Netspar
Network for Studies on Pensions, Aging and Retirement

NETSPAR INDUSTRY SERIES

Genetic health risks, insurance, and retirement

Richard Karlsson Linnér
Philipp D. Koellinger

DESIGN PAPER 142

DESIGN PAPERS are part of the **refereed Industry Paper Series**, which are refereed by the Netspar Editorial Board. Design Papers discuss the design of a component of a pension system or product. A Netspar Design Paper analyzes the objective of a component and the possibilities for improving its efficacy. These papers are easily accessible for industry specialists who are responsible for designing the component being discussed. Authors are allowed to give their personal opinion in a separate section. Design Papers are presented for discussion at Netspar events. Representatives of academic and private sector partners, are invited to these events. Design Papers are published at the Netspar website.

Colophon
Netspar Design Paper 142, April 2020

Design
B-more Design

Lay-out
Bladvulling, Tilburg

Editors
Frans Kooymans, Frans Kooymans-Text and Translation
Netspar

CONTENTS

This study is accompanied by the following supplementary materials:

- *Supplementary Tables: https://www.netspar.nl/assets/uploads/D20200415_Netspar-Design-Paper-1422-Supplementary-Tables.xlsx*
- *Supplementary Figures: https://www.netspar.nl/assets/uploads/D20200415_Netspar-Design-Paper-1422-Supplementary-Figures.docx*

Affiliations
Richard Karlsson Linnér – Vrije Universiteit Amsterdam
Philipp D. Koellinger – Vrije Universiteit Amsterdam

Abstract

Genetic health risks, insurance, and retirement

Genetic health information is becoming increasingly accessible and affordable. Widespread genetic testing for diseases can have major implications for consumer behavior in insurance, pension, and annuity markets, and insurance providers are concerned about adverse selection and escalating premiums. Here we estimate to what extent measures of genetic liability (polygenic scores) are informative about differences in longevity, health expectations, and economic behavior. We construct polygenic scores for 27 common diseases and mortality risks among 9,272 respondents of the Health and Retirement Study, a longitudinal household survey of elderly Americans, by leveraging results from genetic studies in hundreds of thousands of participants. Survival analysis suggests that the median lifespan is up to 4.5 years shorter in the highest decile of genetic risk. In particular, the scores for Alzheimer's disease, parental lifespan, and smoking intensity are strongly associated with survival. Our results highlight that polygenic scores can already be combined to stratify survival similar to or better than some conventional actuarial risk factors, including sex, and we find that greater genetic risk is negatively associated with various economic behaviors, including long-term care insurance. We conclude that the rapid developments in genetic epidemiology pose new challenges for regulating consumer genetics and insurance markets, and that these need to be urgently considered by policymakers.

Samenvatting

Genetische gezondheidsrisico's, verzekeringen en pensioen

Genetische gezondheidsinformatie wordt steeds toegankelijker en betaalbaarder. Uitgebreide genetische tests zouden aanzienlijke gevolgen kunnen hebben voor het gedrag van consumenten ten aanzien van verzekeringen, pensioenen en de verzekeringsmarkt, en aanbieders van verzekeringen maken zich zorgen over negatieve selectie en stijgende premies. Hier geven we een inschatting van het samengaan van genetische voorspellende factoren – *polygene scores* – met levensduur, gezondheidsverwachtingen en economische gevolgen. We creëren polygene scores voor 27 gangbare ziektes en mortaliteitsrisico's, met behulp van 9.272 respondenten van een onderzoek naar gezondheid en pensionering (Health and Retirement Study) en vergroting van de resultaten met behulp van gegevens van genetische onderzoeken onder in totaal honderdduizenden deelnemers. Uit overlevingsanalyses, gecorrigeerd voor leefstijlfactoren, medische diagnoses en sociaal-economische variabelen, blijkt dat de mediane overleving in het hoogste deciel van het genetische risico tot wel 4,5 jaar korter is. Vooral ten aanzien van de scores voor de ziekte van Alzheimer, de levensduur van ouders en roken is er sprake van een sterk verband met overleving. Uit onze resultaten komt naar voren dat polygene scores al zodanig kunnen worden gecombineerd dat ze overleving beter stratificeren dan sommige conventionele actuariële risicofactoren, waaronder geslacht en vroeger gerookt hebben. Verder stellen we vast dat er voor een hoger genetisch risico een negatief verband bestaat met verschillende economische gevolgen, waaronder bij verzekeringen voor langdurige zorg. We concluderen dat de snelle ontwikkelingen op het gebied van genetische epidemiologie nieuwe uitdagingen bieden voor het reguleren van de genetica van consumenten en verzekeringsmarkten en dat beleidsmakers zich hierover dringend zouden moeten buigen.

1. Introduction

Economic behavior can be influenced by expectations about future health and longevity (defined here as length of life)[1]. For example, people who expect to live long and healthy lives appear to save more for retirement, delay their retirement, and desire more comprehensive long-term care insurance[2–7]. A simple survey measure that captures these expectations—subjective life expectancy—has been found to predict actual longevity and various economic outcomes[1,5,16,8–15]. Thus, there is a scholarly interest in identifying factors that can influence expectations about future health and longevity[7]. Previously identified factors include age, sex, ethnicity, education, occupation, lifestyle, medical conditions, parental lifespan, and psychological features such as optimism[5–13]. Here we investigate a new source of health information that has hitherto not been observed by most people (including our participants) but that is fast becoming widely available, namely genetic testing[24–31]. Specifically, we estimate to what extent measures of genetic liability for common diseases and mortality risks, which are called *polygenic scores*, can already be used to distinguish differences in longevity, subjective life expectancy, self-rated health, and economic behaviors, including insurance coverage.

1.1 Genetic testing and insurance behavior

Until recently, genetic testing was largely restricted to healthcare[32]. Presently, however, companies advertise and offer directly to consumers genetic tests to predict the lifetime risk of developing specific diseases[27,33,34]. While the prognostic accuracy of such tests is still limited, this is expected to increase substantially in the coming years, and there are recent reports of polygenic scores that are comparable in accuracy to conventional clinical risk factors[35–40]. Notably, the consumer genetics market is growing exponentially, and it already has several million customers worldwide. Many of them report that exploring their DNA for health information was their main motivation to purchase a genetic test[26,29,30,41–43]. In addition, a growing number of third-party genetic interpretation services offer supplementary health reports to persons who submit their raw genetic data for reanalysis[28,29,43,44]. Some of these services also offer predictions of longevity[45]. Access to genetic health information is thus becoming an affordable and widely available reality, and genetic health risks can now be elucidated early in life, many years before signs or symptoms of disease become evident. Many experts anticipate that novel and widespread access to genetic health information will have major impact on consumer behavior in various markets, in particular for

financial products related to health and longevity, such as insurance, annuities, and pensions[31,46–49].

A handful of studies have investigated whether a positive genetic test result influences insurance purchasing behavior[47,50–53]. While the results are mixed, some studies have indeed found such a relationship. However, a caveat is that most studies on this topic have tested stated preferences rather than actual insurance holdings, and only in small samples. Furthermore, it appears that the genetic literacy of the general public is still limited, even among the highly educated[54–56]. It is therefore also conceivable that genetic health reports may lead to no behavioral changes at all due to disregard or misinterpretation, or, alternatively, to changes that could be considered unexpected or inappropriate[29,43,57,58]. Thus, research is needed to determine the effect that more widespread genetic testing has on economic behavior.

