Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations

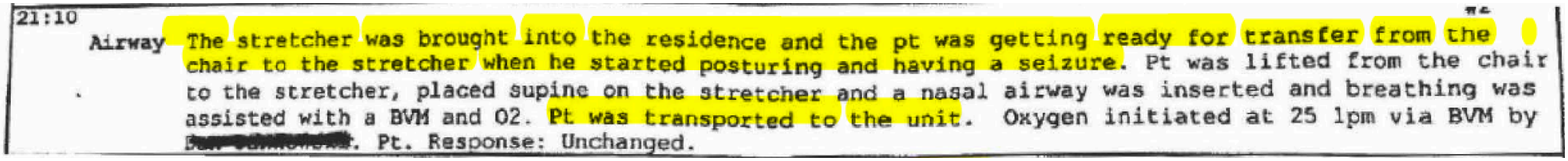
Sekar Kathiresan, M.D.
Director, Center for Genomic Medicine, MGH
Director, Cardiovascular Disease Initiative, Broad

May 9, 2018

Massachusetts General Hospital

Harvard Medical School

Health care scenario: 42yo male with dizziness, profuse sweating



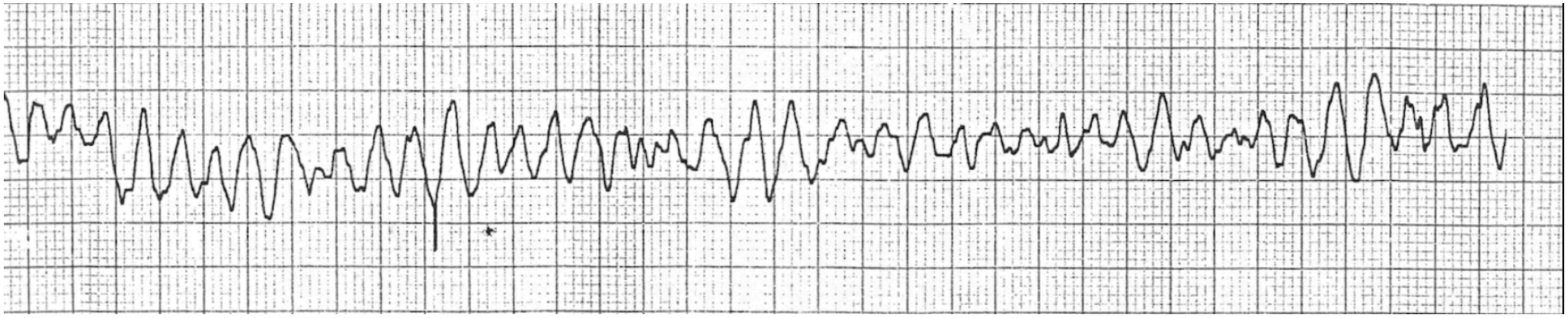
21:10

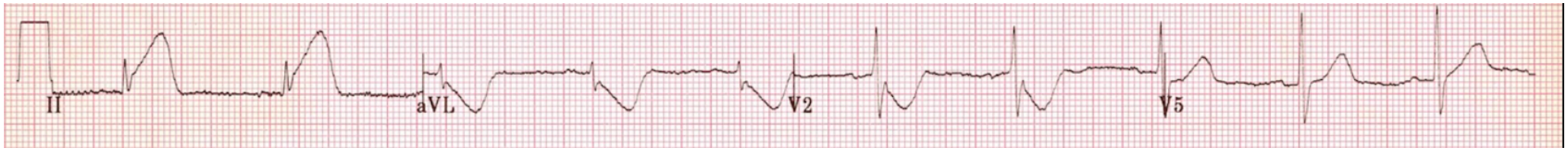
Airway The stretcher was brought into the residence and the pt was getting ready for transfer from the chair to the stretcher when he started posturing and having a seizure. Pt was lifted from the chair to the stretcher, placed supine on the stretcher and a nasal airway was inserted and breathing was assisted with a BVM and O2. Pt was transported to the unit. Oxygen initiated at 25 lpm via BVM by [redacted]. Pt. Response: Unchanged.

42yo male with cardiac arrest due to acute myocardial infarction (MI)

21:10 **Airway** The stretcher was brought into the residence and the pt was getting ready for transfer from the chair to the stretcher when he started posturing and having a seizure. Pt was lifted from the chair to the stretcher, placed supine on the stretcher and a nasal airway was inserted and breathing was assisted with a BVM and O2. Pt was transported to the unit. Oxygen initiated at 25 lpm via BVM by [redacted]. Pt. Response: Unchanged.

21:15 000 **V-Fib** 3041A #2
Initiate IV Once inside the unit the pt went into cardiac arrest, ALS back-up was called, CPR was started, V-Fib was noted on the monitor, precardial thump, CPR continued. Pupils noted to be dialated and fixed. Peripheral IV initiated by [redacted] with 18ga. at LF. Attempts: 1, successful. Authorization: Via Protocol. Pt. Response: Unchanged. 1000cc's of NSS wide open.

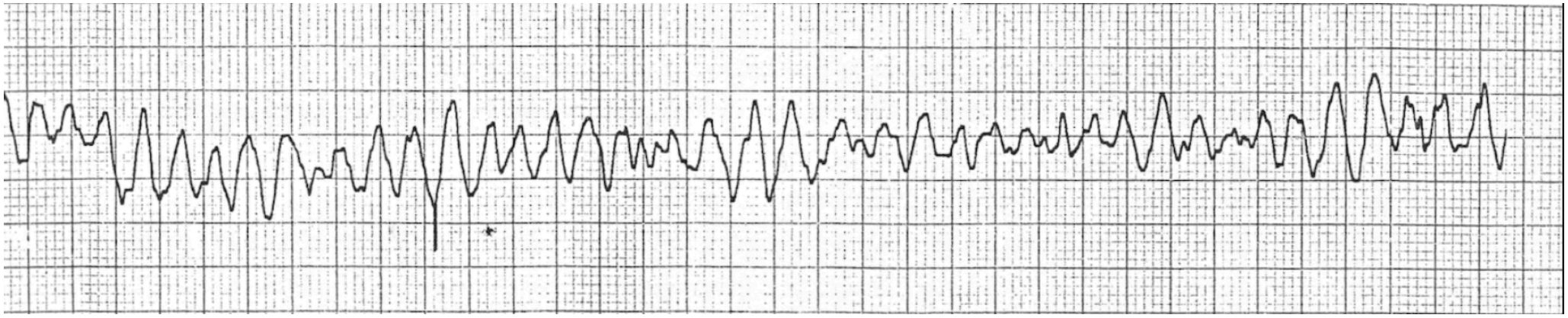


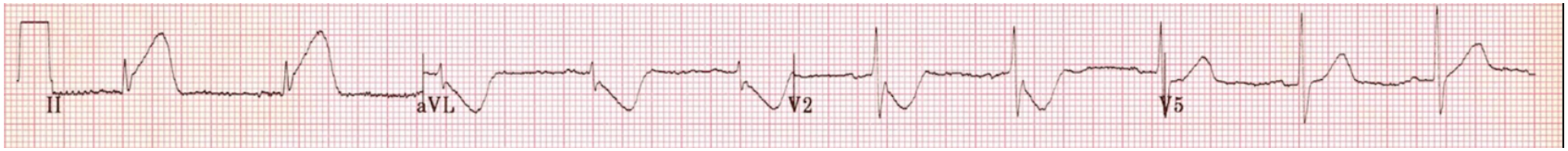


Anoxic brain injury
Expired after 10 days in hospital

21:10 Airway The stretcher was brought into the residence and the pt was getting ready for transfer from the chair to the stretcher when he started posturing and having a seizure. Pt was lifted from the chair to the stretcher, placed supine on the stretcher and a nasal airway was inserted and breathing was assisted with a BVM and O2. Pt was transported to the unit. Oxygen initiated at 25 lpm via BVM by [redacted]. Pt. Response: Unchanged.

21:15 000 V-Fib 3041A #2
Initiate IV Once inside the unit the pt went into cardiac arrest, ALS back-up was called, CPR was started, V-Fib was noted on the monitor, precardial thump, CPR continued. Pupils noted to be dialated and fixed. Peripheral IV initiated by [redacted] with 18ga. at LF. Attempts: 1, successful. Authorization: Via Protocol. Pt. Response: Unchanged. 1000cc's of NSS wide open.





