Evaluating Polygenic Prediction of Common Diseases
Lessons Learned From the Pre-GWAS days and ignored?

A. Cecile J.W. Janssens, PhD
Research professor of epidemiology
Department of Epidemiology
@cecilejanssens
Where we want to be

Identifying (high-)risk groups for targeted intervention using polygenic risk scores

Where we (still) are

...
Many current issues in polygenic risk research seem repetition of discussions in pre-GWAS days.
First mentions of genetic information, susceptibility for common diseases, not yet polygenic models
First mention of risk distributions
- Fitted risk distributions on cancer data from relatives of BC patients
- Describes (carefully) how polygenic risk can be applied in healthcare to make mammography screening more cost-effective
ACCE framework: evaluating genetic tests

- **Key**: Intended use: disorder & Setting: What is predicted in whom, for what purpose?
- Assessment of the test changes if intended use changes (different population or purpose)

Haddow & Palomaki, Human Genome Epidemiology, 2003
First study to show how multiple genes can be combined to predict risk, using regression analysis

Focused on posterior risk for carriers of one or more multiple risk alleles

(very strong per-allele effects by today’s standards (RR 1.5-3.5))
- Evaluation of test performance should include all people, also noncarriers of risk alleles
- Proposed using Area under the Receiver Operating Curve (AUC)
<table>
<thead>
<tr>
<th>Year</th>
<th>Researchers</th>
<th>Disease</th>
<th>Genetic variant</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td>Van der Net et al.</td>
<td>CHD in FH</td>
<td>6 established genes</td>
<td>0.55</td>
</tr>
<tr>
<td>2006</td>
<td>Weedon et al.</td>
<td>Type 2 diabetes</td>
<td>3 established genes</td>
<td>0.55</td>
</tr>
<tr>
<td>2007</td>
<td>Vaxillaire et al.</td>
<td>Type 2 diabetes</td>
<td>3 (out of 19)</td>
<td>0.56</td>
</tr>
<tr>
<td>2008</td>
<td>Lango et al.</td>
<td>Type 2 diabetes</td>
<td>18 established genes</td>
<td>0.60</td>
</tr>
<tr>
<td>2008</td>
<td>Van Hoek et al.</td>
<td>Type 2 diabetes</td>
<td>18 established genes</td>
<td>0.60</td>
</tr>
<tr>
<td>2007</td>
<td>Humphries et al.</td>
<td>CHD</td>
<td>4 (out of 12)</td>
<td>0.62</td>
</tr>
<tr>
<td>2008</td>
<td>Harley et al.</td>
<td>SLE</td>
<td>6 from GWAS</td>
<td>0.67</td>
</tr>
<tr>
<td>2006</td>
<td>Podgoreanu et al.</td>
<td>MI after surgery</td>
<td>3 (out of 48)</td>
<td>0.70</td>
</tr>
<tr>
<td>2006</td>
<td>Maller et al.</td>
<td>AMD</td>
<td>5 established regions</td>
<td>0.80</td>
</tr>
<tr>
<td>2008</td>
<td>Wang et al.</td>
<td>Hypertriglyceridemia</td>
<td>7 established genes</td>
<td>0.80</td>
</tr>
</tbody>
</table>
How to get high AUC: common variants with strong effects

<table>
<thead>
<tr>
<th>Type 2 diabetes</th>
<th></th>
<th>Hypertriglyceridemia</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AUC = 0.60</strong></td>
<td><strong>AUC = 0.80</strong></td>
<td></td>
</tr>
<tr>
<td>TCF7L2</td>
<td>1.36</td>
<td>APOA5 19WW</td>
</tr>
<tr>
<td>KCNJ11</td>
<td>1.25</td>
<td>APOA5 -1131CC</td>
</tr>
<tr>
<td>CDKN2A/2B</td>
<td>1.21</td>
<td>APOE non-e3</td>
</tr>
<tr>
<td>PPARG</td>
<td>1.21</td>
<td>GCKR TT</td>
</tr>
<tr>
<td>ADAM30</td>
<td>1.15</td>
<td>TRIB1 AA</td>
</tr>
<tr>
<td>CDNK2A/2B</td>
<td>1.13</td>
<td>TBL2 CC</td>
</tr>
<tr>
<td>IGF2BP2</td>
<td>1.12</td>
<td>GALNT2 GG</td>
</tr>
<tr>
<td>FTO</td>
<td>1.11</td>
<td></td>
</tr>
<tr>
<td>CDKAL1</td>
<td>1.11</td>
<td></td>
</tr>
</tbody>
</table>

Type 2 diabetes

AUC = 0.60

Don’t treat  Treat

Lango et al Diabetes 2008

AMD

AUC = 0.76

Don’t treat  Treat

Seddon et al IOVS 2009

AUC = degree of separation between risk distributions of affected and unaffected individuals—nothing more, nothing less
When can a risk factor be used as a worthwhile screening test?

N J Wald, A K Hackshaw, C D Frost

Summary points

To be a worthwhile screening test, a risk factor must be strongly associated with a disorder.

