
  - Synthetic
- Want a negative control for each gene/region
- Do replicates, etc.

### Methods Based Validation – The Assumptions

- It's **not possible** to evaluate NGS assay on every possible variant, or in every gene/region which most assays can report (true for panels, exomes, and genomes)
- We understand a lot about the **failure modes of NGS**, in particular that many challenges are **systematic** (not random) and they apply to:
  - Specific classes of variants
  - Specific types of regions
- Analytic performance and clinical interpretation are generally independent
Perform analytic validation on a technically representative set of variants…

…in a technically representative set of genes and regions…

…under uniform data QC criteria.

Extrapolate to other genes/regions/variants which meet the same QC criteria.

The Big Question

What is adequately representative?
A systematic comparison of traditional and multi-gene panel testing for hereditary breast and ovarian cancer in more than 1000 patients

Stephen E. Lincoln1, Yuya Kobayashi1, Michael J. Anderson1, Shan Yang1, Andrea J. Desmond2, Meredith A. Mills3, Geoffrey B. Nilsen1, Kevin B. Jacobs1, Federico A. Monzon1, Allison W. Kurian3, James M. Ford3, Leif W. Ellisen2,4

1. Invitae, San Francisco, CA
2. Massachusetts General Hospital Cancer Center, Boston, MA
3. Stanford University School of Medicine, Stanford, CA
4. Harvard Medical School, Boston, MA

Lincoln et al., J Mol Diag 2015

Hereditary Cancer Panel Validation Study

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Description</th>
<th>Previous Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prospective Clinical</td>
<td>735</td>
<td>Prospective clinical cases</td>
<td>Clinical testing for BRCA1/2, occasionally other genes (depending on case) using traditional methods</td>
</tr>
<tr>
<td>High-Risk Clinical (Total 327)</td>
<td>209</td>
<td>Retrospective cases from a clinical biobank, generally containing higher-risk individuals</td>
<td>Clinical single-site testing</td>
</tr>
<tr>
<td></td>
<td>118</td>
<td>Cases referred due to known pathogenic variant in family</td>
<td>Samples carry known pathogenic variants</td>
</tr>
<tr>
<td>Reference Samples</td>
<td>36</td>
<td>Reference samples from public biobanks (Coriell, NIBSC)</td>
<td>Variants in 29 cancer genes extracted from WGS data; most of these are benign</td>
</tr>
<tr>
<td>Well-Characterized Genomes (WCGs)</td>
<td>7</td>
<td>Reference samples with high-quality whole genome sequencing (WGS) data</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1105</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Lincoln et al., J Mol Diag 2015
Clinical Actionability Study

JAMA Oncology

Original Investigation

Clinical Actionability of Multigene Panel Testing for Hereditary Breast and Ovarian Cancer Risk Assessment

Andrea Desmond, BS; Allison W. Kurtan, MD, MS; Michele Gabere, MS, CGC; Meredith A. Wilke, BA; Michael J. Anderson, PhD; Yuga Kobayashi, PhD; Nora Honoki, MS; Shan Yang, PhD; Kristen M. Shannon, MS, CGC; Nadine Tung, MD; James M. Ford, MD; Stephen E. Lincoln, BS; Leif W. Elsner, MD, PhD

Usefulness of Multigene Testing
Catching the Train That’s Left the Station

Elizabeth M. Swisher, MD

Desmond et al., JAMA Oncol. 2015
Swisher, JAMA Oncol. 2015

Analytic validity of NGS in N=1105 individuals

<table>
<thead>
<tr>
<th>NGS vs. Traditional Methods In 1105 Individuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
</tr>
<tr>
<td>Specificity</td>
</tr>
</tbody>
</table>

750 Comparable Variants (Pathogenic or Otherwise)

- Sequence Changes: 721
- Del/dups (CNVs): 29
- Single Nucleotide: 549
- Small Indel: 156
- Large Indel*: 13
- Complex**: 6

To achieve this, specialized NGS methods, biochemical and bioinformatics, are required.

The most challenging classes of variation tend to be not well represented in other validation studies.

* Large Indel is deletion≥10bp, insertion≥5bp
** Complex includes homo-polymer associated variants, indels in low-complexity sequence, short range haplotypes, etc.

Lincoln et al., J Mol Diag 2015
A Significant Fraction of Pathogenic Variants in Clinical Cases are Technically Challenging

Pathogenic and likely pathogenic variants (n=260) among the clinical cases (n=1062) by variant type:

- 34.2% SNV
- 4.6% Del/dup Multi-exon
- 3.8% Del/dup Single-exon
- 3.5% Large Indel
- 1.5% Complex

Lincoln et al., J Mol Diag 2015

A different study: Analytic validity in 250 individuals

<table>
<thead>
<tr>
<th>NGS vs. Traditional Methods in 250 Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
</tr>
<tr>
<td>Specificity</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3025 Variants Appropriate to Measure Sensitivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single Nucleotide</td>
</tr>
<tr>
<td>Small Indel</td>
</tr>
<tr>
<td>Large Indel *</td>
</tr>
<tr>
<td>Complex **</td>
</tr>
</tbody>
</table>

Note: These numbers assume that Table 2 includes all of the indels and del/dups, and that the remaining variants in the are benign SNPs. This could have been worded more clearly in the paper.