Representatives of various insurance providers have expressed concerns about the viability of certain insurance markets in this era of consumer genetics[30,42,59]. Their main worry is a situation in which insurance applicants with private knowledge of their genetic health risks are sanctioned to withhold that information from underwriters. Conversely, there is the danger of genetic discrimination, meaning denial of access to health care or financial support. Many governments have therefore taken a regulatory stance that favors consumer privacy over corporate interests, by limiting the rights of insurance providers to request and use genetic information to determine premium levels[60–63]. In countries that lack such legislation, the industry has often applied self-regulation by means of voluntary moratoriums[30,64]. Still, experts and stakeholders agree that such sanctioned non-disclosure could be a threat to the fundamental insurance principles of symmetric information and actuarial fairness[31,62,65], and that this can lead to adverse selection and escalating premiums[31,62]. Over time, this development could threaten the affordability and viability of private insurance markets.

Overall, the question of whether or not genetic test results should be disclosed to insurance providers is a pressing controversial topic of societal relevance[60,66–71]. But before polygenic scores can even be considered for use in underwriting, their accuracy and relevance as actuarial risk factors must be determined. Recently, an international expert group of researchers and insurance stakeholders called for more research on this topic[67]. The main objective of this study is to estimate how well polygenic scores for common medical conditions and mortality risks can stratify survival functions compared to conventional actuarial risk factors.

1.2 The current study

We conducted our analyses according to a preregistered analysis plan[a]. In summary, we generated polygenic scores for 13 common medical conditions (including Alzheimer's disease, atrial fibrillation, and type 2 diabetes), and 14 mortality risks (including blood pressure, cholesterol, and smoking), applying these to 9,272 Health and Retirement Study (HRS) respondents of European ancestry. In a series of survival analyses, we found that the combined polygenic scores could distinguish up to 4.4 years of median lifespan, and 2.4 to 4.1 years when statistically conditioned on a range of lifestyle factors, medical conditions, and socioeconomic variables. The polygenic scores were also associated with subjective life expectancy and self-rated health, which suggests that unobserved genetic risks were actually partly observed and captured by these measures. Finally, greater genetic risk was negatively associated with several economic outcomes, including fewer waves of long-term care insurance, while no association was found with life-insurance coverage.

1.3 Genetic health risks

It is well known that genetic factors contribute substantially to the risk of developing a disease[72–74]. At the overall population level, genetic differences may account for as much as 30% of the variation in longevity[75–79]. Genetic tests that screen for severe but rare single-gene disorders have been routine in clinical care for decades. There are already thousands of such tests on the market, and new tests are regularly introduced[80,81]. But only few people are affected by rare genetic disorders[82], and their contribution to the mortality burden from non-communicable diseases (NCDs) is limited, particularly in adults[31]. On the other hand, several common and substantially heritable medical conditions, including cardiovascular disease (CVD), cancer, and diabetes, account for a majority of all NCD deaths[36,83–85]. In addition, a few prevalent and heritable health risks, such as high cholesterol and smoking, also lead to a considerable mortality burden via NCDs[74,86–89]. Accordingly, we restricted our investigation to medical conditions and mortality risks that are common in the overall population.

1.3.1 Genome-wide association studies and polygenic scores

Common medical conditions are rarely, if ever, caused exclusively by a single gene. Instead, they are *complex traits*, influenced by variations in a large number of genes, with individually minute effects on disease risk[36,72,90]. This so-called *polygenicity* also

a The analysis plan is available at *https://osf.io/qzx6p/*

applies to mortality risks and lifestyle factors[90–92]. Still, in combination these small genetic effects add up to the heritability of a trait, which for longevity and many common diseases accounts for ~20–50% of the variance[74,77,93]. In recent years, rapid progress has been achieved in the effort to identify genetic variants that contribute to complex traits[73,94]. The *genome-wide association study* (GWAS) is currently the main method used to identify associated variants[95].

A typical GWAS tests millions of single-nucleotide polymorphisms (SNPs) one at a time, for association with a trait[96]. SNPs refer to a single genetic base pair, which could be considered the simplest type of variation in the genome. Many recent GWAS studies have been performed on hundreds of thousands of participants, and a few in more than a million; larger studies are expected in the near future[73,94,97]. For example, the most recent GWAS on longevity studied more than five hundred thousand people[79]. To date, the GWAS literature has successfully linked tens of thousands of SNPs with hundreds of common diseases, health risks, and lifestyle behaviors, with a respectable replication record[73,94,95,98].

The coefficients estimated in GWAS studies can be used out-of-study to construct indices of genetic risk[96]. The resulting polygenic scores are simple linear combinations of a person's genotype, weighed by each SNP's trait-specific effect[90,96,99,100]. These scores can have substantial predictive accuracy when the underlying effects are estimated in a well-powered study[36,39,101]. For example, recent studies show that polygenic scores can stratify individuals with increased risk of disease when the underlying GWAS coefficients have been estimated in hundreds of thousands of participants[36,102]. Their utility is even greater when analyzed jointly[103], or together with other observable factors, such as family and medical history[99].

1.4 Study aims

Only a handful of studies have investigated whether polygenic scores can stratify survival functions[77–79,104–106]. The largest effect reported thus far is a difference in median lifespan of 3.5 years, achieved by comparing the top versus bottom decile of a score for parental lifespan. (Parental lifespan is a common proxy for longevity in genetic research.) However, estimates are scarce, and most have been obtained without conditioning on confounders, such as income, smoking, or medical conditions. Therefore, the first aim of this study is to estimate how well polygenic scores for common medical conditions and mortality risks can distinguish lifespan in another sample, and to benchmark their performance to conventional actuarial risk factors[62]. As it is essentially unknown how much information polygenic scores can add to conventional risk factors, we will adjust extensively for potential confounders. Next, our

second aim is to investigate whether polygenic scores are associated with subjective life expectancy and self-rated health. Such an association may indicate whether the underlying genetic risks, which we assume not to be directly observed by the respondents, may nonetheless be observed indirectly and captured by these health measures.

Furthermore, there is little information as to whether polygenic scores for common medical conditions and mortality risks are associated with economic outcomes. Thus far, we are only aware of a single recent study that has investigated this, namely that by Shin, Lillard, and Bhattacharya (2019), which tested whether a polygenic score for Alzheimer's disease was associated with wealth composition in the HRS. That study indeed found that greater genetic risk was associated with less wealth[107]. Therefore, our third and final aim is to investigate whether a broader set of polygenic scores might explain variation in various retirement-related economic outcomes, including life insurance and long-term care insurance. Associations with various insurance products may indicate whether there is already some self-selection based on the unobserved genetic risk.

1.5 Insurance – assumptions, definitions, and delimitations

In the empirical analyses, we focus on (a) *life insurance*, defined here as insurance cover that pays a benefit upon the death of the insured, and (b) *long-term care insurance*, defined here as insurance that covers assistance with activities of daily living upon illness or disability, as these two types of insurance have been ascertained consistently in most respondents and waves of the HRS. We also briefly discuss (c) *pension insurance, life annuities, and longevity annuities/insurance*. These we consider together as a class of insurance that pays a benefit in the circumstance that the insured survives until, or continues to live past, a certain age. Thus, for ease of discussion, category (c) includes a range of products that are tied to the survival, rather than the death, of the insured. We classify *defined contribution pension schemes* under (c), since technically the accrued funds are liquidated upon retirement to buy a life annuity that is often lifelong[b]. We do not cover *defined benefit pensions schemes*, since these are being discontinued in most countries[108]. Unfortunately, we could not study (c) empirically because we lacked suitable measures or adequate sample size. Also, we do not study private health insurance since this displayed hardly any variation in coverage. Finally, we do not focus on combined products, such as pension insurance with a death benefit component, since these are basically combinations of

b Lifelong annuities are mandatory for pension schemes in the Netherlands[174].

several simpler products, nor do we study group schemes, as these are not underwritten based on individual risks[109].