Fig 4 Distribution of maternal serum α fetoprotein in pregnancies affected and unaffected by open spina bifida (derived from Wald et al) and distribution of serum cholesterol in men who did and did not die of ischaemic heart disease (derived from Wald et al).
<table>
<thead>
<tr>
<th>Year</th>
<th>Researchers</th>
<th>Disease</th>
<th>Genetic variant</th>
<th>AUC</th>
<th>Δ AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006</td>
<td>Podgoreanu et al.</td>
<td>MI after surgery</td>
<td>3 (out of 48)</td>
<td>0.70</td>
<td>+0.06</td>
</tr>
<tr>
<td>2007</td>
<td>Humphries et al.</td>
<td>CHD</td>
<td>4 (out of 12)</td>
<td>0.66</td>
<td>+0.04</td>
</tr>
<tr>
<td>2007</td>
<td>Morisson et al.</td>
<td>CHD</td>
<td>11 (out of 116)</td>
<td>0.76</td>
<td>+0.01</td>
</tr>
<tr>
<td>2008</td>
<td>Kathiresan et al.</td>
<td>CVD</td>
<td>9 (out of 11)</td>
<td>0.80</td>
<td>+0.00</td>
</tr>
<tr>
<td>2008</td>
<td>Zheng et al</td>
<td>Prostate cancer</td>
<td>5 (out of 16)</td>
<td>0.61</td>
<td>+0.02</td>
</tr>
<tr>
<td>2005</td>
<td>Lyssenko et al.</td>
<td>Type 2 diabetes</td>
<td>3 establ. variants</td>
<td>0.68</td>
<td>+0.00</td>
</tr>
<tr>
<td>2008</td>
<td>Vaxillaire et al.</td>
<td>Type 2 diabetes</td>
<td>3 (out of 19)</td>
<td>0.82</td>
<td>+0.00</td>
</tr>
<tr>
<td>2008</td>
<td>Lyssenko et al</td>
<td>Type 2 diabetes</td>
<td>11 establ. variants</td>
<td>0.74</td>
<td>+0.01</td>
</tr>
<tr>
<td>2008</td>
<td>Van Hoek et al</td>
<td>Type 2 diabetes</td>
<td>18 establ. variants</td>
<td>0.66</td>
<td>+0.02</td>
</tr>
<tr>
<td>2008</td>
<td>Lango et al.</td>
<td>Type 2 diabetes</td>
<td>18 establ. variants</td>
<td>0.78</td>
<td>+0.02</td>
</tr>
<tr>
<td>2008</td>
<td>Meigs et al.</td>
<td>Type 2 diabetes</td>
<td>18 establ. variants</td>
<td>0.90</td>
<td>+0.00</td>
</tr>
</tbody>
</table>

Janssens & van Duijn *Hum Mol Genet* 2008
Discriminative accuracy

Whites

Traditional risk factors 0.764
+ 11 genetic markers 0.766

Difference: 0.002 (0.000-0.006)

Blacks

Traditional risk factors 0.758
+ 11 genetic markers 0.769

Difference: 0.011 (0.002-0.024)

FIGURE 2. Receiver operating characteristic curves using the Atherosclerosis Risk in Communities study cardiovascular risk score (ACRS) alone and incorporating the genetic risk score (GRS) for Whites (A) and Blacks (B), United States, 1986–2001.

Association Between a Literature-Based Genetic Risk Score and Cardiovascular Events in Women

Context: While multiple genetic markers associated with cardiovascular disease have been identified by genome-wide association studies, their aggregate effect beyond traditional factors is uncertain, particularly among women.

Objective: To test the predictive ability of a literature-based genetic risk score in women.

Design, Setting, and Participants: Prospective cohort of 19,313 healthy white women in the Women’s Genome Health Study followed up over 12.3 years (interquartile range, 11.6–12.8 years). Genetic risk scores were constructed from the National Human Genome Research Institute’s catalog of known cardiovascular genes.

Utility of genetic and non-genetic risk factors in prediction of type 2 diabetes: Whitehall II prospective cohort study

Philippa J Talmud, professor of cardiovascular genetics; Aroon D Hingorani, professor of genetic epidemiology; Jackie A Cooper, biostatistician; Michael G Marmot, MRC professor of epidemiology and public health; Eric J Brunner, reader in epidemiology and public health; Meena Kumari, senior research fellow in population health; Mika Kivimäki, professor of social epidemiology; Steve E Humphries, BHF professor of cardiovascular genetics.