42yo male with fatal, early-onset MI

MI risk factors prior to event

Total cholesterol 198 mg/dl

LDL cholesterol 124 mg/dl

HDL cholesterol 40 mg/dl

Triglycerides 170 mg/dl

Blood pressure 122/78

Body mass index 26

Non-smoker

No type 2 diabetes

Family history: father with MI at 54

ACC/AHA10y ASCVD risk calculator typically used for statin allocation decision: 1.7% ('low-risk')

Pooled Cohort Risk Assessment Equations

Predicts 10-year risk for a first atherosclerotic cardiovascular disease (ASCVD) event

ClinCalc.com » Cardiology » Pooled Cohort 10-Year ASCVD Risk Assessment Equations

Risk Factors for ASCVD

Gender	Male / Female	Systolic BP	122 mmHg
Age	42 years	Receiving treatment for high blood pressure (if SBP > 120 mmHg)	No / Yes
Race	White or other	Diabetes	No / Yes
Total Cholesterol	198 mg/dL	Smoker	No / Yes
HDL Cholesterol	40 mg/dL		

Reset | Calculate

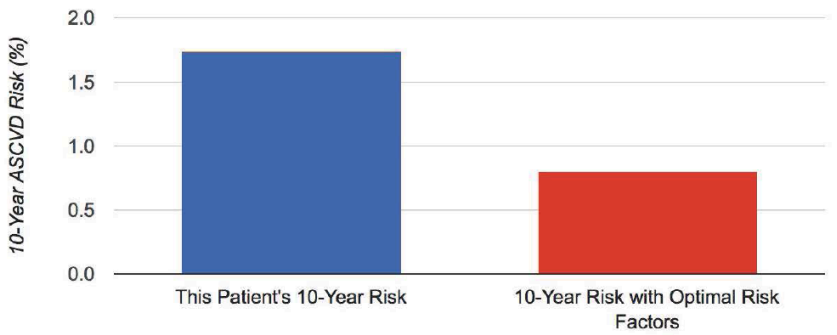
US units

ASCVD Risk Evaluation

10-year risk of atherosclerotic cardiovascular disease: 1.7%

10-year risk in a similar patient with optimal risk factors: 0.8%

Save Image

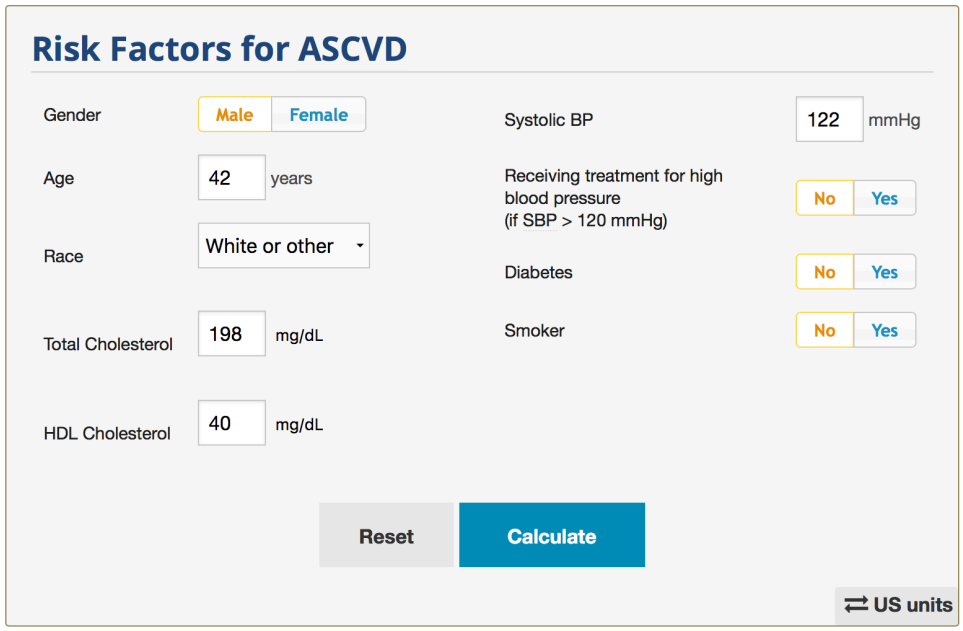


Why is the ACC/AHA pooled cohort equation not useful in young people?

Pooled Cohort Risk Assessment Equations

Predicts 10-year risk for a first atherosclerotic cardiovascular disease (ASCVD) event

ClinCalc.com » Cardiology » Pooled Cohort 10-Year ASCVD Risk Assessment Equations



Model almost entirely driven by 'age'

In population, older you are, more likely you are to have a heart attack!

Health care scenario

What is predicted?	Risk for heart attack
Intended target population	Men/women < 55yo
How	Gene variant(s)
For what purpose	Statin initiation at early age

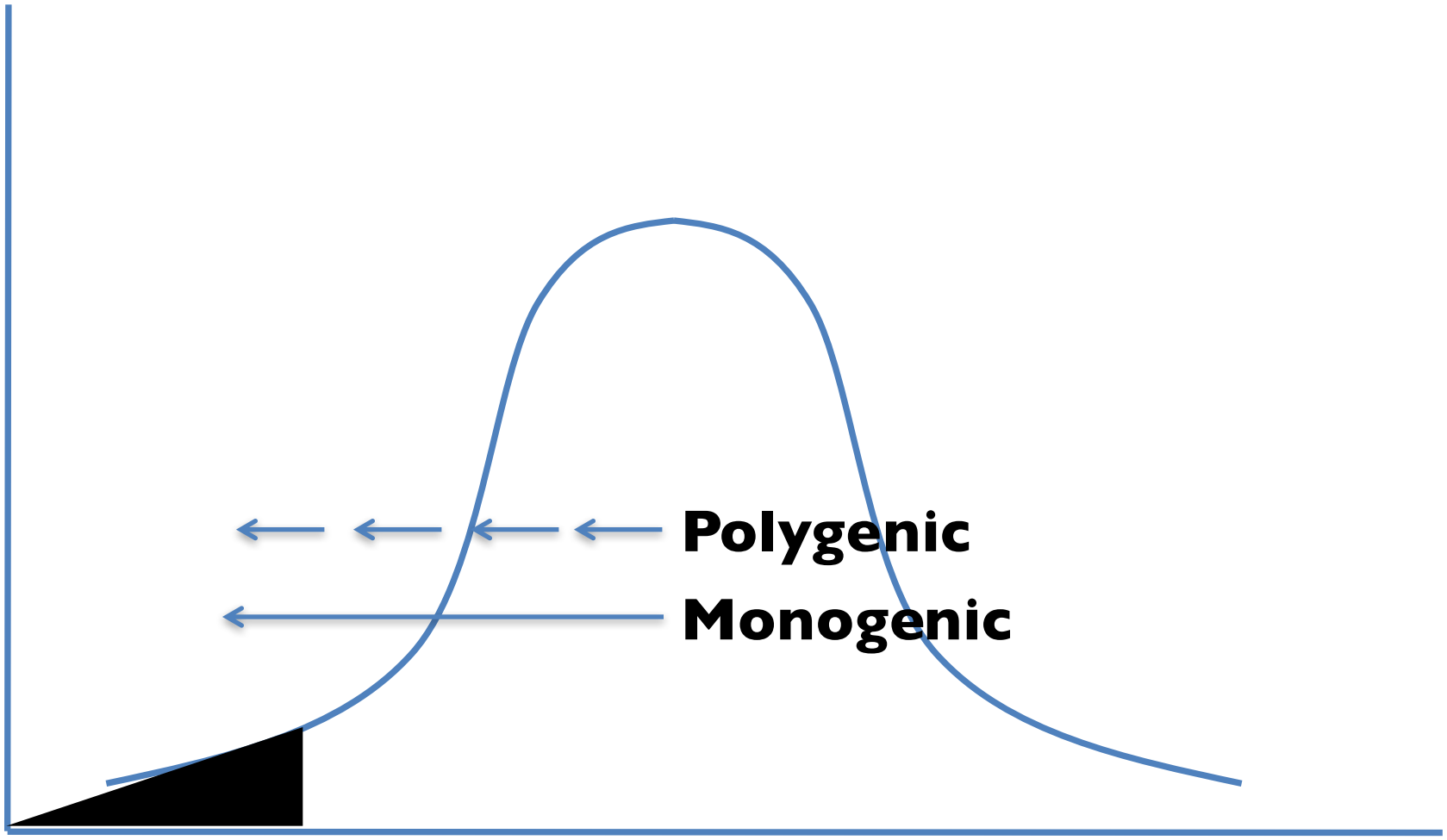
Janssens, Martens, Prediction Research Manual
http://www.cecilejanssens.org/wp-content/uploads/2018/01/PredictionManual2.0.pdf