We assume that the expected value of any kind of insurance that is tied to the health, life, or death of a person should theoretically be influenced by observable genetic risks. In the circumstance that an applicant has private knowledge of an increased risk that is not reflected in the premium (or benefit), then that insurance will be considered either cheap or expensive (that is, actuarially unfair), depending on the product in question[31]. For example, a life insurance policy would be considered cheap by an applicant with private knowledge of reduced life expectancy (genetic or otherwise). Conversely, a pension insurance or life annuity policy would be deemed expensive. Thus, depending on the product, it may be in the interest of the applicant to either withhold or reveal knowledge of genetic risks. At the same time, in practice, not all observable risks can be insured. Reasons for this include negligible influence on mortality, lack of data to accurately define a fair premium, or legal ramifications[62]. For example, in many countries, a controversial topic subject to debate is whether private insurance should be exempted from anti-discrimination law, including price discrimination based on disability or genetic predisposition[62,110].

2. Methods

2.1 Data

We analyzed the Health and Retirement Study (HRS), a longitudinal household survey of elderly Americans, conducted biannually since 1992[111]. The purpose of the HRS is to facilitate studies on how the socioeconomic environment is related to health and aging, for which the study participants have given broad consent. In 2006, the HRS also started collecting genotype data[112]. We analyzed the publicly available HRS Longitudinal File 2016 (v1), curated by the RAND Corporation[111], together with the restricted-access genetic data that are available upon request from the National Center for Biotechnology Information (NCBI) database of Genotypes and Phenotypes (dbGaP)[113]. Overall, thirteen waves of data were available, spanning the years 1992–2018.

The HRS Longitudinal File contains rich data from the following domains: demography; family; health and medical; education, occupation, income, and wealth; and retirement. We extensively searched these and narrowed down a selection of about thirty variables that we considered important to include as covariates. Because of the unbalanced panel structure and many missing observations, we did not implement a panel data model. Instead, to vastly increase the sample size we collapsed the panel structure into a cross-section in the following way: binary, ordinal, and categorical variables were assigned the most frequently occurring value across the waves; continuous variables were assigned the median. All dollar amounts were converted to 2016 values. We report the full set of variables and sample descriptive statistics in **Table 1**.

2.1.1 Respondent inclusion criteria

We restricted our analyses to respondents who self-reported to be of "White/Caucasian" and "Not Hispanic" ancestry, which we hereafter refer to as European[c]. The reason for this inclusion criterion is that the vast majority of GWAS studies have been performed in individuals of European descent, which drastically limits the possibility of constructing accurate polygenic scores in non-Europeans[114–116]. We analyzed the second release of the HRS genotype data, which had been imputed with the 1000 Genomes Project phase 1 (version 3) reference panel[117]. The second release includes 15,620 genotyped respondents, of which 10,958 report to be of European ancestry.

c In this study, we adhere to the definition of European ancestry that is frequently used in genetic epidemiology, which distinguishes "Hispanic/Latin American" as a separate ancestral group[94].

Next, we generated genetic principal components (PCs) to identify genetic outliers[118,119]. Specifically, we projected the European-ancestry subsample of the 1000 Genomes phase 3 (version 5) reference panel[117] onto the PCs and excluded respondents that had a value on any of the first four PCs that exceeded the range of the reference panel by more than 10%. In total, 10,701 individuals remained after this step. Thereafter, as recommended[118], we re-estimated ten genetic PCs in the more homogenous HRS subsample, which were later included as covariates to control for population stratification. Population stratification refers to incidental covariation between allele frequencies and the outcome of interest, which can induce bias if not adjusted for[118,120]. To be extra cautious, we excluded respondents that were more than five standard deviations from the mean on any of the newly generated PCs. This procedure removed 105 additional outliers. In total, 10,596 respondents remained at this stage. In summary, we proceeded conservatively to try to minimize the chance of bias because of ancestry admixture[115].

2.1.2 Mortality selection and restriction of birth years

The HRS first started collecting genotype data in 2006, about 14 years after the first wave was conducted. Thus, the genetic data could be subject to non-random selection based on mortality. In a thorough investigation of mortality selection in the HRS, Domingue et al. verified that the genotyped HRS respondents were indeed healthier, displayed fewer health-risk behaviors, and lived longer than the overall sample[112].

We confirmed that mortality selection exists by comparison of Kaplan-Meier survival functions (**Supplementary Figure 1**). Based on recommendations by Domingue et al., and to exclude birth years with fewer than hundred observations with non-missing birth and death data, we restricted all further analyses to individuals born between 1920 and 1955. At this stage, we retained 9,272 respondents, of which 2,332 had died by the most recent wave. We hereafter refer to these 9,272 genotyped respondents as our "study sample". We acknowledge that mortality selection will most likely lead to underestimation of the influence of the polygenic scores, in particular for mortality risks that manifest before old age and that could have contributed to the mortality selection, such as cardiovascular disease[89].

Table 1. Sample descriptive statistics

Variable	Type	Wave	European HRS respondents		European HRS genotyped study sample	
			N	%, Mean (SD), or range	*N*	%, Mean (SD), or range
Panel A. Demographics						
Sex	Categorical	1–13	27,345		9,272	
– Male			12,256	44.82%	4,036	43.53%
– Female			15,089	55.18%	5,236	56.47%
Birth year	Categorical	1–13	27,345	1890–1995	9,272	1920–1955
HRS birth cohort	Categorical	1–13	27,345		9,272	
– 1. AHEAD, born before 1924			6,112	22.35%	531	5.73%
– 2. the Children of Depression (CODA), born 1924-1930			3,403	12.44%	1,530	16.50%
– HRS, born 1931-1941			7,438	27.20%	3,695	39.85%
– War Babies (WB), born 1942-1947			2,677	9.79%	1,527	16.47%
– Early Baby Boomers (EBB), born 1948-1953			2,666	9.75%	1,642	17.71%
– Mid Baby Boomers (MBB), born 1954-1959			2,583	9.45%	347	3.74%
– Late Baby Boomers, born 1960–1965			2,016	7.37%	0	0.00%
– Born after 1965			450	1.65%	0	0.00%
Deceased by the last wave	Binary	1–13	27,345		9,272	
– Yes			11,300	41.32%	2,332	25.15%
– No			16,045	58.68%	6,940	74.85%
Deceased before genotype data collection (before 2006)	Binary	1–13	27,345		9,272	
– Yes			5,940	21.72%	0	0.00%
– No			21,405	78.28%	9,272	100.00%
Census region, 1992–2014	Categorical	1–13	25,507		9,272	
– Northeast			4,372	17.14%	1,477	15.93%
– Midwest			6,957	27.27%	2,690	29.01%
– South			9,646	37.82%	3,400	36.67%
– West			4,532	17.77%	1,705	18.39%
Years of schooling (17+ years coded as 17)	Continuous	1–13	27,199	12.8 (2.8)	9,246	13.2 (2.5)
Veteran status	Binary	1–13	27,333		9,266	
– Yes			6,440	23.56%	2,354	25.40%
– No			20,893	76.44%	6,912	74.60%
Married or partnered in any wave	Binary	1–13	27,345		9,272	
– Yes			21,220	77.60%	8,042	86.73%
– No			6,125	22.40%	1,230	13.27%
Divorced or separated in any wave	Binary	1–13	27,345		9,272	
– Yes			3,705	13.55%	1,373	14.81%
– No			23,640	86.45%	7,899	85.19%
Widowed in any wave	Binary	1–13	27,345		9,272	
– Yes			7,890	28.85%	2,846	30.69%
– No			19,455	71.15%	6,426	69.31%
Number of children	Continuous	1–13	26,381	2.4 (1.7)	9,263	2.6 (1.6)
Panel B. Health						
Body Mass Index (BMI)	Continuous	1–13	27,284	27.1 (5.5)	9,268	27.7 (5.2)
Subjective life expectancy	Continuous	1–13	23,802	1.2 (1.8)	9,265	1.04 (0.8)
Self-rated health	Categorical	1–13	27,341		9,272	
– 5. Poor			2,586	9.46%	410	4.42%
– 4. Fair			4,834	17.68%	1,242	13.40%
– 3. Good			8,639	31.60%	3,044	32.83%
– 2. Very good			8,195	29.97%	3,439	37.09%
– 1. Excellent			3,087	11.29%	1,137	12.26%