* Large Indel is deletion≥10bp, insertion≥5bp
** Complex includes homo-polymer associated variants, indels in low-complexity sequence, short range haplotypes, etc.

Chong et al., PLOS One 2014
BRCA2 c.9203del126

- Split-read signal at 5' end of deletion
- Split-read signal at 3' end of deletion

Exon target

BRCA2 c.156_insAlu

- Split-read signal of Alu sequence
**MSH2 c.943+3T>C**

- Alignment and Biochemical Artifacts
- Homopolymer-A

**CDKN2A c.9_32dup24**

- Insertion of repeat in correctly mapped NGS reads
- Split-read signal
- Split-read signal
- Translation
- 5' Met
Reference Samples (non-genome) in Validation Study

- **Coriell**
  - 35 samples available
  - Known variants in 9 of 29 genes
  - 7 not available or failed our inbound QC
    - Some of which would have been very useful!

- **NIBSC (UK)**
  - One set of 7 samples
  - Added 2 genes
    - MLH1, MSH2
  - Added most of the CNVs

<table>
<thead>
<tr>
<th>Gene</th>
<th>Variants</th>
</tr>
</thead>
<tbody>
<tr>
<td>APC</td>
<td>1</td>
</tr>
<tr>
<td>BRCA1</td>
<td>20</td>
</tr>
<tr>
<td>BRCA2</td>
<td>7</td>
</tr>
<tr>
<td>CDKN2A</td>
<td>1</td>
</tr>
<tr>
<td>FANCC</td>
<td>2</td>
</tr>
<tr>
<td>MEN1</td>
<td>1</td>
</tr>
<tr>
<td>MLH1</td>
<td>2</td>
</tr>
<tr>
<td>MSH2</td>
<td>5</td>
</tr>
<tr>
<td>NF1</td>
<td></td>
</tr>
<tr>
<td>PTEN</td>
<td>1</td>
</tr>
<tr>
<td>RET</td>
<td>2</td>
</tr>
</tbody>
</table>

Lincoln et al., J Mol Diag 2015

Patient Reference Samples in JMD Study

<table>
<thead>
<tr>
<th>Gene</th>
<th>Variants</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNV</td>
<td>16</td>
</tr>
<tr>
<td>Indel or delins &lt; 5 bp</td>
<td>17</td>
</tr>
<tr>
<td>Indel or delins ≥ 5 bp *</td>
<td>4</td>
</tr>
<tr>
<td>SNV (homopolymer associated)</td>
<td>1</td>
</tr>
<tr>
<td>Single exon deletion</td>
<td>2</td>
</tr>
<tr>
<td>Single exon duplication</td>
<td>1</td>
</tr>
<tr>
<td>Multi exon deletion</td>
<td>2</td>
</tr>
<tr>
<td>Multi exon duplication</td>
<td>0</td>
</tr>
</tbody>
</table>

* BRCA1:del40, CDKN2A:del19, BRCA1:del11, and MEN1:ins5

Lincoln et al., J Mol Diag 2015
Whole Genome Reference Samples in JMD Study

- NA12878 and 6 other **Well-Characterized Genomes** (WCGs) were used
- The 7 WCGs contributed **310 of 750** comparable variants to both the sensitivity and specificity analyses
- But… the 77% coverage of GIAB data was a substantial limitation
  - **No** exonic variants in 5 of 29 panel genes in any of 7 samples
    - Only 1 coding variant each in 2 other genes
    - Reason: (a) missing 23% of GIAB and (b) population genetics
  - **Almost all** GIAB variants are simple SNVs
    - Only 6 of 310 were very small deletions (max 4bp)
    - 0 insertions, 0 other variant types
    - No GIAB CNV data yet (but we’d expect 0 CNVs in these 29 genes)
  - The 77% is **biased** to the “easy” subset of the genome

Lincoln et al., J Mol Diag 2015

There’s Lies, Damned Lies and Statistics*

- Imagine this hypothetical validation study
  - Test genes/exons of medical relevance in GIAB sample NA12878
  - Compare test results to GIAB reference data
  - Count concordance, report metrics
- Imagine an assay which fails to detect all “hard” variants, but which works highly accurately on the easy variants
- For the total spectrum of variants, both sensitivity and specificity will be over 99%, perhaps 99.9% or 99.99%
- **But on a patient sample there is a >10% chance of a false negative!**

*Mark Twain
Not everyone does analytic validation well

- We’ve all seen validation studies that:
  - Include sensitivity analysis only for “easy” variants and/or regions, even when the clinical spectrum is known to include many “hard” ones
    - Over-reliance on the Genome in a Bottle standard
  - Do not include confidence intervals
    - Why? Maybe they have an embarrassingly small N…
  - Use confirmation data to measure sensitivity (!)
  - Confuse reproducibility and accuracy
  - Use confusing or uninformative metrics
  - Suffer from statistical over-fitting

Reference

HOW TO LIE WITH STATISTICS
Darrell Huff
Illustrated by Irving Geis

Over Half a Million Copies Sold. An Honest-to-Goodness Bestseller
Potential Goal of the GET-RM Effort for NGS

1. Make it **much easier** to do a “good” analytic validation study.

2. Maybe… make it easier to identify studies which are not quite as adequate