GENETIC HEALTH RISKS, LONGEVITY, AND RETIREMENT

RICHARD KARLSSON LINNÉR

NETSPAR INDUSTRY EVENT – 26 SEPTEMBER 2019

VU VRIJE UNIVERSITEIT AMSTERDAM

LOOKING FURTHER

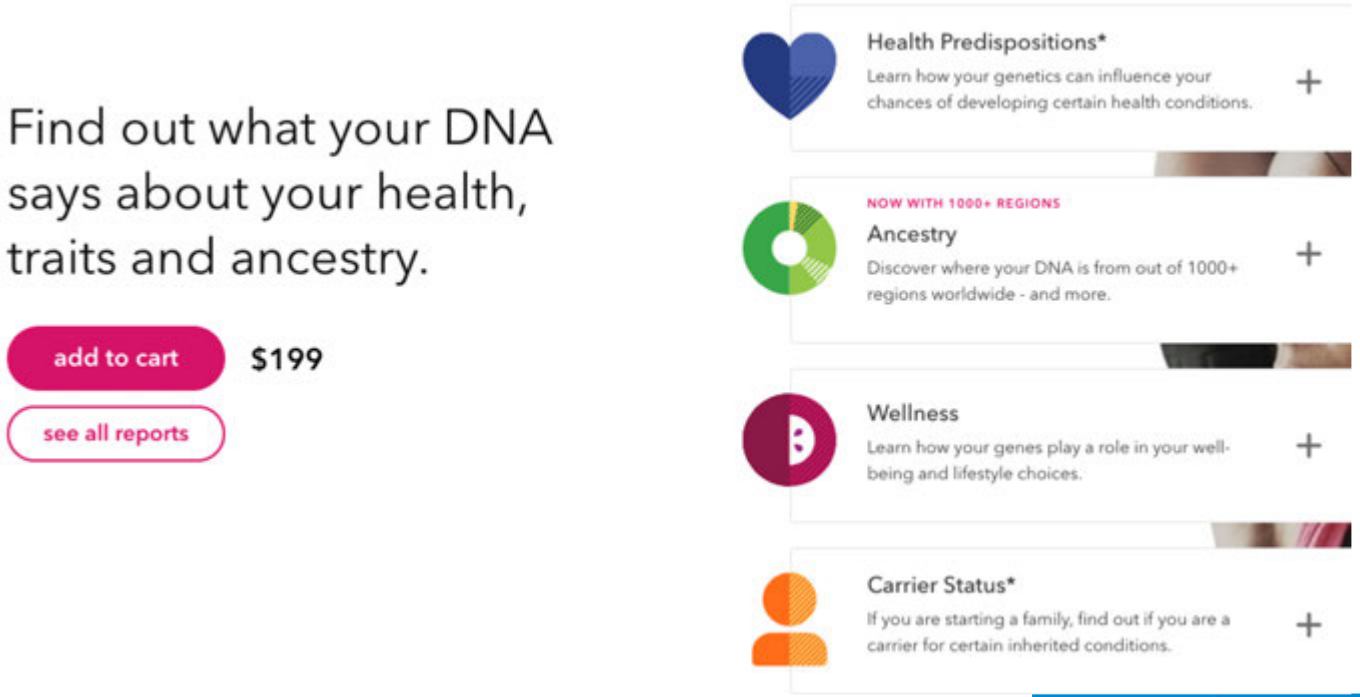
EXPECTATIONS OF HEALTH AND LONGEVITY

- Health expectations influence many economic decisions[1]
  - > Insurance purchases
  - > Investments and savings
  - > Consumption, labor supply, and retirement decisions
- Scholarly interest in factors that shape health expectations
  - > Demographic, socioeconomic, medical, parental lifespan, etc.
- Genes account for much of the variation in health/longevity
  - > But genetic health risks are hitherto unobserved by most people (including our study participants)

1 Seminal paper by Hamermesh. (1985). *Quarterly Journal of Economics*.

GENETIC HEALTH INFORMATION

- Genetic testing is fast becoming accessible and affordable
  - > Limited accuracy today, will increase substantially in the next few years



VU VRIJE UNIVERSITEIT AMSTERDAM School of Business and Economics

GENETIC HEALTH INFORMATION

- Genetic testing is fast becoming accessible and affordable
  - > Limited accuracy today, will increase substantially in the next few years

14 Genetic Risk Reports

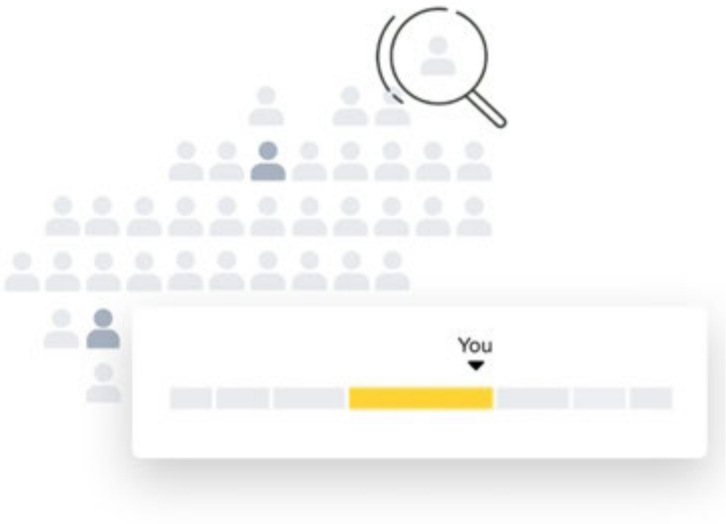
Everyone has some risk of developing genetic conditions. Your reports explain your genetic risk for developing certain conditions, compared to the general population.

Conditions include*:

Heart disease, breast cancer, Alzheimer's disease, and 11 others

Learn more

*Polygenic risk reports for heart disease, breast cancer, and type 2 diabetes are only available for people who are of mainly European ancestry.