Variable	Type	Wave	European HRS respondents		European HRS genotyped study sample	
			N	%, Mean (SD), or range	*N*	%, Mean (SD), or range
Alcoholic drinks per week	Continuous	1–13	25,706	2.8 (6.4)	9,271	2.8 (5.5)
Ever smoker	Binary	1–13	27,345		9,272	
– Yes			16,111	58.92%	5,285	57.00%
– No			11,234	41.08%	3,987	43.00%
Current smoker	Binary	1–13	27,339		9,272	
– Yes			4,690	17.15%	1,233	13.30%
– No			22,649	82.85%	8,039	86.70%
Ever diagnosed with high blood pressure (or hypertension)	Binary	1–13	27,345		9,272	
– Yes			15,954	58.34%	6,193	66.79%
– No			11,391	41.66%	3,079	33.21%
Ever diagnosed with diabetes (or high blood sugar)	Binary	1–13	27,345		9,272	
– Yes			5,916	21.63%	2,365	25.51%
– No			21,429	78.37%	6,907	74.49%
Ever diagnosed with cancer (or malignant tumor except skin cancer)	Binary	1–13	27,345		9,272	
– Yes			5,786	21.16%	2,310	24.91%
– No			21,559	78.84%	6,962	75.09%
Ever diagnosed with chronic lung disease (except asthma such as chronic bronchitis or emphysema)	Binary	1–13	27,345		9,272	
– Yes			4,275	15.63%	1,522	16.42%
– No			23,070	84.37%	7,750	83.58%
Ever diagnosed with heart conditions (such as heart attack, coronary heart disease, angina, congestive heart failure, or other heart problems)	Binary	1–13	27,345		9,272	
– Yes			9,539	34.88%	3,548	38.27%
– No			17,806	65.12%	5,724	61.73%
Ever diagnosed with stroke (or transient ischemic attack)	Binary	1–13	27,345		9,272	
– Yes			3,944	14.42%	1,312	14.15%
– No			23,401	85.58%	7,960	85.85%
Ever diagnosed with emotional, nervous, or psychiatric problems	Binary	1–13	27,345		9,272	
– Yes			1,462	5.35%	352	3.80%
– No			25,883	94.65%	8,920	96.20%
Ever diagnosed with arthritis (or rheumatism)	Binary	1–13	27,345		9,272	
– Yes			16,576	60.62%	6,618	71.38%
– No			10,769	39.38%	2,654	28.62%
Ever diagnosed with Alzheimer's disease	Binary	1–13	27,345		9,272	
– Yes			496	1.81%	299	3.22%
– No			26,849	98.19%	8,973	96.78%
Ever diagnosed with dementia	Binary	1–13	27,345		9,272	
– Yes			786	2.87%	445	4.80%
– No			26,559	97.13%	8,827	95.20%
Ever reported back problems	Binary	1–13	27,345		9,272	
– Yes			15,939	58.29%	6,490	70.00%
– No			11,406	41.71%	2,782	30.00%
Whether health limits work	Binary	1–13	27,345		9,272	
– Yes			15,059	55.07%	6,091	65.69%
– No			12,286	44.93%	3,181	34.31%
Self-reported probability of having a work limiting health problem in the next 10 years	Continuous	1–6	9,984	40.2% (25.0%)	5,368	39.4% (23.9%)

Variable	Type	Wave	European HRS respondents		European HRS genotyped study sample	
			N	%, Mean (SD), or range	*N*	%, Mean (SD), or range
Panel C. Occupation, income, wealth, and retirement						
Labor force status	Categorical	1–13	26,580		9,272	
– Works full-time			8,215	30.91%	2,696	29.08%
– Works part-time			1,259	4.74%	368	3.97%
– Unemployed			295	1.11%	37	0.40%
– Not in labor force			2,084	7.84%	625	6.74%
– Retired			12,909	48.57%	4,971	53.61%
– Partly retired			1,249	4.70%	57	0.61%
– Disabled			569	2.14%	518	5.59%
Household income in 2016 prices (in thousands of US$)	Continuous	1–13	27,345	82.6 (83.6)	9,272	76.8 (76.0)
Total household wealth in 2016 prices (in thousands of US$)	Continuous	1–13	27,345	476.4 (1,299)	9,272	563.7 (957.0)
Received Social Security (OASDI) in any wave	Binary	1–13	27,345		9,272	
– Yes			20,005	73.16%	8,220	88.65%
– No			7,340	26.84%	1,052	11.35%
Self-reported probability of working full-time after age 65	Continuous	1–13	17,255	25.4% (31.9%)	7,195	22.6% (30.5%)
Plans to continue paid work in retirement	Binary	1	7,400		3,771	
– Yes			5,368	72.54%	2,761	73.22%
– No			2,032	27.46%	1,010	26.78%
Concerned about having enough retirement income	Categorical	1	6,275		3,331	
– 1. A lot			1,762	28.08%	868	26.06%
– 2. Somewhat			1,799	28.67%	982	29.48%
– 3. Little			1,518	24.19%	867	26.03%
– 4. Not at all			1,196	19.06%	614	18.43%
Planned retirement age (i.e., planned retirement year minus birth year)	Continuous	1–13	6,331	64.3 (5.3)	3,463	64.6 (5.3)
Satisfied with retirement	Categorical	1–13	15,868		7,484	
– 1. Very			8,680	54.70%	4,381	58.54%
– 2. Moderately			5,800	36.55%	2,664	35.60%
– 3. Not at all			1,388	8.75%	439	5.87%
Expectation of total retirement wealth in 2016 prices (in thousands of US$)	Continuous	1	6,346	201.9 (571.4)	3,354	214.4 (580.3)
Life-insurance coverage in any wave	Binary	1–13	27,212		9,272	
– Yes			22,050	81.03%	8,330	89.84%
– No			5,162	18.97%	942	11.31%
Percentage of waves covered by life insurance	Continuous	1–13	27,212	65.4% (40.0%)	9,272	68.9% (35.9%)
Long-term care insurance coverage in any wave	Binary	1–13	27,165		9,272	
– Yes			6,561	24.15%	3,107	33.51%
– No			20,604	75.85%	6,165	66.49%
Percentage of waves covered by long-term care insurance, 1992–2014	Continuous	1–13	27,165	10.4% (23.4%)	9,272	13.4% (25.9%)
Financial planning horizon	Categorical	1, 4–8, 11–13	21,905		9,196	
– 1. Next few months			3,245	14.81%	1,154	12.55%
– 2. Next year			2,433	11.11%	979	10.65%
– 3. Next few years			6,463	29.50%	2,862	31.12%
– 4. Next 5–10 years			7,364	33.62%	3,293	35.81%
– 5. Longer than 10 years			2,400	10.96%	908	9.87%

2.2 Statistical analyses

2.2.1 Quality control, heritability, and genetic correlations

We performed an extensive search of the published GWAS literature to identify studies of common medical conditions and mortality risks performed in at least one hundred thousand participants. We had preregistered that particular threshold to avoid constructing underpowered polygenic scores. Our selection of traits was guided by the academic literature on established predictors of mortality[77,105]. As reported in **Table 2**, the collected GWAS results span many medical domains, including cardiology, oncology, neurology, and psychiatry. Also, we collected results for the following recognized mortality risks: three measures of blood pressure[121], body mass index (BMI)[122], four measures of blood cholesterol[123], educational attainment[89,104,124], height[105,125], parental lifespan[79,106], smoking initiation and cigarettes per day (smoking intensity)[83], and alcoholic drinks per week[126,127]. The average *N* of the GWAS was ~455,000, and the largest included more than one million participants (atrial fibrillation)[128].

