Defining high risk of CKD: Genetic Risk Scores as New Players

Krzysztof Kiryluk MD
Associate Professor of Medicine
Columbia University
DISCLOSURES

- Research collaborations: AstraZeneca, Visterra, Vanda, Aevi Genomics
- Consulting or advisory boards: HiBio, Vera, Travere
Genetic Architecture of Human Disease

Mendelian Traits

- Rare variants
- Large effects

Complex Traits

- Common variants
- Small effects
Thousands of GWAS Loci for Complex Traits

Genetic Risk Score (GRS)

\[ GRS = \beta_1 \times \text{snp}_1 + \beta_2 \times \text{snp}_2 + \ldots + \beta_n \times \text{snp}_n \]

Standardized GRS = \( \frac{GRS - \text{mean}_{ctr}}{sd_{ctr}} \)

OR = 2.0
Genome-wide Polygenic Score (GPS)

Methods to account for linkage disequilibrium (LD):
- P+T – P-value thresholding and LD pruning to select independent SNPs
- LD Pred – adjusts SNP weights to account for non-independence

Extended PRS formulation captures the effects of all ~10M common variants across the genome

\[ GPS_i = \sum_{j} \hat{\beta}_j \times \text{dosage}_{ij} \]
Polygenic Risk Models are Phenotype-specific

GWAS for eGFR and CKD
N ~ 1 million

308 genome-wide significant loci
~7% of variance in eGFR

Wuttke et al. *Nature Genetics* 2019

GWAS for Membranous Nephropathy
3K cases/9K controls

4 genome-wide significant loci
~30% of disease risk

Xie et al. *Nature Commun.* 2020
Major Limitation: Ancestry Bias

80% of GWAS participants are of European ancestry
(Europeans represent only 16% of the global population)

GPS accuracy by ancestry relative to Europeans
(17 quantitative traits from UKBB)

Reasons for poor cross-ancestry transferability:
• European over-representation in GWAS, bias in array design, bias in imputation
• LD differences between populations
• Differences in the environment (via GxE interactions)
• Actual differences in the genetic architecture (e.g., APOL1)

Electronic Medical Records & Genomics Phase IV (eMERGE-IV)

- Optimization of polygenic scores for 10 common diseases
- Recruitment of 25,000 participants of diverse ancestries for PRS + Monogenic screening
- Return of results with prospective collection of outcomes

- Coronary Artery Disease
- Chronic Kidney Disease
- Type 1 Diabetes
- Type 2 Diabetes
- Atrial Fibrillation
- Obesity
- Asthma
- Breast Cancer
- Colorectal Cancer
- Prostate Cancer
eMERGE-IV PRS Testing Process (CLIA Lab)

Major challenge: to implement standardized GPS calculations in CLIA-certified testing labs, establish actionable thresholds and standardize reporting.

DNA is genotyped with Global Diversity Array (1.8M SNPs).

Genotyping data is phased and imputed (1000G reference)

Raw PRS scores are calculated for each condition

Ancestry calibration based on AoU

‘high risk’ defined for adjusted scores crossing the threshold

Clinical report is generated, signed and delivered to the study portal

Genomic Integrative Risk Assessment (GIRA)

**Genome-wide Polygenic Score (GPS) for CKD**

### GPS Optimization

**Summary Statistics**
CKDGen GWAS for eGFR (Wuttke et al)

**1000G Reference Panel**
All Populations (N=2,504)

**Derive & select the best GPS**
UK Biobank European ancestry (70%)

- 6,573 cases 170,635 controls
- Best GPS: P+T ($r^2=0.2$, $P=0.03$)

**APOL1 effects derivation**
UK Biobank African ancestry

- 967 cases 6,191 controls

**GPS Validation**

<table>
<thead>
<tr>
<th>Study</th>
<th>UK Biobank</th>
<th>eMERGE-III</th>
<th>BioMe</th>
<th>UAB (African-American ONLY)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>European</td>
<td>European</td>
<td>European</td>
<td>HyperGEN</td>
</tr>
<tr>
<td></td>
<td>2,759 cases</td>
<td>10,572 cases</td>
<td>870 cases</td>
<td>109 cases, 619 controls</td>
</tr>
<tr>
<td></td>
<td>72,968 controls</td>
<td>8,030 controls</td>
<td>1,851 controls</td>
<td></td>
</tr>
<tr>
<td></td>
<td>East Asian</td>
<td>African American</td>
<td>African American</td>
<td>729 cases, 1,149 controls</td>
</tr>
<tr>
<td></td>
<td>26 cases, 1,525 controls</td>
<td>1,143 cases, 1,600 controls</td>
<td>729 cases, 1,149 controls</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SW Asian</td>
<td>East Asian</td>
<td>East Asian</td>
<td>REGARDS</td>
</tr>
<tr>
<td></td>
<td>209 cases, 6,258 controls</td>
<td>96 cases, 97 controls</td>
<td>61 cases, 353 controls</td>
<td>1055 cases, 4314 controls</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Latinx</td>
<td>Latinx</td>
<td>GenHAT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>382 cases, 533 controls</td>
<td>1,004 cases, 1,706 controls</td>
<td>924 cases, 2,454 controls</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>WARFARIN</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>308 cases, 140 controls</td>
</tr>
</tbody>
</table>
Additive effects of *APOL1* and polygenic background

Six validation cohorts of African ancestry (4,268 cases and 10,276 controls)

Outcome of CKD stage 3+, covariates: age, sex, diabetes, cohort, and ancestry PCs