VU VRIJE UNIVERSITEIT AMSTERDAM School of Business and Economics

ADVERSE SELECTION

- Insurance industry is concerned about DTC genetic testing[1]
  - Adverse selection could lead to escalating premiums
  - Could threaten the long-term affordability and viability of private insurance
- Fundamental principles of private insurance:
  - Symmetric information about observable risks
  - Actuarial fairness and evidence-based underwriting
- Question whether genetic test results should be disclosed to insurance companies is a current and controversial topic

1 Nabholz & Rechfeld. (2017). Swiss Re Centre for Global Dialogue.

STUDY OVERVIEW

- Preregistered study protocol with Open Science Framework[1]
- Main RQ: How well can genetic predictors for *common* medical conditions and health risks distinguish longevity compared to conventional actuarial risk factors?
- Analyzed the Health and Retirement Study (HRS)
  - > Rich genetic, demographic, socioeconomic, and health data
  - > 9,272 *genotyped* respondents of European ancestry (2,332 deceased)
  - > Mortality selection—healthier, less health-risk behaviors, and long-lived

1 Available at: https://osf.io/c7uem/
2 Also referred to as “expected longevity” or “subjective survival probabilities.”

GENETIC HEALTH RISKS

VU VRIJE UNIVERSITEIT AMSTERDAM

School of Business and Economics

GENETICS OF COMMON DISEASE

- Genetic screening for rare monogenic disorders is not new
- But a majority of NCD deaths are caused by a few *common* medical conditions and health risks[1]
  - > Cardiovascular disease, cancers, diabetes, smoking, etc.
- These conditions are heritable and *polygenic*[2]
  - > Influenced by a very large number of genetic variants with small effects
  - > Most risk-conferring variants are yet unlinked to disease
- Ongoing revolution in genetic discovery of common disease[3]

1 Bloom et al. (2011). *World Economic Forum and the Harvard School of Public Health.*
2 Visscher & Wray. (2016). *Human Heredity.*
3 Mills & Rahal. (2019). *Communications Biology.*

GENOMICS REVOLUTION

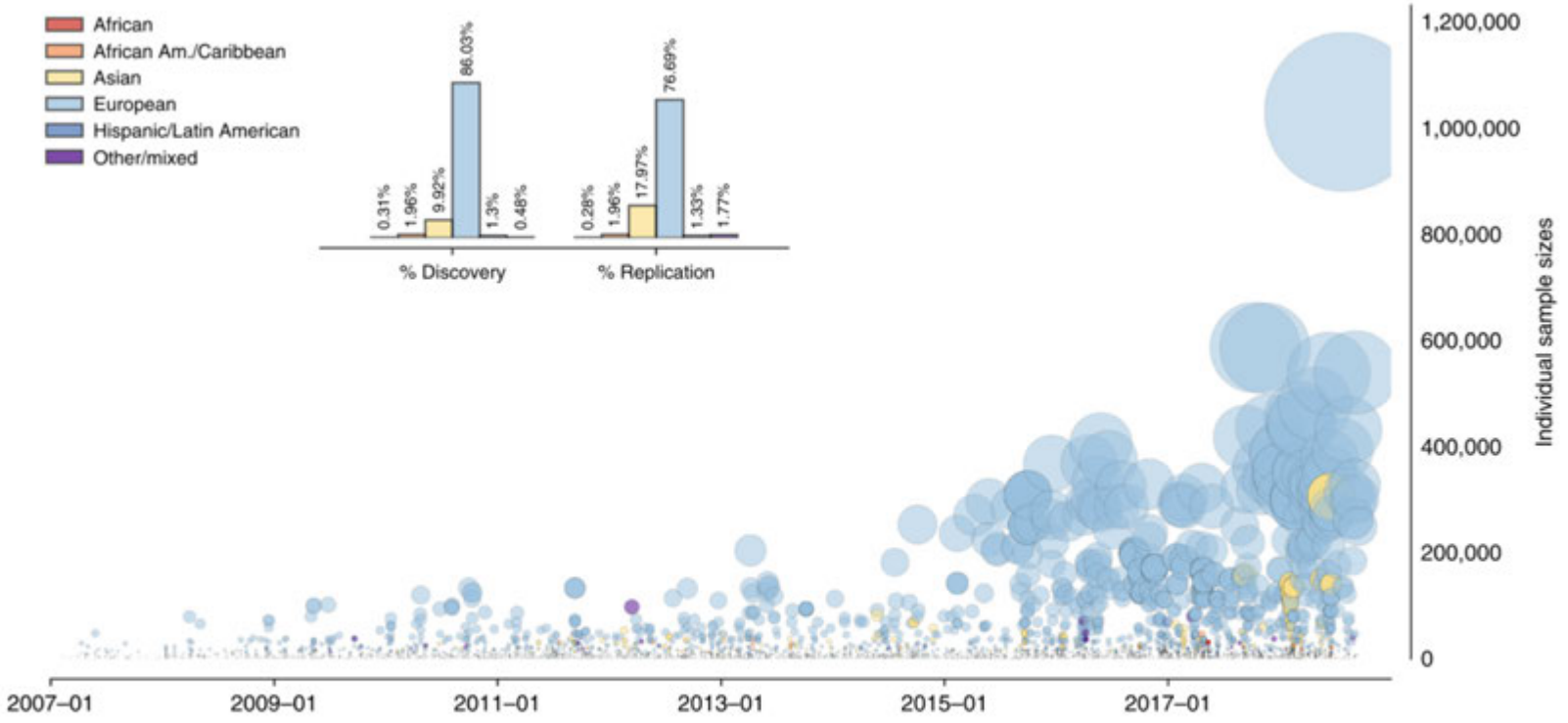
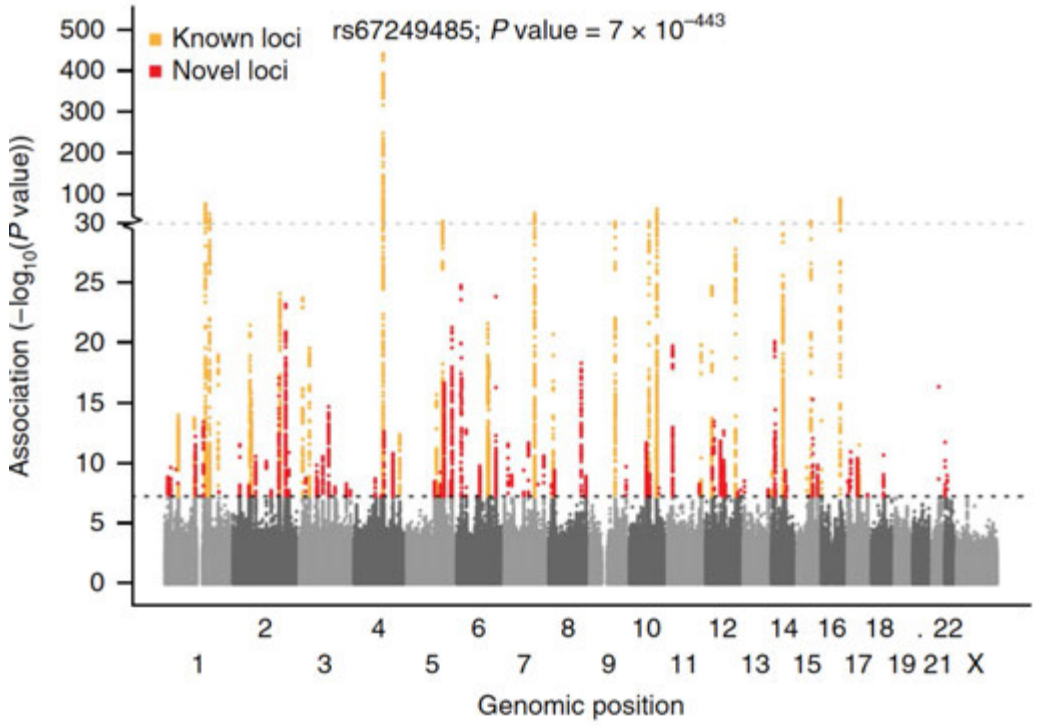