Next, we performed quality-control analyses to exclude rare and low-quality SNPs (minor allele frequency less than 0.01 and imputation quality less than 0.9), in accordance with the standards in the field[92,100,129]. It is recommended to exclude such SNPs to increase the signal-to-noise ratio in the polygenic scores[130]. We estimated SNP heritabilities (the proportion of variation explained by a set of SNPs) by applying LD Score regression on the GWAS results[131,132]. This method utilizes the fact that, under a polygenic model, genetics variants that are correlated—in *linkage disequilibrium* (LD)—with many other variants are more likely to capture genetic signal across the genome. A variant's LD Score – the sum of its squared correlations with other variants – has a proportional relationship with the expectation of its association test statistic. This relationship can be used to approximate a lower bound SNP heritability and to test for population stratification.

To eliminate GWAS with negligible genetic signal, we excluded two traits for which the LD Score heritability was not distinguishable from zero, at *P* less than 0.001: (1) large artery stroke (h^2 = 0.07%; P = 0.59) and (2) small vessel stroke (h^2 = 0.25%; P = 0.037). The heritability of the 27 remaining traits ranged from 0.7% (cardioembolic stroke) to 45.9% (height). Next, using the same method, we estimated pairwise genetic correlations (r_g)[132] (**Supplementary Table 1** and **Supplementary Figure 2**). LD Score regression estimates genetic correlations by utilizing the overlap in association test statistics across SNPs as a measure of genetic covariance, while adjusting for LD and sample overlap. It has been shown that this method can robustly estimate genetic correlations across a range of plausible confounding scenarios[133]. Importantly,

we found that most of the common medical conditions and mortality risks that we collected were moderately genetically correlated with parental lifespan, which suggests that they may be able to explain variation in survival[79,104,105,134].

2.2.2 Polygenic scores

In our study sample, we computed polygenic scores for 27 traits as linear combinations of individual-level genotypes weighed by trait-specific GWAS effects. Thus, polygenic scores aggregate an individual person's genetic risk (or propensity) towards a trait or disease into a genetic predictor[36,99,135]. The *i*th respondent's polygenic score for the *k*th trait, $\hat{S}_{ik}$, was computed as:

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

where the respondent's genotype g_{ij} at SNP j was weighed by the corresponding trait-specific GWAS effect, $\hat{\beta}_{jk}$, and then added up across M SNPs. In the regression analyses below, we entered multiple polygenic scores on the right-hand side of the regression equation, thereby increasing the predictive accuracy of genetically correlated outcomes[103].

Before computing the scores for each of the 27 traits, we first excluded weak associations at P value greater than 0.05, to reduce noise from estimation error[136]. We then constructed polygenic scores using the remaining SNPs that overlapped with the 1.4 million SNPs in the high-quality consensus genotype set established by the International HapMap 3 Consortium[137,138]. This is a common approach that achieves reasonable genome-wide coverage without including too many correlated SNPs, which could reduce the predictive accuracy[92,100]. The final number of SNPs in the scores ranged from 45,077 to 271,528 (**Table 2**).

2.2.3 Univariate survival analysis and Cox proportional-hazards regression

We first performed a series of non-parametric, univariate analyses of respondent, maternal, and paternal survival by estimating stratified Kaplan-Meier survival functions (**Supplementary Table 2**)[151,152]. A benefit of analyzing parental survival is reduced censoring and mortality selection, but the disadvantage is that the offspring genotypes are a noisy measure of the parents' genotype[153]. Also, mortality risks change over time, and it is questionable that the same genetic and environmental risks would be relevant now. We analyzed monthly survival in the respondents but only

Table 2. GWAS results for common medical conditions and mortality health risks

Traits	Domain	Relevance	GWAS *N*	LD Score h^2	SE (h^2)	*P* value (h^2)	λ GC	Mean χ^2	*M* SNPs in polygenic score	Reference
Alzheimer's disease	Neuropsychiatry	Medical condition	455,258	1.4%	0.20%	**0.000**	1.086	1.126	58,587	Jansen et al. (2019)[139]
Any ischemic stroke	Neurology	Medical condition	446,696	1.3%	0.13%	**0.000**	1.121	1.140	61,911	Malik et al. (2018)[140]
Any stroke	Cardiology/ Neurology	Medical condition	446,696	1.3%	0.14%	**0.000**	1.124	1.148	62,712	Malik et al. (2018)[140]
Atrial fibrillation	Cardiology	Medical condition	1,030,836	2.4%	0.21%	**0.000**	1.293	1.517	89,934	Nielsen et al. (2018)[128]
Body mass index (BMI)	Metabolic/ Lifestyle	Mortality risk	795,640	19.0%	0.54%	**0.000**	2.787	3.986	208,732	Yengo et al. (2018)[141]
Breast cancer	Oncology	Medical condition	228,951	13.3%	1.04%	**0.000**	1.362	1.683	95,858	Michailidou et al. (2017)[142]
Cardioembolic stroke	Cardiology	Medical condition	446,696	0.7%	0.12%	**0.000**	1.108	1.120	59,268	Malik et al. (2018)[140]
Chronic kidney disease	Urology	Medical condition	444,971	1.4%	0.17%	**0.000**	1.146	1.199	67,634	Wuttke et al. (2019)[143]
Cigarettes per day	Oncology/ Lifestyle	Mortality risk	109,804	9.5%	1.01%	**0.000**	1.184	1.216	68,207	the Neale lab (2017)[144]
Coronary artery disease	Cardiology	Medical condition	547,261	5.9%	0.32%	**0.000**	1.362	1.619	100,938	van der Harst et al. (2017)[145]
Depression	Psychiatry	Medical condition	500,199	6.0%	0.23%	**0.000**	1.453	1.604	107,705	Howard et al. (2019)[146]
Diastolic blood pressure	Cardiology	Mortality risk	757,601	13.0%	0.53%	**0.000**	2.047	3.191	186,290	Evangelou et al. (2018)[121]
Drinks per week	Metabolic/ Lifestyle	Mortality risk	414,343	6.9%	0.30%	**0.000**	1.407	1.584	106,587	Karlsson Linnér et al. (2019)[92]
Educational attainment	Behavior/ Lifestyle	Mortality risk	756,382	10.7%	0.27%	**0.000**	2.090	2.642	184,905	Lee et al. (2018)[100]
HDL cholesterol	Metabolic	Mortality risk	187,167	21.4%	2.85%	**0.000**	1.020	1.210	45,927	Willer et al. (2013)[91]
Height	Cross-domain	Mortality risk	709,706	45.9%	2.13%	**0.000**	3.613	9.042	271,528	Yengo et al. (2018)[141]
Insomnia	Psychiatry	Medical condition	386,533	4.6%	0.20%	**0.000**	1.310	1.367	84,329	Jansen et al. (2019)[147]
Large artery stroke	Cardiology	Medical condition	446,696	0.1%	0.13%	0.538	1.108	1.113	Excluded	Malik et al. (2018)[140]
LDL cholesterol	Metabolic	Mortality risk	173,082	20.3%	3.24%	**0.000**	1.014	1.194	45,388	Willer et al. (2013)[91]
Parental lifespan	Cross-domain	Mortality risk	640,189	2.4%	0.13%	**0.000**	1.300	1.342	82,676	Timmers et al. (2019)[79]
Prostate cancer	Oncology	Medical condition	140,254	14.8%	1.81%	**0.000**	1.217	1.460	60,248	Schumacher et al. (2018)[148]
Pulse pressure	Cardiology	Mortality risk	745,820	11.4%	0.40%	**0.000**	1.914	2.822	171,127	Evangelou et al. (2018)[121]
Schizophrenia	Psychiatry	Medical condition	105,318	42.0%	1.52%	**0.000**	1.691	1.977	134,764	Pardinas et al. (2018)[149]
Small vessel stroke	Cardiology	Medical condition	446,696	0.3%	0.12%	0.037	1.080	1.094	Excluded	Malik et al. (2018)[140]
Smoking initiation	Oncology/ Lifestyle	Mortality risk	518,633	8.9%	0.29%	**0.000**	1.637	1.929	136,392	Karlsson Linnér et al. (2019)[92]
Systolic blood pressure	Cardiology	Mortality risk	745,820	13.3%	0.49%	**0.000**	2.090	3.154	187,442	Evangelou et al. (2018)[121]
Total cholesterol	Metabolic	Mortality risk	187,365	21.2%	2.73%	**0.000**	1.014	1.237	47,713	Willer et al. (2013)[91]
Triglycerides	Metabolic	Mortality risk	177,861	21.5%	3.49%	**0.000**	1.002	1.219	45,077	Willer et al. (2013)[91]
Type 2 diabetes	Metabolic	Medical condition	231,426	19.4%	0.86%	**0.000**	1.626	1.958	130,042	Mahajan et al. (2018)[150]

observed yearly survival in the parents. All survival analyses were performed using analytical packages that are freely available for the R software environment[154–158].