For early-onset disease, stratifying individuals based on inborn DNA variation an option

Most diseases inherited component

Stratify individuals based on inherited DNA variation

Inherited component to early heart attack

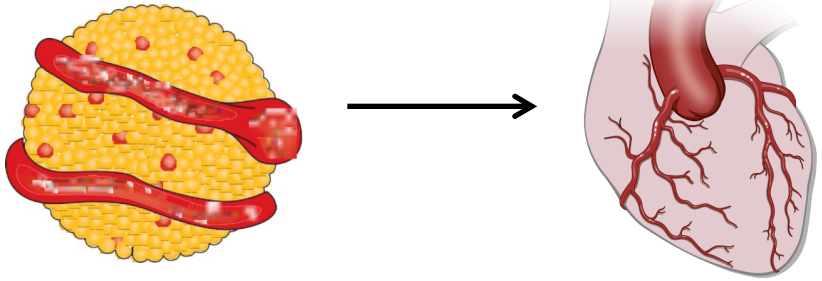


MI at age < 55 **Age onset at MI**

Traditional approach:
Genetic prediction focuses on rare, monogenic mutations

Traditional approach:
Genetic prediction focuses on rare, monogenic mutations

Familial hypercholesterolemia



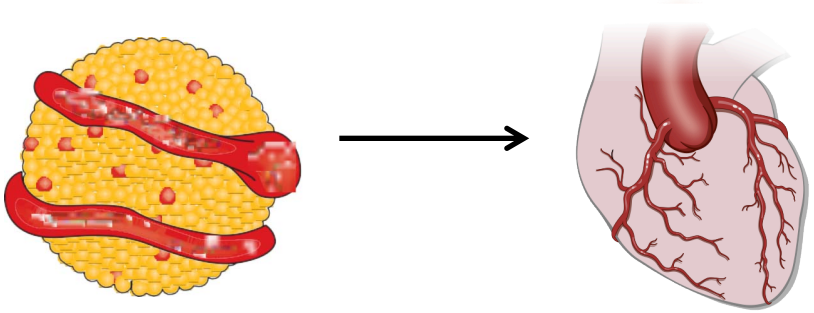
↑ **Cholesterol**

Heart attack
3x increased risk

0.4% of the population

Traditional approach:
Genetic prediction focuses on rare, monogenic mutations

Familial hypercholesterolemia



↑ **Cholesterol**

Heart attack
3x increased risk

Identify this risk group early in life
Target statin intervention

Testing for familial hypercholesterolemia mutations: CDC Tier 1 Genomics Application

Tier 1

- FDA label requires use of test to inform choice or dose of a drug
- FDA cleared or approved companion diagnostic device
- CMS covers testing
- Clinical practice guidelines based on systematic review supports testing

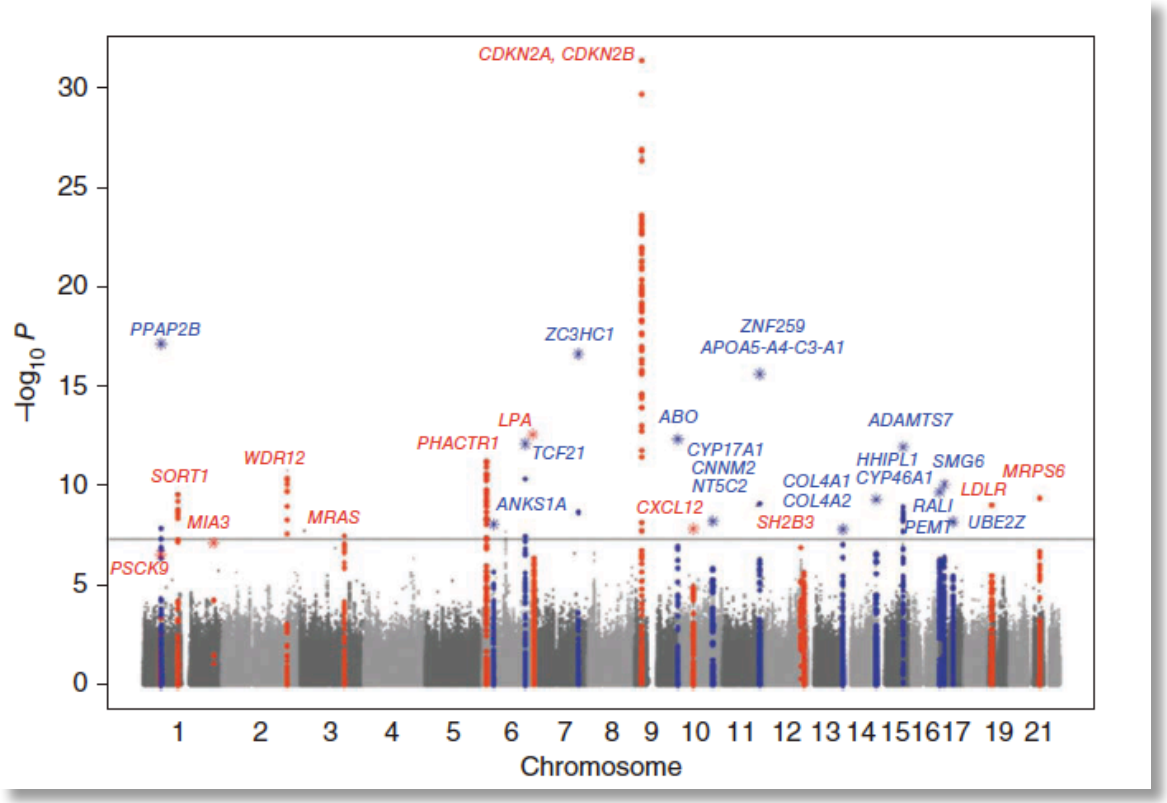
Genomics Application General Information	
Tier Classification	Tier 1
Disease/Disorder:	Familial hypercholesterolemia (FH)
Test or Application:	DNA testing and LDL-C concentration measurement
Target Population:	
Intended Use:	Cascade testing of relatives of people diagnosed with FH
Application Type	Other
Basis:	Clinical Practice Guideline Systematic Review
Entered Date:	08/19/2015
Last Updated Date:	08/19/2015

Question: Can we identify additional patients with a polygenic risk model?