Khan et al. *Nature Medicine* 2022
GPS for CKD: tail cut-off selection

OR=3.0, equivalent to + family history of kidney disease

European Ancestry
14,201 cases, 82,849 controls
3 cohorts

African Ancestry
4,268 cases, 10,276 controls
6 cohorts

Asian Ancestry
392 cases, 8,233 controls
4 cohorts

Latinx Ancestry
1,386 cases, 2,239 controls
2 cohorts

All Cohorts
20,247 cases, 103,597 controls
15 cohorts

Khan et al. Nature Medicine 2022
Monogenic risk of CKD

Diagnostic ES
N=3,315
CKD patients

307 with Monogenic Disease
(9.3% diagnostic yield by ACMG criteria)

Groopman at al. NEJM 2019
Interplay of polygenic and monogenic risk for CKD

- E-phenotyping for CKD
- Imputation using 1000G and quality control of imputed data
- Calculation of GPS for CKD
- Ancestry correction and GPS standardization using 1000G
- Testing GPS as a predictor of CKD in M1, M2, M3 QV carriers

- SNP array N=488,377
- Exome Sequence N=469,835
- Meta-analysis

- UK Biobank
- All of Us (release 1)
- SNP array N=165,208
- Genome Sequence N=98,622

Khan et al. *Nature Comm (in press)*
Polygenic risk and ADPKD

UKBB + All-of-Us
~600K individuals with ES/GS

pLOF and ClinVar ‘Pathogenic’ variants
PKD1 and PKD2 genes

Exclude:
1. MAF>10-5 in any ancestry (UKBB, AoU, gnomAD)
2. Single submitter ‘Pathogenic’ in ClinVar
3. Conflict of pathogenicity in ClinVar

Monogenic ADPKD variants used in the analysis (Model 1)

N=206 carriers (0.03%)

Khan et al. Nature Comm (in press)
## Outcome of CKD stage 3+, covariates: age, sex, diabetes, cohort, and ancestry PCs

<table>
<thead>
<tr>
<th>PKD variants</th>
<th>GPS tertile</th>
<th>OR, 95% CI, P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carrier</td>
<td>High</td>
<td>54.4 (26.1-113.0), P=9.64E-27</td>
</tr>
<tr>
<td>Carrier</td>
<td>Intermediate</td>
<td>35.8 (16.7-76.4), P=1.96E-20</td>
</tr>
<tr>
<td>Carrier</td>
<td>Low</td>
<td>3.03 (1.03-8.95), P=4.37E-02</td>
</tr>
<tr>
<td>Noncarrier</td>
<td>High</td>
<td>1.82 (1.75-1.89), P=3.44E-208</td>
</tr>
<tr>
<td>Noncarrier</td>
<td>Intermediate</td>
<td>Reference</td>
</tr>
<tr>
<td>Noncarrier</td>
<td>Low</td>
<td>0.62 (0.59-0.65), P=3.95E-96</td>
</tr>
</tbody>
</table>

**Khan et al. Nature Comm (in press)**
### Polygenic risk and COL4A-AN

<table>
<thead>
<tr>
<th>COL4A-AN variants</th>
<th>GPS tertile</th>
<th>OR, 95% CI, P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carrier</td>
<td>High</td>
<td>2.53 (1.66-3.85), P=1.44E-05</td>
</tr>
<tr>
<td>Carrier</td>
<td>Intermediate</td>
<td>1.66 (1.03-2.68), P=3.70E-02</td>
</tr>
<tr>
<td>Carrier</td>
<td>Low</td>
<td>1.08 (0.63-1.86), P=7.78E-01</td>
</tr>
<tr>
<td>Noncarrier</td>
<td>High</td>
<td>1.82 (1.75-1.89), P=3.44E-208</td>
</tr>
<tr>
<td>Noncarrier</td>
<td>Intermediate</td>
<td>Reference</td>
</tr>
<tr>
<td>Noncarrier</td>
<td>Low</td>
<td>0.62 (0.59-0.65), P=1.13E-97</td>
</tr>
</tbody>
</table>

Outcome of CKD stage 3+, covariates: age, sex, diabetes, cohort, and ancestry PCs

Khan et al. *Nature Comm (in press)*
Summary:

- GPS offers a promising tool for kidney disease risk stratification: top 2% associated with >3-fold higher risk (a family history risk equivalent).

- Monogenic risk, APOL1, and polygenic risk appear to have additive effects; Added value of polygenic risk over family history and other known risk factors still unknown for kidney disease.

- Ongoing work includes improvements in the overall predictive performance and cross-ancestry portability (new GWAS, new methods) and prospective testing of clinical utility (prospective eMERGE-IV).
Acknowledgements

Kiryluk Lab

Lili Liu  Atlas Khan  Katherine Xu
Junying Zhang  Francesca Zanoni  Chen Wang
Cecilia Berrouet  German Varela  Ning ‘Sunny’ Shang

Columbia University

Nephrology and Center for Precision Medicine and Genomics: Ali Gharavi, Simone Sanna-Cherchi, Maya Sabatello

Statistical Genetics: Iuliana Ionita-Laza

Glomerular Center: Jerry Appel, Jai Radhakrishnan, Andy Bomback, Pietro Canetta

Transplant Center: David Cohen, Sumit Mohan

Renal Pathology: Vivette D’Agati, Ibrahim Batal

Biomedical Informatics: Chunhua Weng, George Hripcsak

Clinical Genetics: Wendy Chung, Vaidehi Jobanputra
Questions?