Power and instrument strength requirements for Mendelian randomization studies using multiple genetic variants

Tyler J. VanderWeele
Departments of Epidemiology and Biostatistics
Harvard School of Public Health
Mendelian Randomization (MR)
Genes as instrumental variables

• **Goal of MR:** estimate causal effect of X on Y, using data on a known genetic determinant (G) of X

• Need data on G, X, and Y
  – But not U (if assumptions hold)

• Key MR assumptions: The IV (G) is
  – (1) associated with X
  – (2) independent of X-Y confounders (U)
  – (3) independent of Y given X and U (X-Y confounders)
What will be the implications for Mendelian Randomization of...

1. More potential IVs
2. Using knowledge of genetic architecture

**Goal of this work:**
Evaluate power and IV strength requirements for MR studies that utilize multiple variants
Methods for Simulations

• Simulate data on a gene (G), exposure (X), and outcome (Y)
  – G is biallelic, X and Y ~N(0,1)
  – With and without confounding (U)
  – 10,000 datasets

• Vary parameters:
  – \( f(G), \beta_{gx}, \beta_{xy}, n, \beta_{ux}, \beta_{uy} \)

• Two-stage least squares regression (2SLS)
  – Stage 1: regress X on G (the IV)
  – Stage 2: regress Y on fitted X values
  – Equivalent to the “Wald Estimator”

Wald Estimator:

\[ \hat{\beta}_{MR} = \frac{\hat{\beta}_{gy}}{\hat{\beta}_{gx}} \]

• Retain F statistic from the first-stage regression
  – Rule of thumb: A “strong IV” has F>10 (Stock et al, 2002)
  – Weak IVs bias towards the confounded OLS estimate

• Retain estimates and p-values from stage 2, determine empirical power
Power estimates for MR with a single IV

<table>
<thead>
<tr>
<th>(\beta_{gx})</th>
<th>(f(g) = 0.1)</th>
<th>(f(g) = 0.3)</th>
<th>(f(g) = 0.5)</th>
<th>(f(g) = 0.1)</th>
<th>(f(g) = 0.3)</th>
<th>(f(g) = 0.5)</th>
<th>(f(g) = 0.1)</th>
<th>(f(g) = 0.3)</th>
<th>(f(g) = 0.5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\beta_{xy})</td>
<td>(0.1)</td>
<td>(0.3)</td>
<td>(0.5)</td>
<td>(F)</td>
<td>(0.1)</td>
<td>(0.3)</td>
<td>(0.5)</td>
<td>(F)</td>
<td>(0.1)</td>
</tr>
<tr>
<td>(0.1)</td>
<td>1.8</td>
<td>0.00</td>
<td>0.02</td>
<td>0.06</td>
<td>9.2</td>
<td>0.02</td>
<td>0.13</td>
<td>0.31</td>
<td>17.8</td>
</tr>
<tr>
<td>(0.3)</td>
<td>16.1</td>
<td>0.04</td>
<td>0.24</td>
<td>0.52</td>
<td>81.8</td>
<td>0.15</td>
<td>0.78</td>
<td>(0.98)</td>
<td>&gt;100</td>
</tr>
<tr>
<td>(0.5)</td>
<td>46.2</td>
<td>0.10</td>
<td>0.51</td>
<td>0.88</td>
<td>&gt;100</td>
<td>0.30</td>
<td>(0.99)</td>
<td>1.00</td>
<td>&gt;100</td>
</tr>
</tbody>
</table>

\[\begin{align*}
G & \xrightarrow{\beta_{gx}} X \\
X & \xrightarrow{\beta_{xy}} Y
\end{align*}\]
Power estimates for MR with a single IV

Well-powered scenarios (>80%) and weak IV scenarios (F<10) are mutually exclusive (for this single IV scenario)

Well-powered studies will typically need to be quite large for modest effects
Multi-IV MR

• Multiple IVs in the 2SLS regression

• $R^2$ as summary measure of the effects of $G$s on $X$

• If $n$ and $\beta_{xy}$ are held constant, $R^2$ determines power
Using many IVs will result in low F values, resulting in an MR estimate that is biased towards the confounded association.
Reducing the number of IVs