While univariate analysis can suggest factors that appear to be associated with survival, such analysis is not statistically conditioned on other variables[159]. Therefore, we performed multiple regression of respondent survival by estimating a series of Cox proportional-hazards models (**Table 3** and **Supplementary Table 3**)[160,161]. We did not perform Cox regression of parental survival because we could not observe their covariates. In these analyses, we included ten standardized genetic PCs. Because of the HRS household structure, we clustered the standard errors at household level. We estimated four nested regression models hierarchically. In summary, the models included the following regressors, in addition to the genetic PCs:

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

Overall, we regard model (3) as our preferred model for the development of a prognostic index of genetic risk that could be evaluated early in life before any signs or symptoms of disease (see below). At the same time, model (4) is indicative of whether the polygenic scores can be used to stratify survival above and beyond the inclusion of intermediate variables that lie on the causal pathway between genetic risk and mortality, such as manifested medical conditions.

2.2.4 Model diagnostics and fit

We evaluated a series of model diagnostics for the Cox models[162–164]. First, we checked whether the models or any of the regressors violated the proportional-hazards (PH) assumption by testing the scaled Schoenfeld residuals for association with the time to event. Second, we examined whether we had chosen a suitable functional form for the covariates by visually inspecting their relationship with the Martingale residuals. Third, we examined the Deviance residuals for outliers or influential observations.

We then assessed the model fit. We computed likelihood-ratio tests, Wald tests, and (score) log-rank tests to evaluate whether the regressors improved model fit above the null model (with only the baseline hazard). Next, we computed the

Cox-Snell pseudo-R^2, the Harrel's c statistic[165], and the Gönen & Heller's K statistic[165]. The latter two are concordance measures that compare the observed time to event with the ranks of the respondents' hazards as predicted by the fitted model, akin to an area under the ROC curve (AUC) measure[161]. In the next section, we explain how we computed the Royston & Sauerbrei R_D^2 measure of model fit[161,163]. Finally, we performed likelihood-ratio tests between the nested models to examine whether additional regressors improved model fit. Because of the large number of events, $N_{deceased}$ = 2,332, no stepwise covariate selection was needed, and we confirmed that there was no near-perfect collinearity between any of the regressors.

2.2.5 Prognostic indices of survival

To investigate how well the polygenic scores, when combined, could stratify survival relative to (i) the genetic PCs and to (ii) the covariates, we computed three prognostic indices (PI) for each of the four Cox models[159,161,163]. In this context, a PI is a weighted sum of multiple variables, weighed by their Cox coefficients. In other words, we aggregated the influence of sets of variables into hazard indices. The *i*th respondent's PI across the polygenic scores was computed as:

$$\widehat{PI}_{PGS,i} = \sum_{k=1}^{K} \hat{\alpha}_k \hat{S}_{ik} \quad (2)$$

where the respondent's polygenic score, $\hat{S}_{ik}$, was weighed by its Cox regression coefficient, $\hat{a}_k$, and then added up across the K polygenic scores. The PIs for the genetic PCs ($\widehat{PI}_{PC}$) and the covariates ($\widehat{PI}_{COVAR}$) were computed analogously, except that $\widehat{PI}_{COVAR}$ excluded the sex-specific birth-year dummies and the birth-month dummies, because we consider these to capture time and sampling effects rather than meaningful individual differences.

To evaluate the relative influence of the three PIs: $\widehat{PI}_{PGS}$, $\widehat{PI}_{PC}$, and $\widehat{PI}_{COVAR}$, we computed the Royston & Sauerbrei R_D^2 measure of model fit[161,163,166]. This method orders the respondents according to a PI and then projects them onto a normal distribution to attain so-called rankits (expected Z-scores based on the order and number of the individuals). Then, an auxiliary Cox regression is performed on the rankits alone, and the resulting regression coefficient can be transformed into a measure of explained variation on the log hazard scale. Thus, the R_D^2 is more similar to the traditional coefficient of determination (R^2) of linear regression than to the Cox-Snell pseudo-R^2, which is computed as ratios of log-likelihoods. We performed this comparison across all four Cox models (**Supplementary Table 3**).

2.2.6 Comparison of stratified survival functions

Next, for Cox models (3) and (4), we compared stratified survival functions in two ways: the first comparison (a) was stratified by the top versus lower nine deciles of the PI distribution, and the second comparison (b) by the top versus bottom decile. Comparison (a) was deliberately chosen to mirror medical underwriting, where primarily individuals with a substantially increased risk are classified as substandard and charged a higher premium[31,62], while comparison (b) is similar to a traditional extreme-groups approach[79,167]. We report both comparisons but focus our discussion on comparison (a). In these analyses, the $\widehat{PI}_{PGS}$ based on model (3) is our preferred genetic predictor of interest (**Table 4** and **Figure 1**). We performed a log-rank test to see whether the survival functions (not the median) of the two groups in each comparison differed, and we evaluated the size of the difference in median lifespan[152]. Thereafter, we stratified survival by the three PIs simultaneously, but only for comparison (a) since simultaneously stratifying by the top and bottom decile of all three PIs, which are only weakly correlated, would not include enough respondents (**Tables 5** and **Figures 2**).

2.2.7 Benchmark to conventional actuarial risk factors

We then stratified Kaplan-Meier survival functions by each of the following separate conventional actuarial risk factors, that were selected as benchmarks to the $\widehat{PI}_{PGS}$ (**Supplementary Table 4** and **Supplementary Figure 6**): smoking (never, current, and former); BMI; years of schooling; log of household income; sex; and ever diagnosed with (a) high blood pressure (or hypertension), (b) diabetes (or high blood sugar), (c) cancer (or malignant tumor except skin cancer), (d) chronic lung disease (except asthma such as chronic bronchitis or emphysema), (e) heart conditions (such as heart attack, coronary heart disease, etc.), or (f) stroke (or transient ischemic attack). To avoid mortality selection from the genotyping procedure, we applied this analysis to both our study sample as well as to the full sample of 27,345 European HRS respondents.

2.2.8 Cross validation of the preferred Cox model

To maximize sample size, we applied the prognostic index analyses to the same sample that was used to estimate the Cox coefficients. To evaluate the possibility of overfitting, we performed a cross validation with 1,000 iterations with our preferred model (3)[166]. In each iteration, we trained the coefficients in a random sample that contained 65% of the households ($H_{training}$ = 4,325) and used the remaining households as a validation sample ($H_{validation}$ = 2,329). First, we evaluated the coefficient

of the $\widehat{PI}_{PGS}$ estimated in an auxiliary Cox regression—the so-called "calibration slope"[161]. If the slope is different from one, then the cross-validation discrimination is either better (>1) or worse (<1). To test this hypothesis, we evaluated the median $Z = \frac{\hat{\alpha}_{PI.PGS}-1}{SE_{robust}}$ across the iterations. Secondly, we calculated the median R_D^2 of the $\widehat{PI}_{PGS}$. Finally, we evaluated the median of the difference in median lifespan analogous to comparisons (a) and (b), described in **section 2.2.6**.