Fig. 2 GWAS Participant Ancestry over Time, 2007-2017. The main panel shows a disaggregation of our broad ancestral categories field, which is a direct mapping from the 17 broad ancestral categories identified in the Catalog. We drop all rows where any proportion of the ancestry is not recorded, and for combinations of ancestries (e.g., European and African) we create a new field: Other/Mixed. The inset aggregates this across the entire sample but partitions the data across discovery and replication phases. 2007-2017 is selected since only 10 entries occurred before 2007 and we have complete information for the year 2017. Source: NHGRI-EBI GWAS Catalog and author mapping

Mills & Rahal. (2019). *Communications Biology.*

GENOME-WIDE ASSOCIATION STUDIES (GWAS)

- GWAS test millions of SNPs for association with an outcome
  - > Recent GWAS of atrial fibrillation in >1 million individuals[1]
  - > Results can stratify severalfold increased risk of developing disease[2]



Linear regression framework:

$$y = \alpha + \beta_j g_{ij} + \boldsymbol{X}\gamma + \varepsilon$$

where g_{ij} is SNP j,
and $\boldsymbol{X}\gamma$ are control variables.

1 Nielsen et al. (2018). *Nature Genetics*.
2 Khera et al. (2018). *Nature Genetics.*

COLLECTION OF GWAS RESULTS

- Performed an extensive search of the GWAS literature
  - Guided by the medical literature on recognized predictors of mortality
  - Restricted to GWAS in at least 100,000 individuals
- Identified 13 GWAS on common medical conditions:
  - Alzheimer's disease, cardiovascular disease, cancers, stroke, etc.
- Identified 14 GWAS on mortality health risks:
  - Blood pressure, BMI, cholesterol, smoking, parental lifespan, etc.
- Average N = 455,000; Largest N > 1 million (atrial fibrillation)

POLYGENIC SCORES

- Polygenic scores summarize genetic risk (or propensity) towards a trait into a genetic predictor
  - > Direct and indirect causal pathways
- We constructed 27 polygenic scores ($\hat{S}_{ik}$):

Linear combination of genetic effects on trait *k*:

$$\hat{S}_{ik} = \sum_{j=1}^{M} \hat{\beta}_{jk} g_{ij}$$

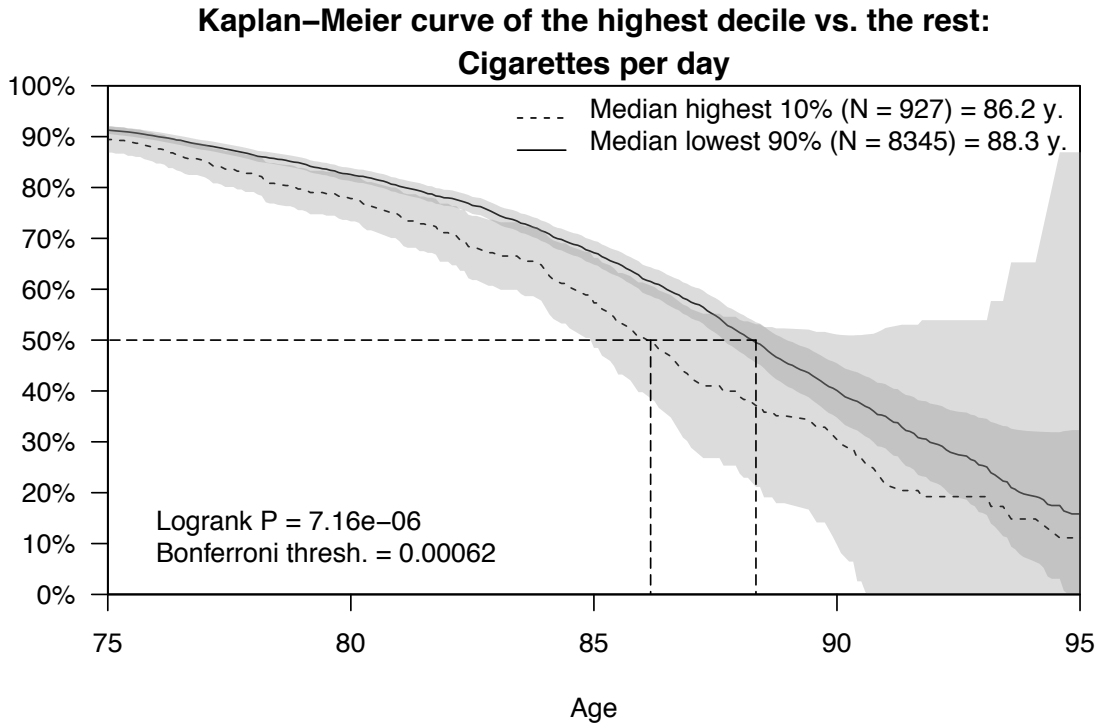
where g_{ij} (the SNPs) are weighed by $\hat{\beta}_{jk}$, the trait-specific GWAS coefficient, and summed across *M* SNPs.

RESULTS


VU
VRIJE
UNIVERSITEIT
AMSTERDAM
School of Business
and Economics

UNIVARIATE SURVIVAL ANALYSIS

- Kaplan-Meier estimation of respondent and parental survival
- ~20 polygenic scores could distinguish survival functions
  - Compared (a) top decile versus the rest; and (b) top versus bottom decile



MULTIPLE REGRESSION OF SURVIVAL

- Estimated four nested Cox proportional-hazards models of respondent survival, with the following variables:

1. all polygenic scores except the score for parental lifespan*;
2. model (1) together with sex-specific birth-year dummies, birth-month dummies, and several demographic and socioeconomic covariates;
3. **model (2) together with the polygenic score for parental lifespan (preferred model);**
4. model (3) together with many covariates from the health risk domain: including BMI, current and former smoker, subjective life expectancy and self-rated health, and 11 categories of diagnosed medical conditions.

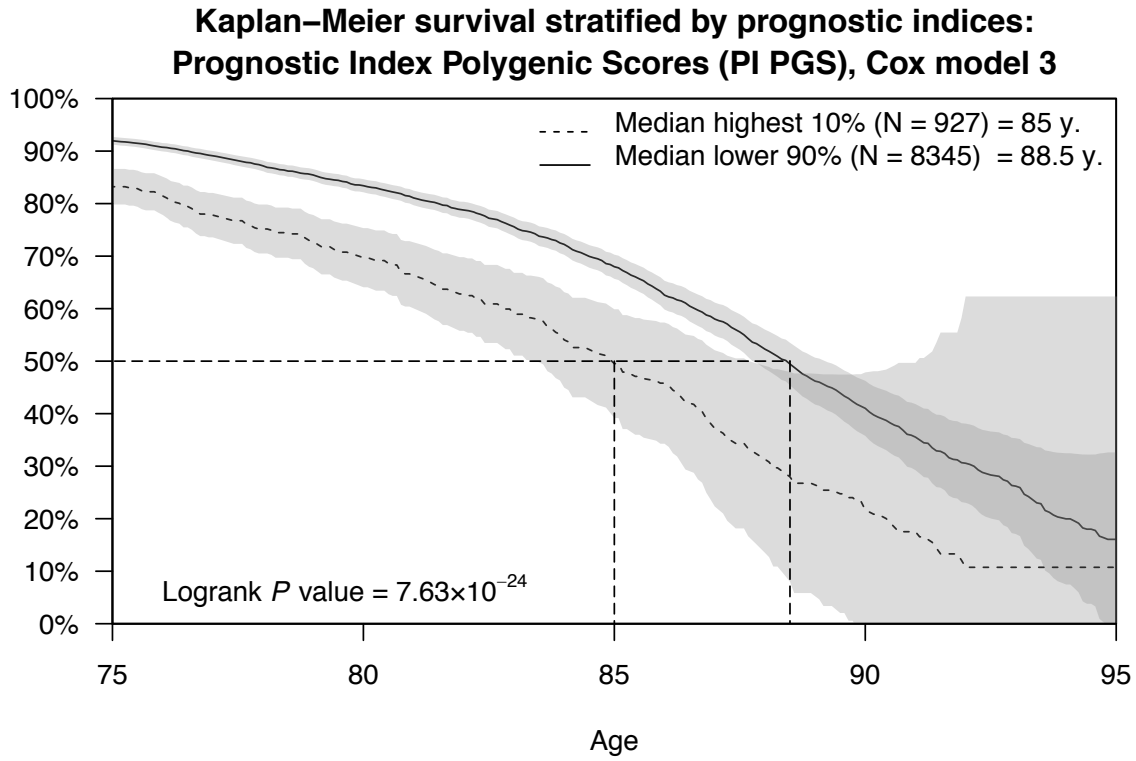
* All models included 10 genetic PCs to control for population stratification. All standard errors were clustered at the household level.