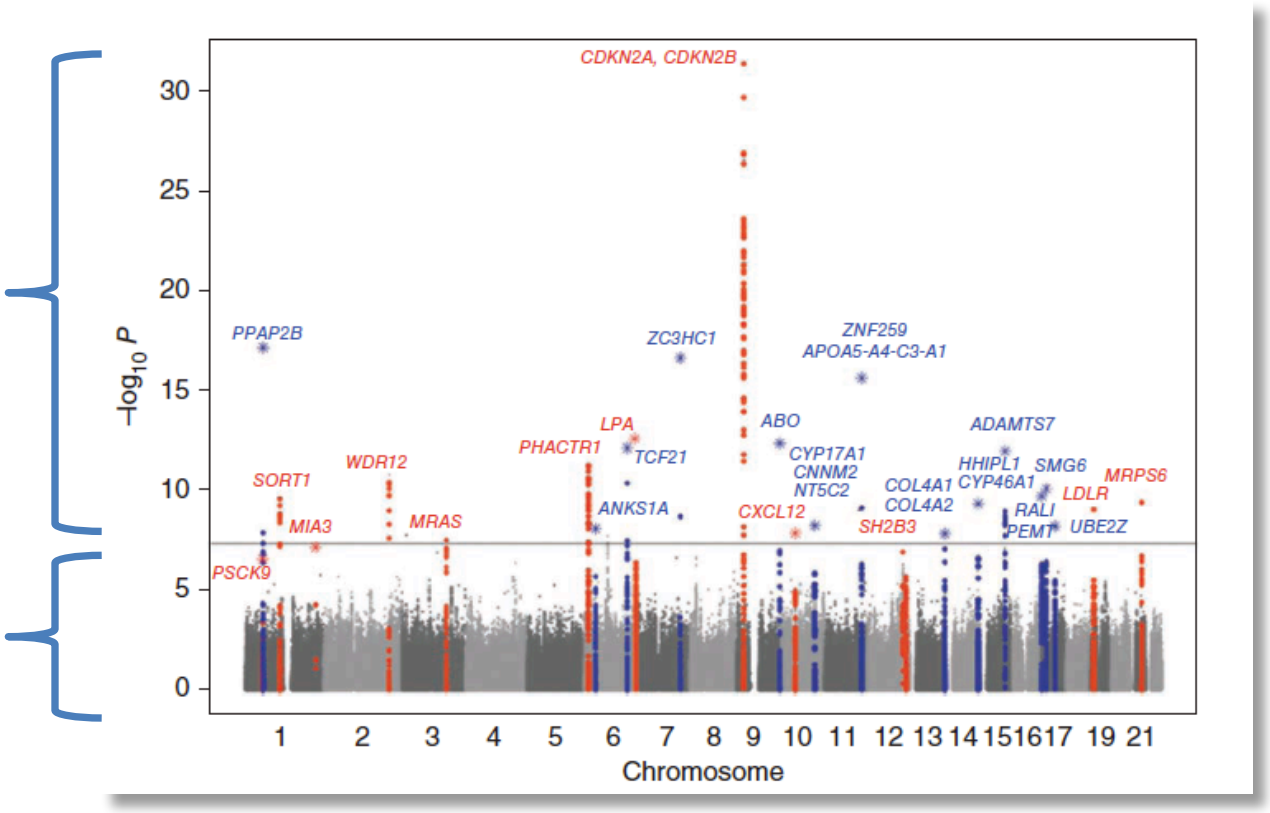
Concept: polygenic risk scores

'N' polymorphisms → 0, 1, or 2 copies of the risk allele → Score ranging from 0 to 2N for each person

Kathiresan, *N Engl J Med* (2008)
Ripatti, *Lancet* (2010)
Khera, *N Engl J Med* (2016)



Polygenic risk scores: move from top SNPs to a genome-wide set of 6.6M for prediction

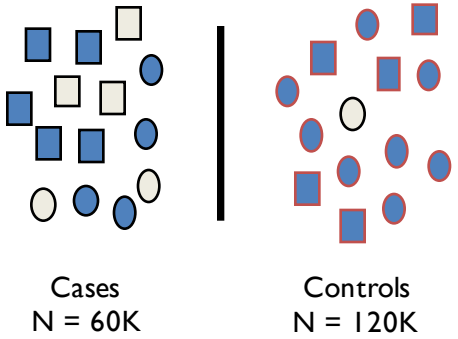


Khera*, Chaffin*, bioRxiv 2017

Amit V. Khera

Hypothesis: a polygenic score including a genome-wide set of SNPs can identify individuals with risk equivalent to a monogenic mutation

Step 1
Training data set: effect sizes for 6.6 million variants from genome-wide association study

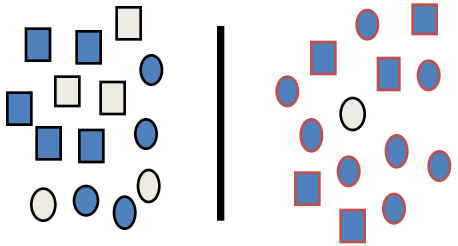


Genotypes: from arrays + imputation

Khera*, Chaffin*, *bioRxiv* (2017)

Hypothesis: a polygenic score including a genome-wide set of SNPs can identify individuals with risk equivalent to a monogenic mutation

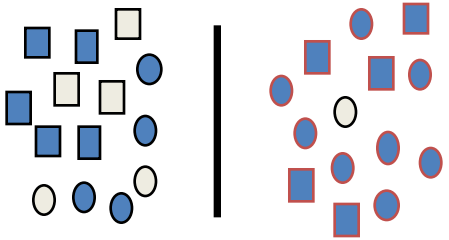
Step 1
Training data set: effect sizes for 6.6 million variants from genome-wide association study



Cases
N = 60K

Controls
N = 120K

Step 2
Validation Dataset: ~125K



Cases
N = 4K

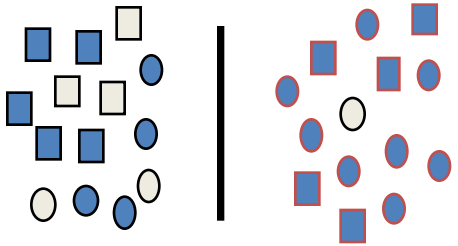
Controls
N = 120K

Genotypes: from arrays + imputation

Khera*, Chaffin*, *bioRxiv* (2017)

Hypothesis: a polygenic score including a genome-wide set of SNPs can identify individuals with risk equivalent to a monogenic mutation

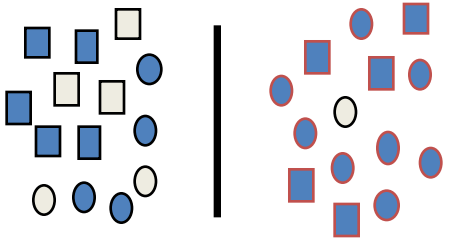
Step 1
Training data set: effect sizes for 6.6 million variants from genome-wide association study



Cases
N = 60K

Controls
N = 120K

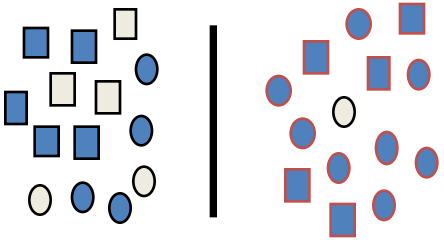
Step 2
Validation
Dataset: ~125K



Cases
N = 4K

Controls
N = 120K

Step 3
Testing
Dataset: ~300K



Cases
N = 8.7K

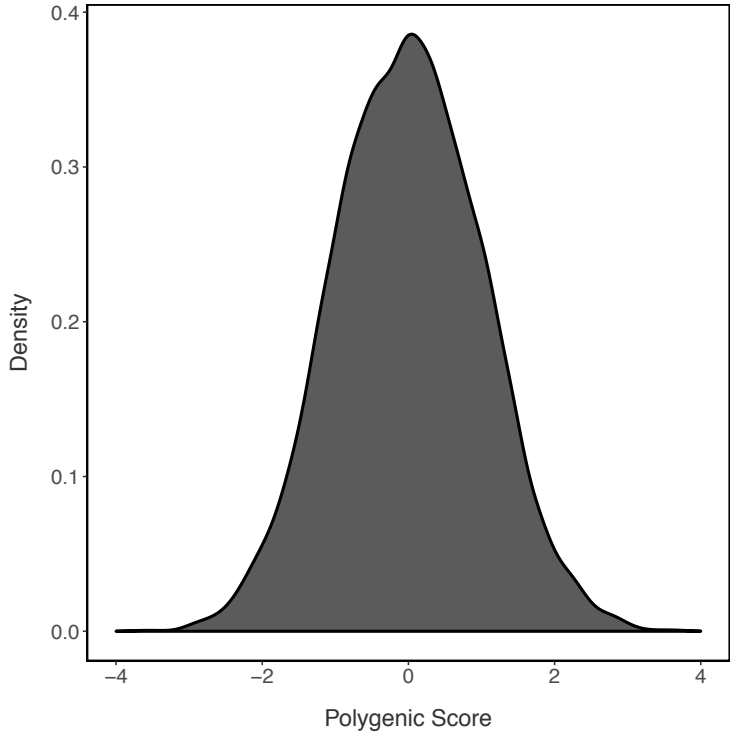
Controls
N = 288K

Genotypes: from arrays + imputation

Khera*, Chaffin*, *bioRxiv* (2017)