2.2.9 Subjective life expectancy, self-rated health, and economic outcomes

We tested whether the $\widehat{PI}_{PGS}$ based on Cox model (3) was associated with subjective life expectancy and self-rated health (**Supplementary Table 5**). Subjective life expectancy is defined as the ratio between the respondents' self-reported probability of surviving to a specific age, divided by the life table prediction adjusted for age and sex. Because subjective life expectancy is normally distributed but left censored on 0, we performed both OLS and Tobit regressions on that outcome. Self-rated health is measured on a five-point Likert scale with the following categories: 1. Excellent; 2. Very good; 3. Good; 4. Fair; and 5. Poor, so we estimated an ordinal logit regression. Both models controlled for the same covariates as our preferred model (3), and we clustered the standard errors at household level.

We also tested the $\widehat{PI}_{PGS}$ based on Cox model (3) for association with the following eleven preregistered variables of relevance to economic research (**Supplementary Table 6**)[d]: (a) whether health limits work; (b) the self-reported probability of having a work-limiting health problem in the next ten years; (c) the self-reported probability of working full-time after age 65; (d) plans to continue paid work in retirement; (e) concerned about having enough retirement income; (f) planned retirement age (defined as planned retirement year minus birth year); (g) retirement satisfaction; (h) expectation of total retirement wealth; (i) the percentage of waves covered by private life insurance or (j) by long-term care insurance; and finally (k) financial planning horizon. Depending on the distribution of each outcome, we applied either OLS, logit, or ordinal logit regression, again with model (3) covariates and clustered standard errors.

d We had also preregistered that we would study private health-insurance coverage. But in the HRS, there is little variation in health insurance coverage, so we dropped this particular analysis out of power considerations.

3. Results

3.1 Results from the univariate survival analysis

For each of the 27 polygenic score that we constructed, we performed univariate survival analyses of respondent, maternal, and paternal survival. The results are reported in **Supplementary Table 2** and **Supplementary Figures 3–5.** When stratifying survival by the top versus lower nine deciles of the score distribution, we found that 18 polygenic scores significantly discriminated survival in either the respondents or their parents (at $P < 0.05$) and that eight were Bonferroni-significant corrected for 27 traits. With respect to respondent survival, the following polygenic scores had a strong and significant influence (defined here as >1 year difference in median survival): (a) Alzheimer's disease (1.2 y), (b) any ischemic stroke (1.3 y), (c) any stroke (1.6 y), (d) BMI (1.3 y), (d) cigarettes per day (2.2 y), (e) educational attainment (1.3 y), (f) prostate cancer (1.2 y), and (g) type 2 diabetes (1.6 y). As to the parents, the score for parental lifespan had the strongest influence (4 and 6 y in the mothers and fathers, respectively, but only 0.6 y in the respondents).

All the univariate associations with respondent and parental survival were in the expected direction, except the association between respondent survival and the score for prostate cancer (this score was not associated with parental survival). Therefore, we performed an ad-hoc robustness check to ensure that it associated with the likelihood of reporting a cancer diagnosis, which it indeed did in the expected direction ($P = 0.0007$). Considered on its own, prostate cancer has an overall high survival rate as long as it is detected before metastasizing[168]. We speculate that the genetic risk of prostate cancer may lead to certain healthcare benefits if diagnosed early, such as more frequent doctor checkups or changes in lifestyle. This may explain the unexpected direction of effect, but replication is necessary to ascertain this.

3.2 Results from the Cox proportional-hazards regression

A selection of the multivariate survival analysis results is presented in **Table 3** and the complete results in **Supplementary Table 3**. In our preferred model (3), we identified associations with the polygenic scores for Alzheimer's disease ($\hat{\alpha} = 0.052$; $P = 0.022$), atrial fibrillation ($\hat{\alpha} = 0.054$; $P = 0.019$), cigarettes per day ($\hat{\alpha} = 0.073$; $P = 0.001$), height ($\hat{\alpha} = 0.049$; $P = 0.046$), type 2 diabetes ($\hat{\alpha} = 0.054$; $P = 0.036$), and parental lifespan ($\hat{\alpha} = -0.087$; $P < 0.001$). All estimated effects were in the anticipated direction. Importantly, we could not detect any violation of the proportional-hazards (PH) assumption (P of the global $\chi^2 = 0.404$), and the model attained satisfying fit (e.g., a Cox and Snell R^2 of 0.23). The R^2_D of $\widehat{PI}_{PGS}$ was 0.039 (95% CI = 0.028–0.052), while the R^2_D of the $\widehat{PI}_{PC}$ and $\widehat{PI}_{COVAR}$ were 0.006 (0.002–0.011) and 0.113 (0.095–0.133),

Table 3. Selection of hierarchical Cox regression results[1]

Regressors (polygenic scores shown in bold)	Model 3 (N = 9,246; deaths = 2,321) Coefficient (Robust SE)	Model 4 (N = 9,007; deaths = 2,223) Coefficient (Robust SE)
Alzheimer's disease	0.052 (0.023)*	0.057 (0.024)*
Atrial fibrillation	0.054 (0.023)*	0.046 (0.025)†
Cigarettes per day (smoking intensity)	0.073 (0.023)***	0.064 (0.024)**
Height	0.049 (0.025)*	0.045 (0.026)
Type 2 diabetes	0.054 (0.026)*	0.029 (0.028)
Parental lifespan	−0.087 (0.025)***	−0.058 (0.028)*
Sex (0 = Male; 1 = Female)	−0.872 (0.241)***	−0.568 (0.302)†
Census region (base: Northeast)		
– Midwest	−0.107 (0.072)	−0.035 (0.075)
– South	0.048 (0.068)	0.034 (0.072)
– West	−0.063 (0.080)	0.023 (0.083)
Years of schooling	−0.058 (0.010)***	−0.024 (0.010)*
Log of total household income (2016 prices)	−0.201 (0.046)***	−0.089 (0.047)†
Labor force participation (base: Works full-time)		
– Works part-time	−0.167 (0.169)	−0.066 (0.165)
– Unemployed	0.276 (0.532)	0.243 (0.552)
– Partly retired	−0.335 (0.111)**	−0.355 (0.118)**
– Retired	−0.251 (0.073)***	−0.385 (0.076)***
– Disabled	1.119 (0.270)***	0.725 (0.314)*
– Not in labor force	−0.148 (0.117)	−0.198 (0.118)†
Received "Social Security" (OASDI) in any wave	−0.430 (0.117)***	−0.496 (0.124)***
Married or partnered in any wave	−0.200 (0.071)**	−0.158 (0.072)*
Widowed in any wave	−0.357 (0.056)***	−0.348 (0.057)***
Sex-specific birth-year dummies	YES	YES
Birth-month dummies	YES	YES
Subjective life expectancy		−0.086 (0.029)**
Self-rated health (Base: 1. Excellent)		
– 2. Very good		−0.050 (0.089)
– 3. Good		0.304 (0.091)***
– 4. Fair		0.633 (0.102)***
– 5. Poor		1.069 (0.134)***
BMI		0.013 (0.006)*
Alcoholic drinks per week		0.012 (0.005)*
Current smoker		0.812 (0.071)***
Ever smoker		0.193 (0.053)***
Maternal max attained age		−0.003 (0.002)†
Paternal max attained age		−0.003 (0.002)*
Ever diagnosed with high blood pressure		−0.022 (0.054)
Ever diagnosed with diabetes		0.136 (0.054)*
Ever diagnosed with cancer		0.244 (0.049)***
Ever diagnosed with chronic lung disease		0.348 (0.058)***
Ever diagnosed with heart conditions		0.069 (0.049)
Ever diagnosed with stroke		0.091 (0.058)
Ever diagnosed with psychiatric disorder		0.082 (0.105)
Ever diagnosed with arthritis		−0.165 (0.056)**
Ever diagnosed with Alzheimer's disease		0.208 (0.090)*
Ever diagnosed with dementia		−0.180 (0.082)*
Ever diagnosed with back problems		−0.356 (0.053)***
Global test if model violates PH assumption	χ^2 (128 df) = 131.0; P = 0.404	χ^2 (150 df) = 203.0; P < 0.001
Cox-Snell R^2	0.231	0.286
Harrel's c statistic	0.816 (0.004)	0.850 (0.004)
Gönen & Heller's K statistic	0.857 (0.003)	0.862 (0.003)

*Notes: † $P \leq 0.1$; * $P \leq 0.05$: ** $P \leq 0.01$; *** $P \leq 0.001$.* Robust standard errors were clustered at the household level.