• Combined IVs could take several forms:
  – allele count
  – weighted allele count
  – major gene/polygene

• **Goal:** maximize $R^2$, while maintaining acceptable F values
A continuum of effects

<table>
<thead>
<tr>
<th>No. of variants</th>
<th>β_{xy}</th>
<th>R^2</th>
<th>F</th>
<th>Power Estimates</th>
<th>β_{xy}</th>
</tr>
</thead>
<tbody>
<tr>
<td>IV(s)</td>
<td></td>
<td></td>
<td></td>
<td>0.1</td>
<td>0.3</td>
</tr>
<tr>
<td>5 variants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 IVs</td>
<td>0.189</td>
<td>0.055</td>
<td>11.6</td>
<td>0.10</td>
<td>0.61</td>
</tr>
<tr>
<td>Allele count</td>
<td>0.189</td>
<td>0.045</td>
<td>47.2</td>
<td>0.09</td>
<td>0.56</td>
</tr>
<tr>
<td>Weighted count</td>
<td>0.189</td>
<td>0.051</td>
<td>53.5</td>
<td>0.11</td>
<td>0.58</td>
</tr>
<tr>
<td>10 variants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 IVs</td>
<td>0.134</td>
<td>0.060</td>
<td>6.3</td>
<td>0.11</td>
<td>0.64</td>
</tr>
<tr>
<td>Allele count</td>
<td>0.134</td>
<td>0.045</td>
<td>46.5</td>
<td>0.09</td>
<td>0.53</td>
</tr>
<tr>
<td>Weighted count</td>
<td>0.134</td>
<td>0.051</td>
<td>53.7</td>
<td>0.11</td>
<td>0.57</td>
</tr>
<tr>
<td>20 variants</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20 IVs</td>
<td>0.094</td>
<td>0.069</td>
<td>3.6</td>
<td>0.14</td>
<td>0.71</td>
</tr>
<tr>
<td>Allele count</td>
<td>0.094</td>
<td>0.045</td>
<td>47.0</td>
<td>0.10</td>
<td>0.53</td>
</tr>
<tr>
<td>Weighted count</td>
<td>0.094</td>
<td>0.051</td>
<td>53.7</td>
<td>0.11</td>
<td>0.58</td>
</tr>
</tbody>
</table>

- Weighted count outperforms allele count (R^2 and F)
- Model misspecification decreases R^2 and F
## Major gene/polygene model

<table>
<thead>
<tr>
<th>No. of variants</th>
<th>IV(s)</th>
<th>$\beta_{gx}$</th>
<th>$R^2$</th>
<th>$F$</th>
<th>0.1</th>
<th>0.3</th>
<th>0.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 main effects + 8 polygenes (n=500)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 variants</td>
<td>10 IVs</td>
<td>0.081</td>
<td>0.118</td>
<td>6.6</td>
<td>0.13</td>
<td>0.67</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>Allele count</td>
<td>0.081</td>
<td>0.063</td>
<td>33.6</td>
<td>0.07</td>
<td>0.42</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td>Weighted count</td>
<td>0.081</td>
<td>0.102</td>
<td>57.0</td>
<td>0.10</td>
<td>0.60</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>2 major IVs +</td>
<td>0.081</td>
<td>0.105</td>
<td>19.7</td>
<td>0.11</td>
<td>0.62</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>allele count</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Major-gene/polygene has slightly higher $R^2$ than the weighted count, lower $F$ (but acceptable)

- Allows flexibility in model assumptions
Conclusions

• $R^2$, power are maximized when using one IV per variant. But...
• For fixed $R^2$, IV strength (F) decreases as # of IV increases

• Constructing optimal IV set is a balancing act:
  – Maximize $R^2$, power
  – Minimize bias (adequate F values)
• Weighted allele counts may accomplish this

• GWAS-based MR requires careful treatment of weak IV problem

• BUT... with multiple instruments, we have multiple exclusion restrictions