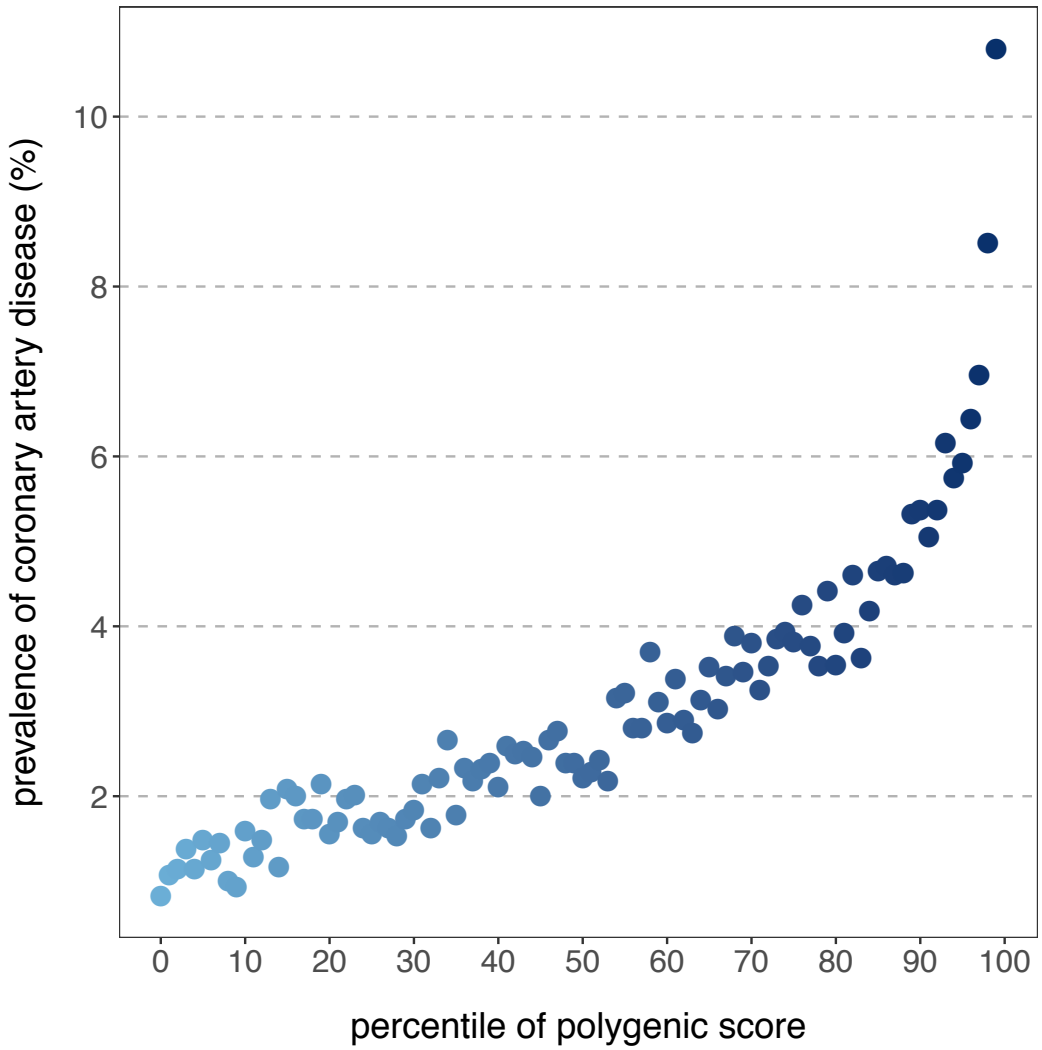
A new quantitative metric of genetic liability to heart attack

Polygenic score of 6.6 million common variants



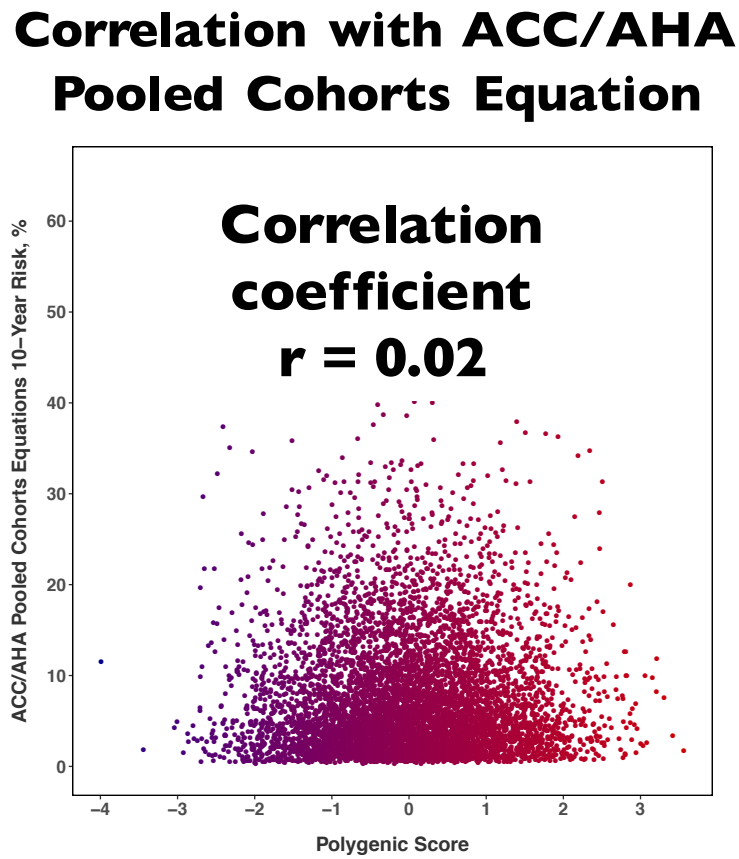
Khera*, Chaffin*, *bioRxiv* (2017)

>20-fold risk gradient across percentile bins of score



Khera*, Chaffin*, *bioRxiv* (2017)

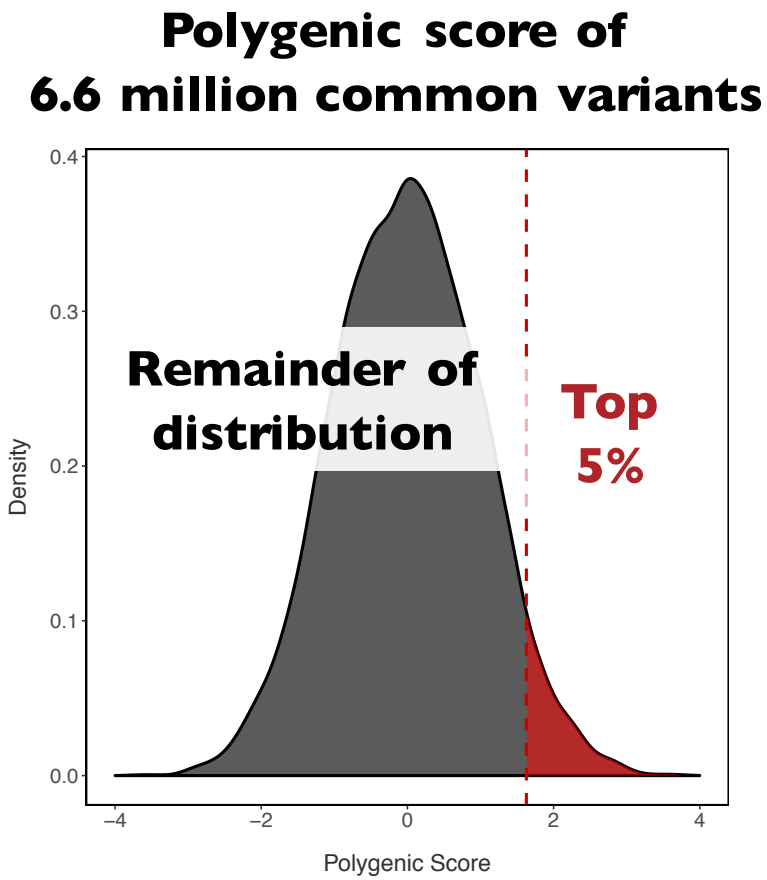
Genome-wide polygenic score: little correlation with currently measured MI risk factors



Khera*, Chaffin*, *bioRxiv* (2017)

Using polygenic model, can we identify group with risk for MI equivalent to monogenic mutations?