MULTIPLE REGRESSION OF SURVIVAL

- Model (3) is our preferred model for developing a prognostic index that could be evaluated early in life
  - Satisfied model assumptions and fit (Cox-Snell R^2 = 0.231)
- Polygenic scores associated in model (3):
  - Alzheimer's disease ($\hat{\alpha} = 0.052$; P = 0.022)
  - Atrial fibrillation ($\hat{\alpha} = 0.054$; P = 0.019)
  - Cigarettes per day (smoking intensity; $\hat{\alpha} = 0.073$; P = 0.001)
  - Height ($\hat{\alpha} = 0.049$; P = 0.046)
  - Type 2 diabetes ($\hat{\alpha} = 0.054$; P = 0.036)
  - Parental lifespan ($\hat{\alpha} = -0.087$; $P < 0.001$)

PROGNOSTIC INDEX – POLYGENIC SCORES

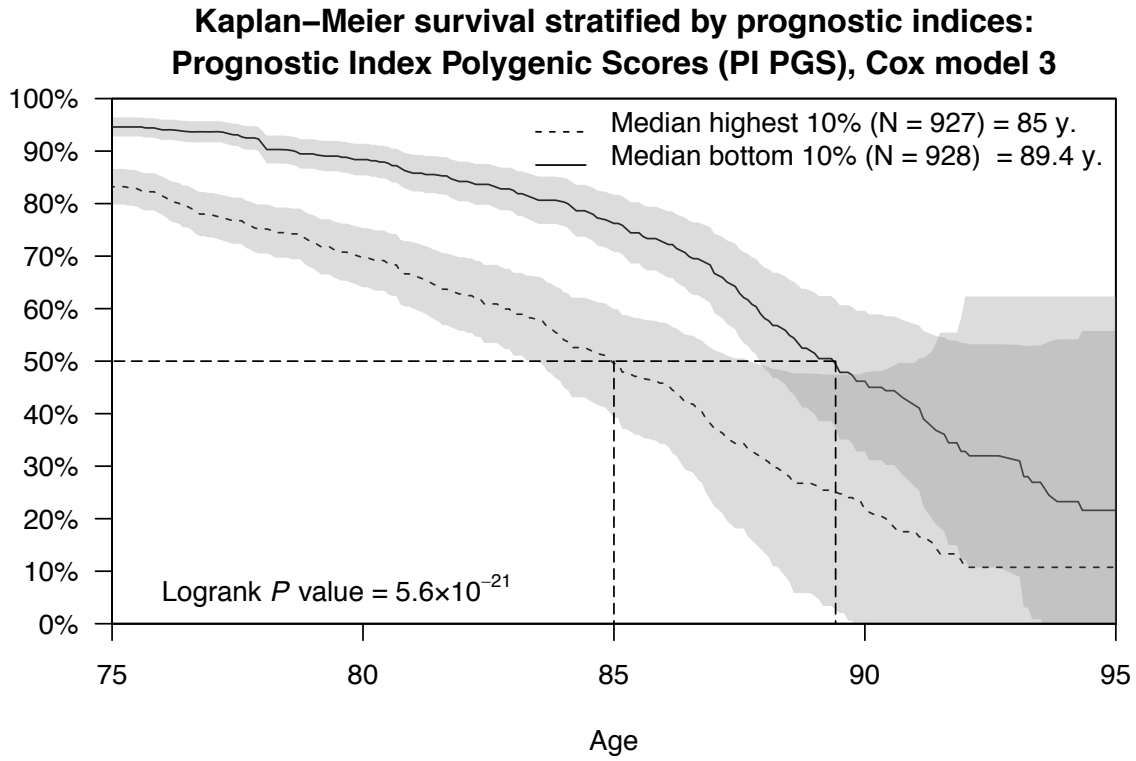
- Prognostic indices (PIs)—linear combinations of sets of regressors, weighed by their Cox coefficients ($\hat{\alpha}$)



- 3.5 y difference in median survival between top decile versus the rest
- 4.4 y difference top vs. bottom