COGENT (COlorectal cancer GENeTics): an international consortium to study the role of polymorphic variation on the risk of colorectal cancer

IFM Tomlinson1, M Dunlop2, H Campbell3, B Zanke1-3, S Gallinger1, T Hudson4, T Koestler5, C van Leeuwen3, S Nitzkyowsk6, S Tuzun7, K Arrighi1, K Herrmann8, A Lindstrom1, A Forsti1, D Sekeres1, T van Wezel1, H Morreale1, JT Wijnhoven1, P van den Bemt1, P Matsuda1, Y Nakamura1, S Castelvecchio1, C Ruiz-Ponte1, A Casteldale1, A Carracedo1, JW Ho1, P Shum1, RWH Hofstra1, P Vodicka1, J Hampe1, C Schiffmayer1, T Tjonneland1, H Vliegher1, MM Lech1, C Schnickel1, V Moreno1, CM Villanueva1, P Peterlongo1, P Radice1, MH Echeverry1, A Velez1, L Carvajal1, R Scott1, S Penegar1, P Broderick1, A Temper1 and RS Houston1.
It was early days

- Researchers unfamiliar with methods—still learning
- ‘Few’ genes identified, no impact should have been expected yet
- Little consideration of intended use
- Non-representative populations
- No relevant comparison with clinical risk models
- Relying on p-value more than actual improvement AUC
- Reported limited predictive value, while leaving the door wide open for GWAS to deliver
Multi-gene test predicts early heart disease risk

Date: January 8, 2018
Source: American Heart Association

February 21, 2018

Polygenic Risk Scores Show Utility for Stratifying Disease Risk
Polygenic Risk Scores Offer a Glimpse of True Clinical Utility
Brandon May

Genetic score can identify Alzheimer's risk much earlier
By Catharine Paddock PhD | Published Wednesday 7 March 2018
Fact checked by Jasmin Collier

Precision education: DNA test for intelligence could be on the horizon
Antonio Regalado | MIT Technology Review | April 10, 2018

Scientists predict reading ability from DNA analysis alone
Evaluation of polygenic risk scores for ovarian cancer risk prediction in a prospective cohort study

New Results

PRS of 96 snps

AUC 0.58

AUC 0.60

<table>
<thead>
<tr>
<th>Polygenic score</th>
<th>Derivation strategy</th>
<th>N Variants</th>
<th>Area under the curve</th>
<th>Odds ratio (per SD Increment)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tada et al. (7)</td>
<td>Variants that had achieved genome-wide levels of statistical significance in prior GWAS ($p&lt;5 \times 10^{-8}$)</td>
<td>50</td>
<td>0.59</td>
<td>1.38</td>
</tr>
<tr>
<td>Abraham et al. (8)</td>
<td>Linkage-disequilibrium based thinning of variants from prior GWAS</td>
<td>49,310</td>
<td>0.59</td>
<td>1.38</td>
</tr>
<tr>
<td>FSnet</td>
<td>Pruning based on statistical significance ($p &lt; 0.05$) and linkage disequilibrium ($r^2 &lt; 0.4$) of variants from prior GWAS</td>
<td>116,859</td>
<td>0.62</td>
<td>1.34</td>
</tr>
<tr>
<td>FSow</td>
<td>LDPred computational algorithm to assign weights to all available variants from prior GWAS via explicit modeling of linkage disequilibrium</td>
<td>6,630,150</td>
<td>0.64</td>
<td>1.67</td>
</tr>
</tbody>
</table>
When geneticists use the word *prediction*, they give it a different meaning than the rest of us do. We usually think of predictions as accurate forecasts for particular situations. At a carnival, you might encounter a man who promises to predict your weight simply by looking you over. If you weighed, say, 130 pounds, and he guessed 132, you might be impressed. If he guessed 232, you’d expect to walk away with a giant teddy bear.

Geneticists are a lot more forgiving about predictions. When they try to predict a trait from a set of genes, their prediction may be dead on, or it may be no better than random. Or, as is almost always the case, it is somewhere in between.
New analyses in polygenic risk research

• Statistical significance of polygenic risk score
  → If significant, risk score predicts

• Odds ratio highest risk group relative to lowest or rest
  – E.g., highest versus lowest quintile or highest 2.5% versus rest
  – Compare to OR of mutation carriers: risk is similar then polygenic risk score is predictive (FYI: there easily is a tip of the tail that has this OR)
Shortcomings of studies same as decade ago

Still often:

- No consideration of intended use, so no clue whether predictive performance is sufficient

- Risk thresholds, if chosen, based on little to nothing
- No calibration
- No validation
- No appropriate comparison with relevant clinical models
Where we want to be

Identifying (high-)risk groups for targeted intervention using polygenic risk scores

Where we (still) are

Doing (too much) research without knowing whether results will ever be relevant
→ Not contributing to evidence base needed to responsibly introduce genomics in healthcare
WILL GENETICS REVOLUTIONIZE MEDICINE?

In our rush to fit medicine with the genetic mantle, we are losing sight of other possibilities for improving the public health. Differences in social structure, lifestyle, and environment account for much larger proportions of disease\textsuperscript{42,43} than genetic differences. Although we do not contend that the genetic mantle is as imperceptible as the emperor’s new clothes were, it is not made of the silks and ermines that some claim it to be. Those who make medical and science policies in the next decade would do well to see beyond the hype.

\begin{flushright}
NEIL A. HOLTZMAN, M.D., M.P.H. \\
THERESA M. MARTEAU, PH.D.
\end{flushright}