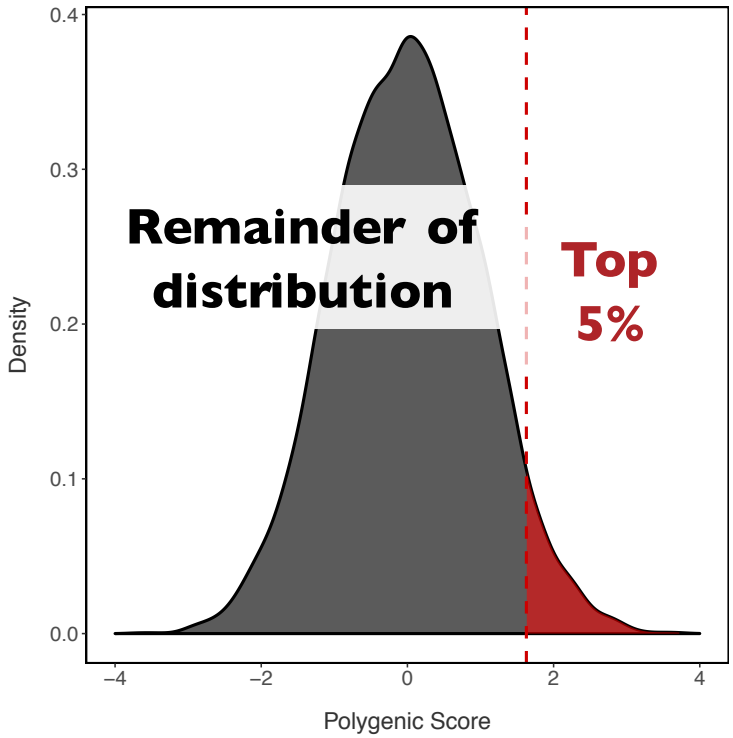
What if we label top 5% tail of distribution as ‘carriers’ and remainder as ‘non-carriers’?



Khera*, Chaffin*, *bioRxiv* (2017)

Top 5% of polygenic MI score: risk equivalent to monogenic mutations

Polygenic score of 6.6 million common variants

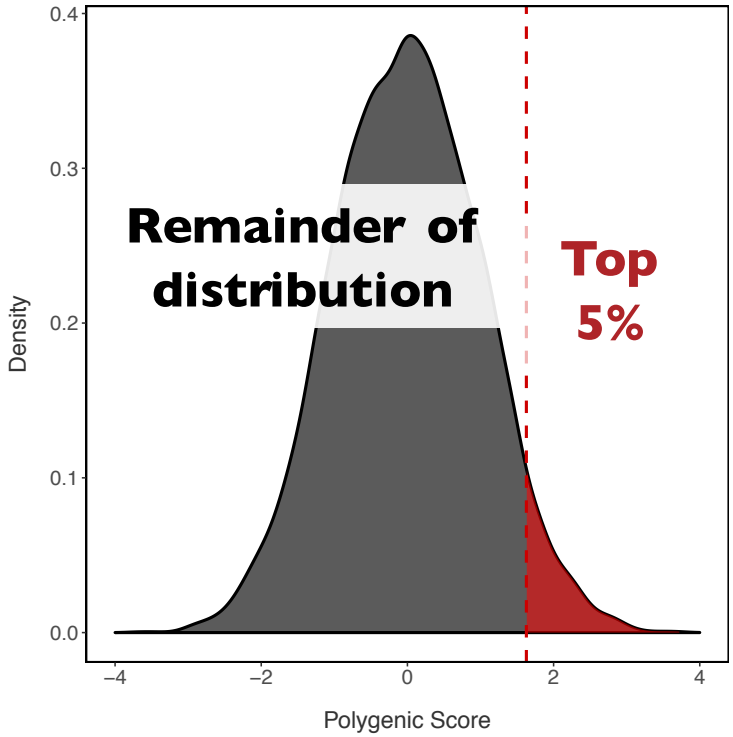


High polygenic score definition	Odds ratio
Top 5%	**3.3**
Top 1%	4.7

Khera*, Chaffin*, *bioRxiv* (2017)

In UK Biobank, top 5% of polygenic score risk equivalent to monogenic mutations but what about external validation?

Polygenic score of 6.6 million common variants



High polygenic score definition	Odds ratio
Top 5%	**3.3**
Top 1%	4.7

Khera*, Chaffin*, *bioRxiv* (2017)

2,081 Early-onset MI patients | 3,761 Controls

MI Cases:

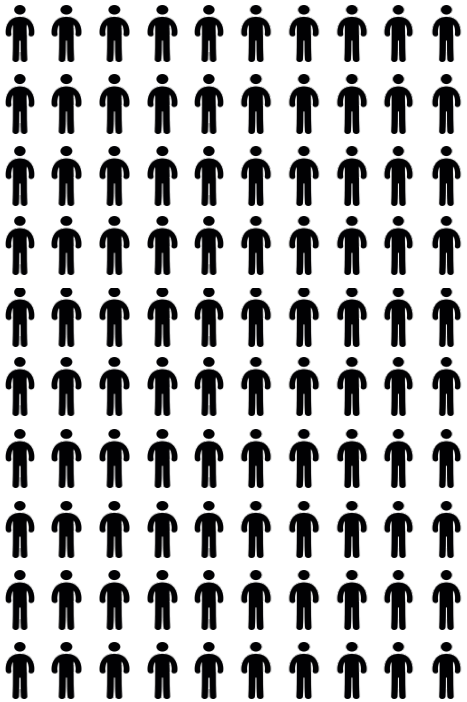
- **VIRGO:** Patients hospitalized across US with first MI at age ≤ 55 years

Controls:

- **MESA:** Multiethnic population free of cardiovascular disease

Contributions of monogenic and polygenic models to early MI

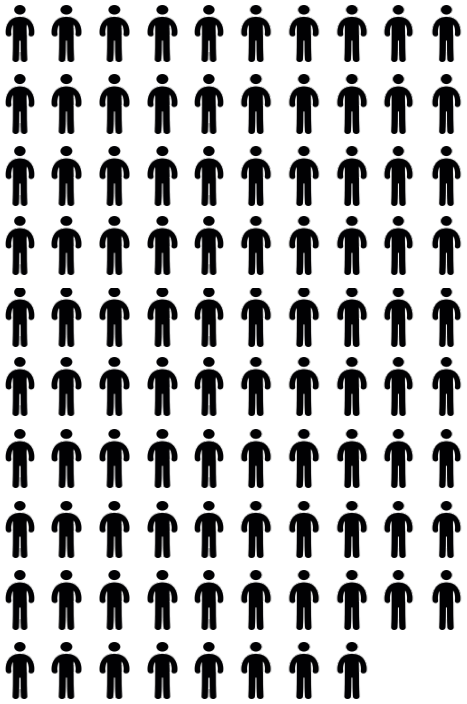
100 patients with myocardial infarction



Khera*, Chaffin*, *under review*

Monogenic familial hypercholesterolemia mutation identified in 1.7% patients -> 3.8-fold increased risk

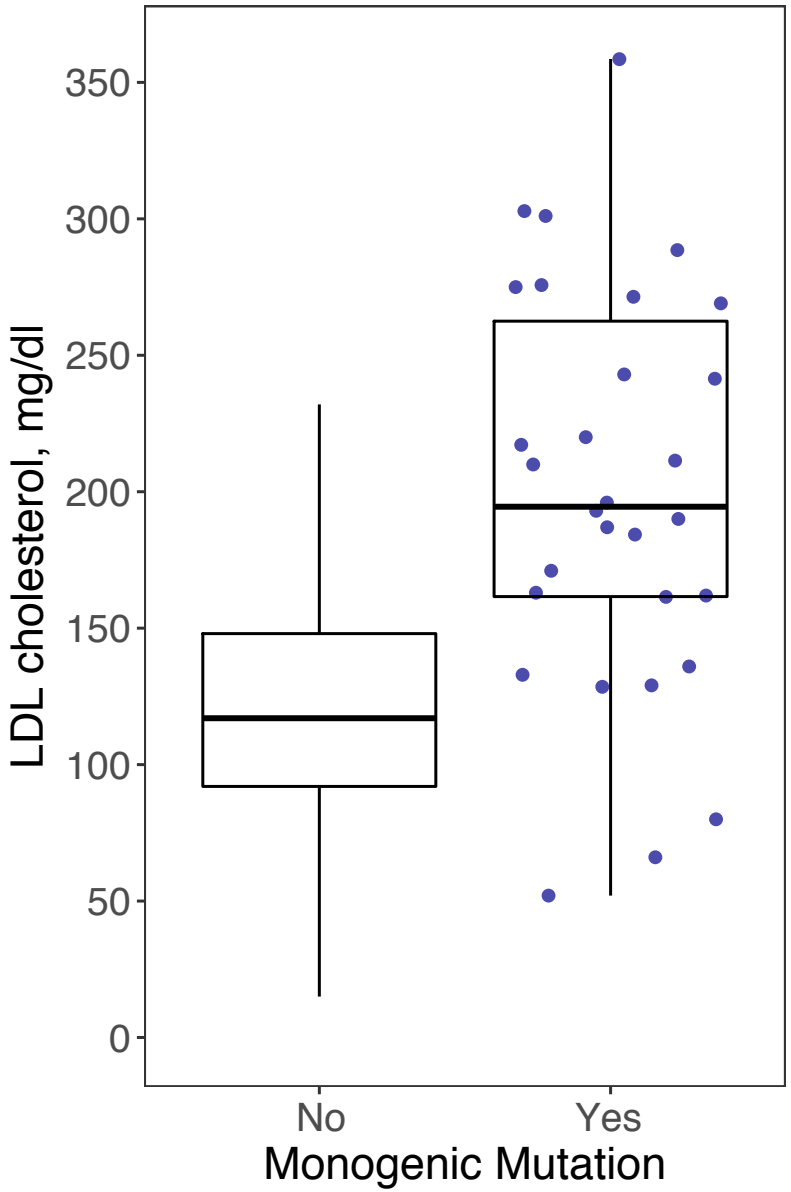
100 patients with myocardial infarction



	↑ Risk
Monogenic	3.8-fold

Khera*, Chaffin*, *under review*

Carriers of familial hypercholesterolemia mutations can be distinguished by high LDL cholesterol