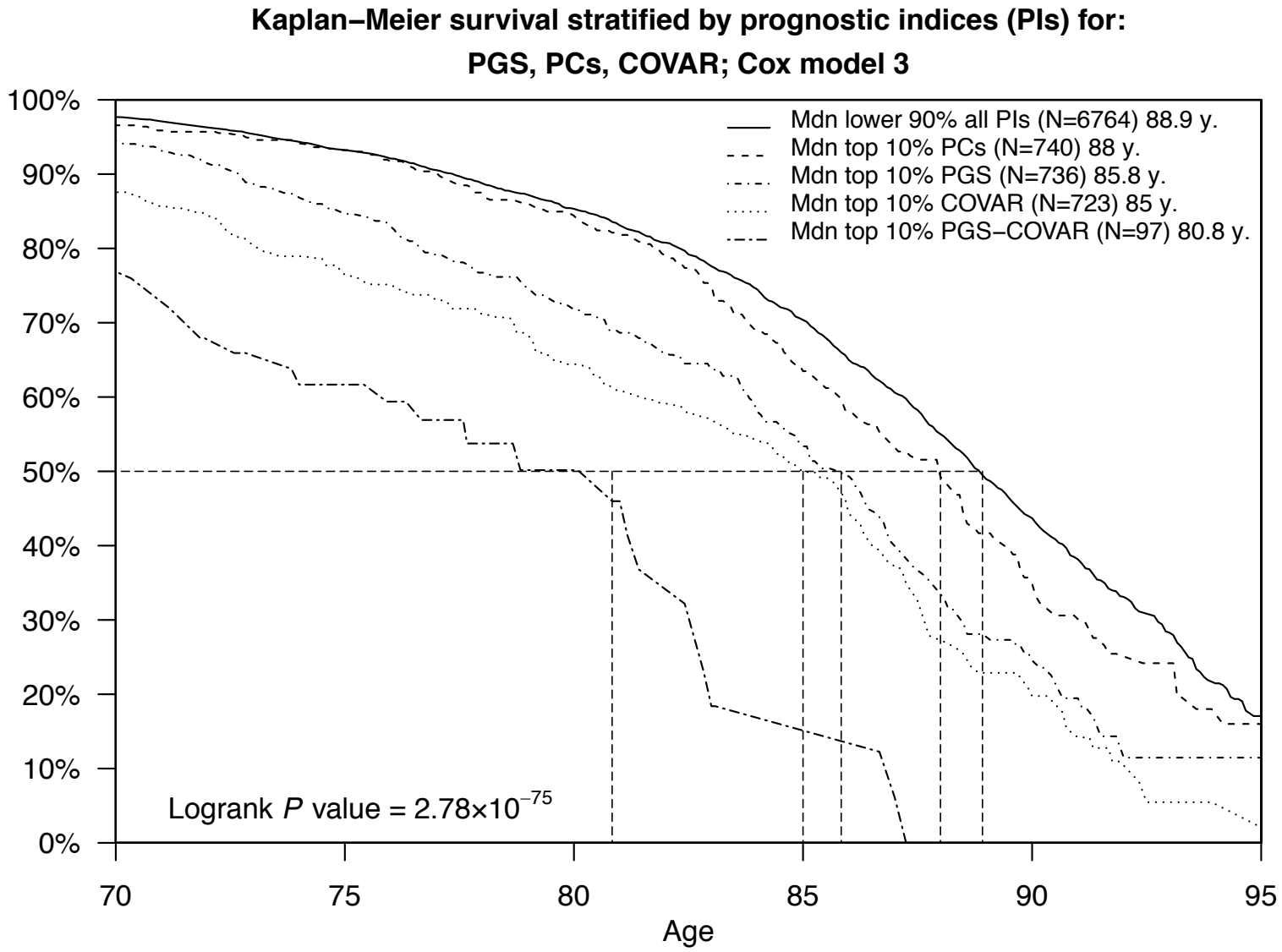
PROGNOSTIC INDEX – POLYGENIC SCORES

- Prognostic indices (PIs)—linear combinations of sets of regressors, weighed by their Cox coefficients ($\hat{\alpha}$)



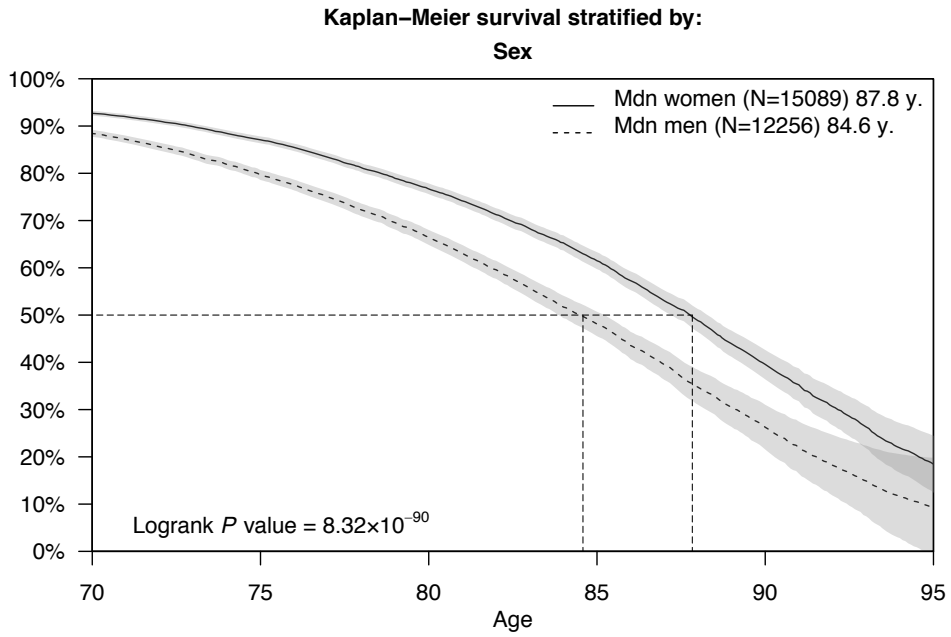
- 3.5 y difference in median survival between top decile versus the rest
- 4.4 y difference top vs. bottom