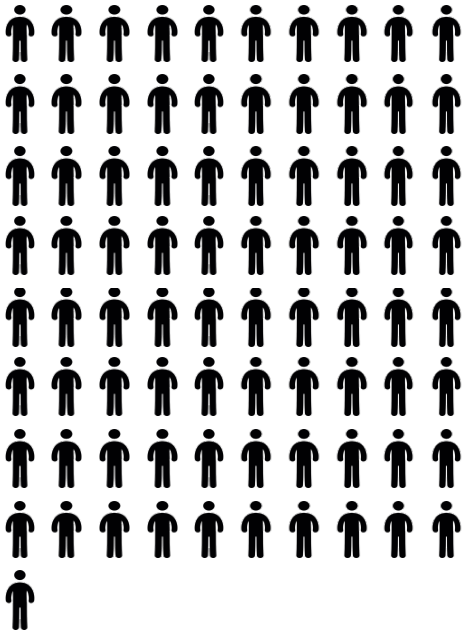
Mean LDL Cholesterol

Carriers: 206 mg/dl

Non-carriers: 124 mg/dl

High polygenic score identified in 17% of patients and confers a 3.7-fold increase in risk

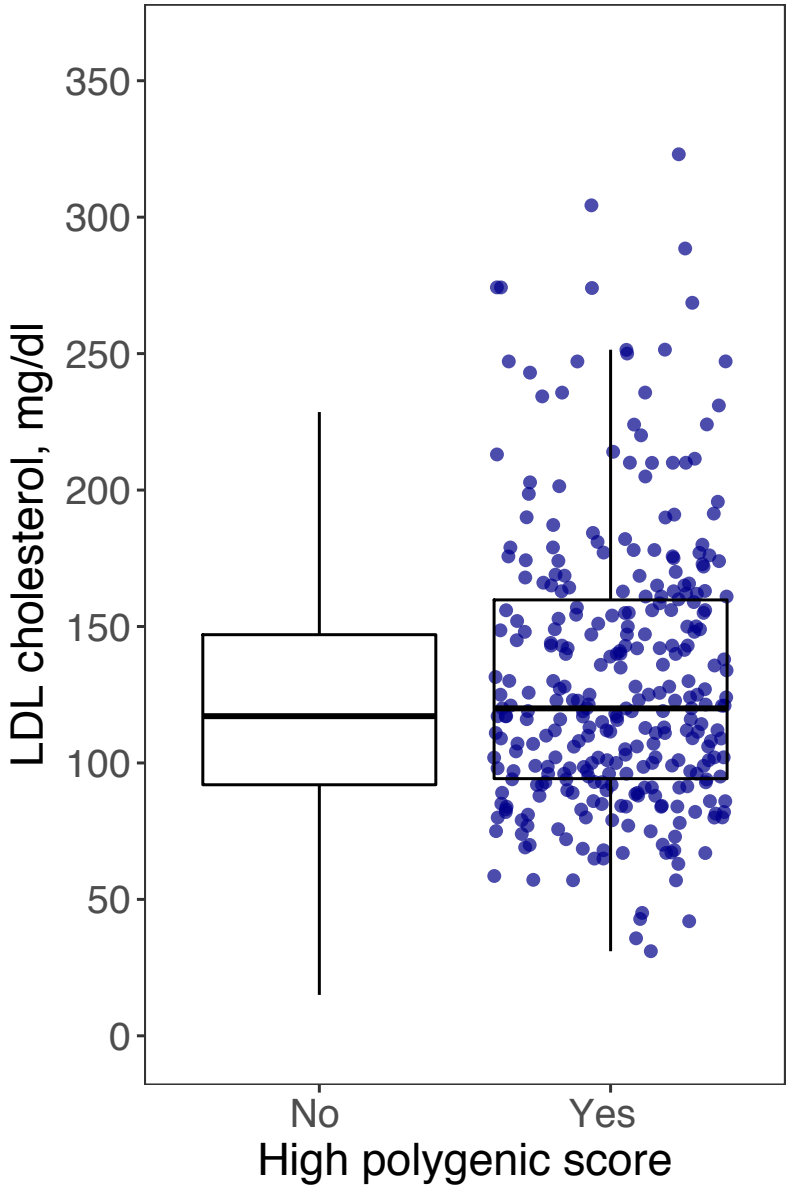
100 patients with myocardial infarction



	↑ Risk
Monogenic	**3.8-fold**
High polygenic	**3.7-fold**

Khera*, Chaffin*, *under review*

High polygenic score individuals can NOT be distinguished by high LDL cholesterol



Mean LDL Cholesterol

High polygenic: 132 mg/dl

Non-carriers: 124 mg/dl

High polygenic score individuals can NOT be distinguished by clinical risk factors

	Neither	High Polygenic Score	FH Mutation
N	1690	355	32
Race, N (%)			
White	1232 (72.9)	281 (79.2)	20 (62.5)
Black	296 (17.5)	35 (9.9)	5 (15.6)
Hispanic	129 (7.6)	32 (9.0)	7 (21.9)
Asian	33 (2.0)	7 (2.0)	0 (0.0)
Male sex, N (%)	563 (33.3)	123 (34.6)	21 (65.6)
Age, years; Mean (SD)	47.6 (5.9)	47.8 (5.7)	46.8 (6.5)
Hypertension, N (%)	1075 (63.9)	243 (68.5)	24 (75.0)
Diabetes, N (%)	593 (35.3)	134 (37.7)	6 (18.8)
Current Smoking, N (%)	848 (50.4)	190 (53.5)	14 (43.8)
Statin Use, N (%)	445 (26.5)	113 (31.8)	15 (46.9)
Lipid Levels, mg/dl			
LDL Cholesterol; Mean (SD)*	122.1 (45.75)	130.4 (51.0)	201.5 (82.0)
HDL Cholesterol; Mean (SD)	40.7 (13.75)	38.9 (13.0)	37.6 (8.1)
Triglycerides; Median (IQR)	133 (91 – 205)	155 (105 – 222)	162 (91 – 246)

Some traditional risk factors are slightly elevated, but not enough to be useful

Polygenic score identifies 10x than monogenic mutations

	Monogenic	Polygenic
Prevalence among early MI cases	1.7%	17%
Odd ratio for MI	3.8	3.7
Mode of detection	↑ LDL cholesterol	**Currently UNAWARE**
Mechanism of risk	apoB lipoproteins	'Gemish'

Monogenic, polygenic contributions to early MI

	Monogenic	Polygenic
Prevalence among early MI cases	1.7%	17%
Odd ratio for MI	3.8	3.7
Mode of detection	↑ LDL cholesterol	**Currently UNAWARE**
Mechanism of risk	apoB lipoproteins	'Gemish'
Intervention	Lifestyle Medications	**?**

Is polygenic risk for MI modifiable? Yes

Lifestyle

↓48%

Khera, *N Engl J Med* (2016)

Medicines

↓44%

Mega*, Stitziel*, *Lancet* (2015)
Natarajan, *Circulation* (2017)

Do those at high polygenic risk derive greater benefit from statin therapy?