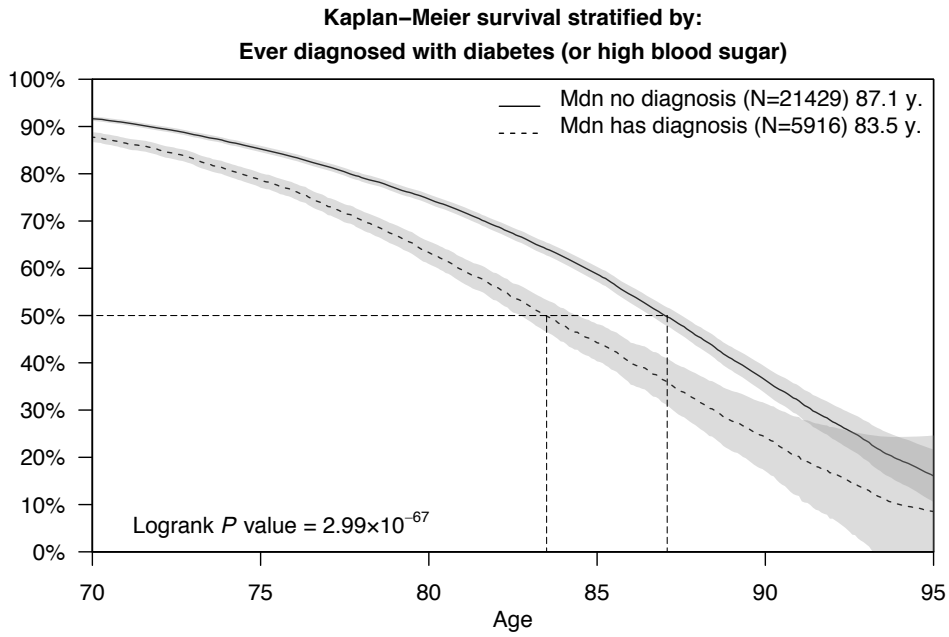
PROGNOSTIC INDEX – ALL COVARIATES



COMPARISON WITH ACTUARIAL RISK FACTORS

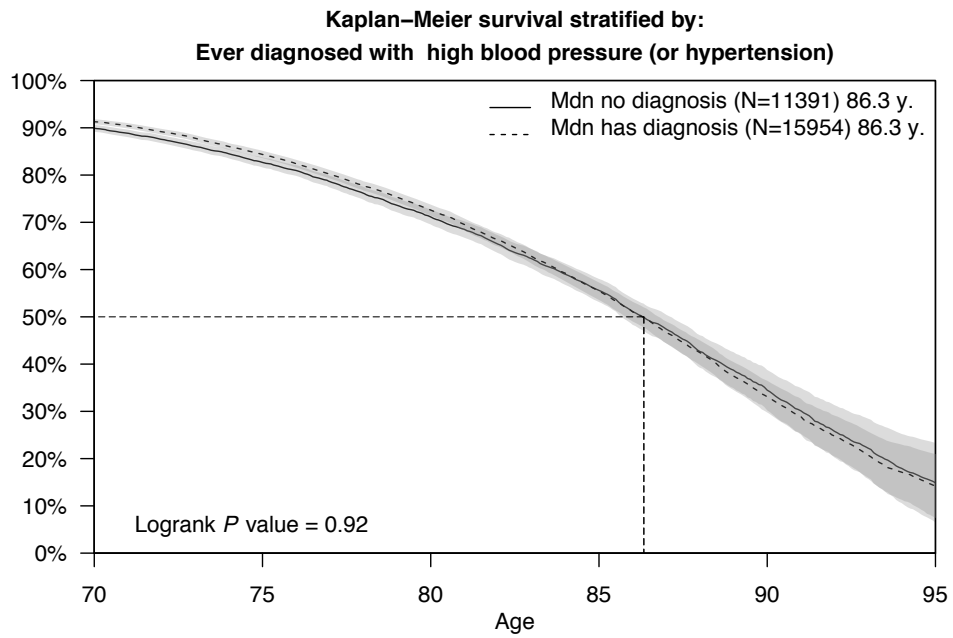
- PI_{PGS} distinguished longevity similar to sex (2.8–3.2y), diabetes (or high blood sugar; 1.7–3.6y), and former smoking (2.5–3.4y)

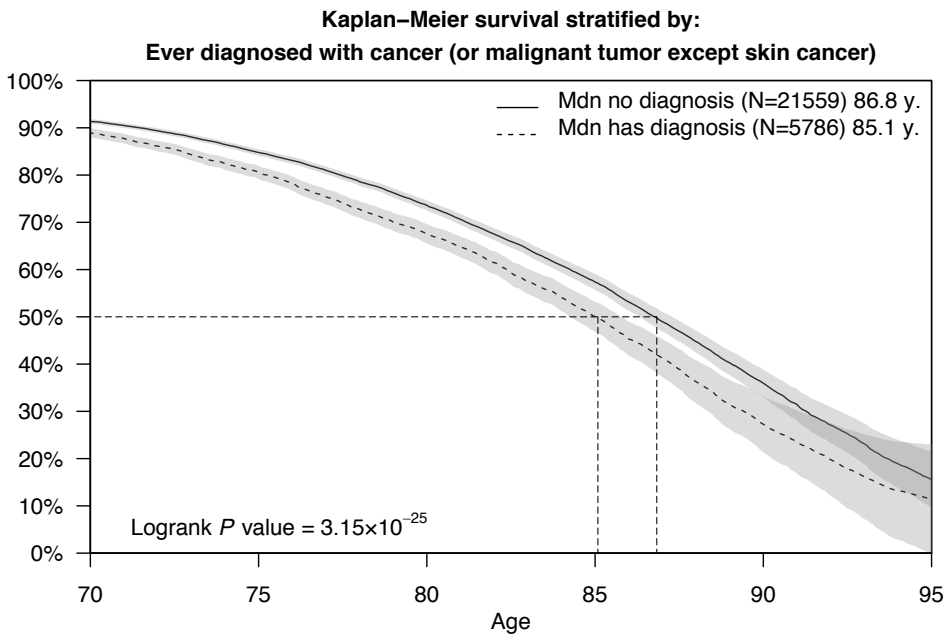




COMPARISON ACTUARIAL RISK FACTORS

- PI_{PGS} distinguished longevity better than education* (1.3–2y), and several medical diagnoses, including cancer (1.2–1.7y)