Determined polygenic risk score for participants of three statin RCTs to prevent first heart attack

The New England
Journal of Medicine

Volume 333 NOVEMBER 16, 1995 Number 20

PREVENTION OF CORONARY HEART DISEASE WITH PRAVASTATIN IN MEN WITH HYPERCHOLESTEROLEMIA

ARTICLES

Articles

Prevention of coronary and stroke events with atorvastatin in hypertensive patients who have average or lower-than-average cholesterol concentrations, in the Anglo-Scandinavian Cardiac Outcomes Trial—Lipid Lowering Arm (ASCOT-LLA): a multicentre randomised controlled trial

*Peter S Sever, Björn Dahlöf, Neil R Poulter, Hans Wedel, Gareth Beevers, Mark Caulfield, Rory Collins, Sverre E Kjeldsen, Arni Kristinsson, Gordon T McInnes, Jesper Mehlsen, Markku Nieminen, Eoin O'Brien, Jan Östergren, for the ASCOT investigators**

The NEW ENGLAND JOURNAL of MEDICINE

ESTABLISHED IN 1812 NOVEMBER 20, 2008 VOL. 359 NO. 21

Rosuvastatin to Prevent Vascular Events in Men and Women with Elevated C-Reactive Protein

Paul M Ridker, M.D., Eleanor Danielson, M.I.A., Francisco A.H. Fonseca, M.D., Jacques Genest, M.D.,

Evaluate clinical benefit of statin therapy in genetic risk subgroups:

High genetic risk versus all others

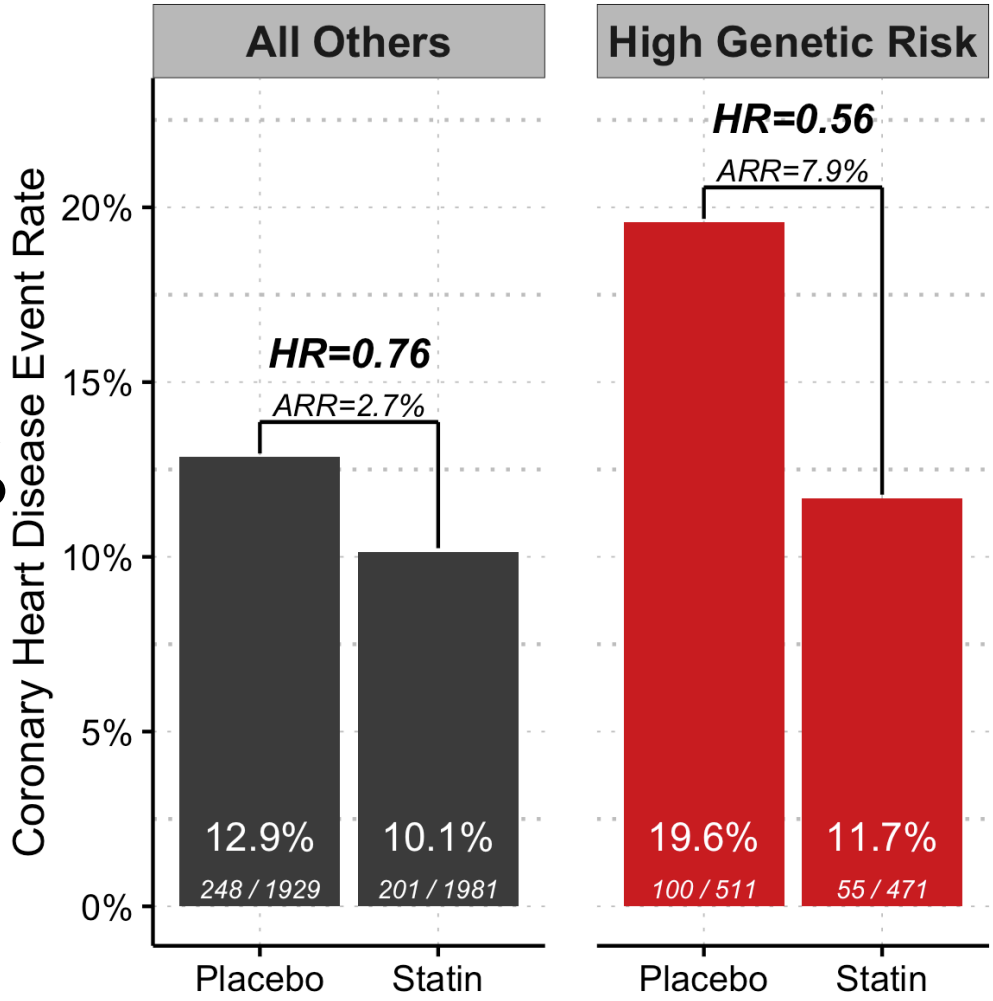
Mega*, Stitziel*, *Lancet* (2015)

Natarajan*, Young*, *Circulation* (2017)

Among those at high polygenic risk, statins confer greater benefit (to prevent first MI)

Pradeep Natarajan

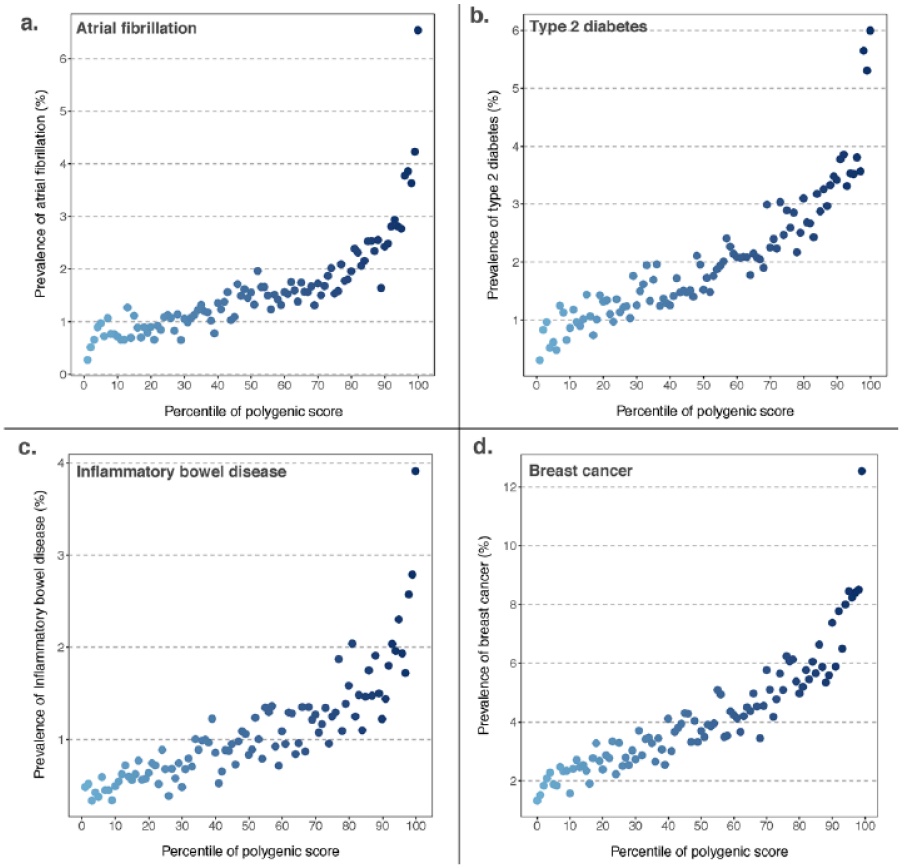
RRR = 24%



RRR = 44%

Natarajan*, Young*, *Circulation* (2017)

Approach works for other common diseases . . . including those without monogenic risk factors



	% of population at >3-fold risk
Atrial fibrillation	**6.1%**
Diabetes	**3.5%**
Inflammatory Bowel	**3.2%**
Breast cancer	**1.5%**

Why much better prediction now?

- Larger genome-wide association studies, more precise effect estimate for each variant
- Better computational methods to create genome-wide polygenic scores
- Larger cohorts to validate and test genome-wide polygenic scores (e.g., UK Biobank, 500K participants with GWAS data)

Conclusions:

- Now possible to score polygenic component to any complex trait (from genotyping array data, simultaneous for many diseases, at birth)

- Those in extremes of score: at risk for disease approaching or exceeding monogenic mutations

- For MI, top 5% tail of polygenic score equivalent risk to monogenic mutations and this risk modifiable by lifestyle, statin

If we care about screening for monogenic MI mutation (Tier 1 CDC), we should also consider polygenic risk score

What is predicted?	Risk for heart attack
Intended target population	Men/women < 55yo
How?	Genome-wide polygenic score (top 5%)
For what purpose?	Statin initiation at early age