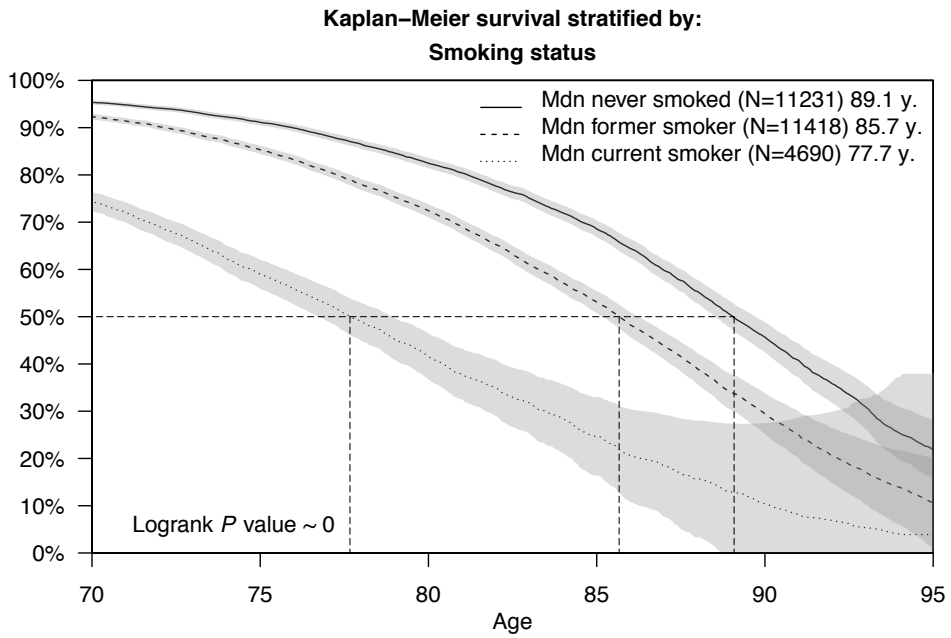


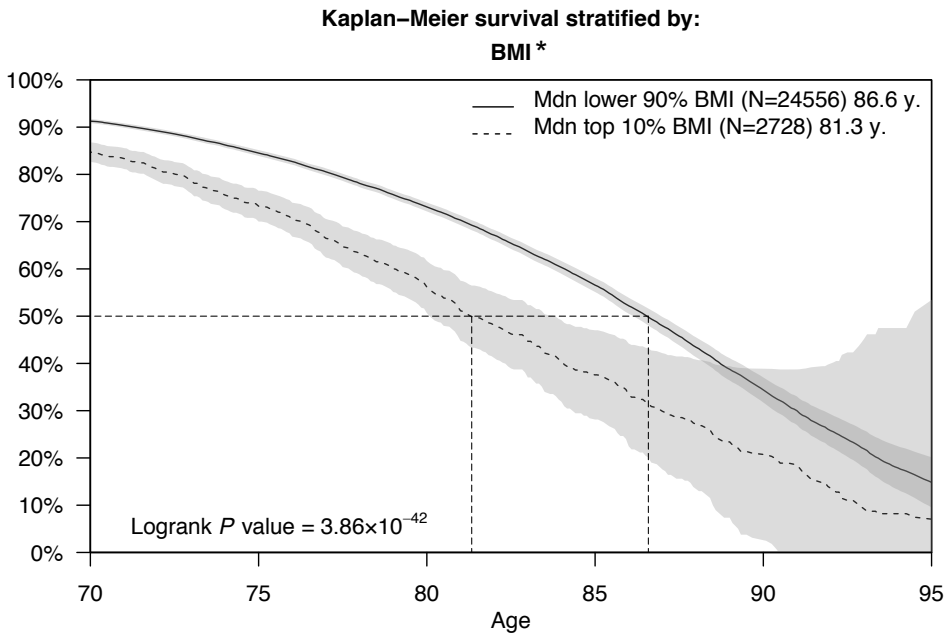
* Top 10% of educational attainment is >=17 years of schooling.

VU VRIJE UNIVERSITEIT AMSTERDAM School of Business and Economics

COMPARISON ACTUARIAL RISK FACTORS

- PI_{PGS} distinguished longevity worse than current smoker (9.9–11.4y) and severe obesity (4.4–5.3y)





* Top 10% of BMI is >38.6. BMI >30 is typically classified as obese, and >35 as severe obesity.

VU VRIJE UNIVERSITEIT AMSTERDAM School of Business and Economics

EXPECTATIONS OF HEALTH AND LONGEVITY

- RQ2: Are the genetic predictors associated with subjective life expectancy and self-rated health?
- Can indicate whether the unobserved genetic risk is still captured by manifested medical conditions and overall health
- PI_{PGS} was associated with both measures (model 3 covariates):
  - Subjective life expectancy: –0.052 per SD (SE = 0.008; $P = 4.02 \times 10^{-11}$)
  - Self-rated health*: OR = 1.215 per SD (*SE(log[OR])* = 0.019; $P = 1.03 \times 10^{-24}$)
- Suggests that the genetic risk has indeed manifested

* Coded as ”1. Excellent; 2. Very Good; 3. Good; 4. Fair; 5. Poor”.

RETIREMENT-RELATED ECONOMIC OUTCOMES

- RQ3: Are the genetic predictors associated with economic outcomes related to retirement?
- PI_{PGS} was associated with (model 3 covariates):
  1. Whether health limits work: OR = 1.2 per SD ($P = 1.08 \times 10^{-13}$)
  2. Self-reported probability of having a work-limiting health problem in the next 10 years: 1.079 pp per SD ($P = 0.0008$)
  3. Retirement satisfaction*: OR = 1.13 per SD ($P = 1.08 \times 10^{-6}$)
  4. % waves covered by long-term care insurance[1]: –0.007 per SD ($P = 0.007$)
  5. Financial planning horizon**: OR = 0.953 per SD ($P = 0.016$)

* Coded as ”1. Very; 2. Moderately; 3. Not at all”.
** Coded as “1. next few months; 2. next year; 3. next few years; 4. next 5-10 years; 5. longer than 10 years".
1 The effect size corresponds to ~2 months shorter coverage per SD.

CONCLUSIONS

- Genetically-informed research design estimated that genetic health risks could jointly distinguish up to 4.4 y of median survival
  - > Lower bound because of mortality selection and still limited GWAS *N*
  - > Yet, distinguished longevity comparable to conventional actuarial risk factors
- The unobserved genetic risk:
  - > was partly captured in subjective life expectancy, self-rated health, and manifested medical conditions
  - > was associated with some retirement outcomes, including coverage by long-term care insurance but not life-insurance coverage

THANK YOU!

Questions?

r.karlssonlinner@vu.nl
p.d.koellinger@vu.nl

We gratefully acknowledge:
The Health and Retirement Study
Netspar
Bas Werker
Anja De Waegenaere
Aysu Okbay
Casper Burik
the SURF cooperative

VU VRIJE UNIVERSITEIT AMSTERDAM | School of Business and Economics