1 The complete results are reported in **Supplementary Table 3**.

respectively. Thus, the polygenic scores explained much more of the variation on the log hazard scale than the genetic PCs, but only about a third compared to the other model covariates.

The only difference between model (2) and our preferred model (3) is the score for parental lifespan. We had decided to add that particular score separately because it may capture the influence of other scores, as suggested by the genetic correlations (**Supplementary Table 1** and **Supplementary Figure 2**). However, adding that score did not notably influence the parameter estimates of the other scores, and the set of significant scores was the same between models (2) and (3). Notably, the effect of the score for parental lifespan in model (3) was the largest effect estimated for any of the polygenic scores across all four models ($\hat{\alpha}$= −0.087; P = 0.0004).

Model (4) is our most adjusted model. This additionally included subjective life expectancy and self-rated health, various lifestyle factors (such as smoking and drinking), observed parental lifespan, and eleven indicators for categories of medical diagnoses, such as "ever diagnosed with cancer (or malignant tumor except skin cancer)". Unfortunately, the medical diagnoses were only available as binary indicators. Nevertheless, our primary interest is in model (3), which could be evaluated early in life. As could be expected when including directly observed health variables, model (4) drastically improved the model fit over both the null model and model (3) (both $P \sim 0$). However, model (4) also strongly violated the PH assumption (P of the global χ^2= 0.00003). Reassuringly, the estimates of the polygenic scores, which we were particularly interested in, appeared highly stable across all four model specifications.

3.3 Results from the prognostic index analysis

Using the Cox coefficients, we computed prognostic indices (PIs) for three sets of regressors: (i) the polygenic scores ($\widehat{PI}_{PGS}$), (ii) the genetic PCs ($\widehat{PI}_{PC}$), and (iii) the covariates ($\widehat{PI}_{COVAR}$). We first evaluated the relative influence of the three PIs within each model, using the Royston & Sauerbrei R^2_D (**Supplementary Table 3**). Across the four models, the proportion of variation explained on the log hazard scale by the $\widehat{PI}_{PGS}$ was between 0.03 and 0.041. The R^2_D of the $\widehat{PI}_{PGS}$ was stable across the four models, which suggests that the polygenic scores explained a non-negligible proportion of variation even when adjusted for socioeconomic variables, observable health risks, and other potential confounders.

Next, we used the PIs based on models (3) and (4) to stratify Kaplan-Meier survival functions, first using only the $\widehat{PI}_{PGS}$ (**Table 4** and **Figure 1**). The $\widehat{PI}_{PGS}$ based on model (3) could distinguish a 3.5-year difference in median lifespan in the first comparison (a) between the top decile (N = 927) and the lower nine deciles (N = 8,345), and a

4.4-year difference in the second comparison (b) between the top and the bottom decile (N = 928). As expected, the $\widehat{PI}_{PGS}$ based on model (4) could distinguish somewhat less: 2.9 and 4.1 years, in comparisons (a) and (b), respectively. The P values of the tests of no difference between the groups were all below 7.91×10^{-13}. Thus, the $\widehat{PI}_{PGS}$ could distinguish a greater difference in median lifespan than any of the scores could on their own in the univariate analysis, even when based on the most extensively adjusted model (4).

Thereafter, we stratified survival using the three PIs simultaneously (**Table 5** and **Figure 2**). Here we performed only the first comparison (a) (see **Methods**). With respect to the PIs based on model (3), respondents' median lifespan was 3.1 years shorter in the top decile of $\widehat{PI}_{PGS}$ (N = 736) than in the bottom nine deciles (N = 6,674) and was similar to the $\widehat{PI}_{COVAR}$ (3.9 y; N = 723). The few individuals who were in the top decile of both the $\widehat{PI}_{PGS}$ and the $\widehat{PI}_{COVAR}$ (N = 97) had an 8.1 years shorter median lifespan. The log-rank test was highly significant ($P = 2.78\times10^{-75}$).

As expected, in the analogous analysis based on model (4), the capacity of the $\widehat{PI}_{PGS}$ was somewhat reduced. Compared to the lower nine deciles of all three PIs (N = 6,605), the median lifespan was 2.4 years shorter in the top decile of the $\widehat{PI}_{PGS}$ (N = 694), 10.7 years shorter for the $\widehat{PI}_{COVAR}$ (N = 696), and an astonishing 14.8 years shorter for respondents in the top decile of both the $\widehat{PI}_{PGS}$ and the $\widehat{PI}_{COVAR}$ (N = 110). Respondents in the top decile of all three PIs had an 18.5 years shorter median lifespan, but we caution that the relevant cell is tiny (N = 11). Again, the log-rank test was highly significant ($P = 3.59\times10^{-271}$). Overall, adding health variables and medical conditions drastically increased the discriminatory capacity of the $\widehat{PI}_{COVAR}$, while it only slightly reduced the capacity of the $\widehat{PI}_{PGS}$, suggesting that the polygenic scores were able to add information above and beyond the inclusion of these intermediate variables.

3.3.1. Benchmark to conventional actuarial risk factors

We benchmarked how well the $\widehat{PI}_{PGS}$ could distinguish median lifespan in comparison to conventional actuarial risk factors (**Supplementary Table 4** and **Supplementary Figure 6**). In that capacity, the $\widehat{PI}_{PGS}$ was comparable to sex (2.8 and 3.2 years in our study sample and the full sample of European HRS respondents, respectively), former smoker (2.5 and 3.4 years), and ever diagnosed with diabetes or high blood sugar (1.7 and 3.6 years). The $\widehat{PI}_{PGS}$ distinguished a greater difference than several of the conventional risk factors, including the top decile of years of schooling (corresponding to more than 16 years of schooling; 1.3 and 2 years), ever diagnosed with cancer (or malignant tumor except skin cancer; 1.2 and 1.7 years), and ever diagnosed with

Table 4. Kaplan–Meier survival estimates stratified by a prognostic index (PI)

Panel A. Top versus lower nine deciles (Cox model 3 coefficients)						
Respondents in PI_{PGS}	*N*	Events	Median (years)	95% lower CL	95% upper CL	Log-rank *P* (Ho: no difference)
Lower nine deciles	8345	2025	88.5	88.1	88.8	**0.000**
Top decile	927	307	85.0	84.0	86.2	
Panel B. Top versus bottom decile (Cox model 3 coefficients)						
Respondents in PI_{PGS}	*N*	Events	Median (years)	95% lower CL	95% upper CL	Log-rank *P* (Ho: no difference)
Bottom decile	928	213	89.4	88.4	90.7	**0.000**
Top decile	927	307	85.0	84.0	86.2	
Panel C. Top versus lower nine deciles (Cox model 4 coefficients)						
Respondents in PI_{PGS}	*N*	Events	Median (years)	95% lower CL	95% upper CL	Log-rank *P* (Ho: no difference)
Lower nine deciles	8345	2039	88.3	88.0	88.7	**0.000**
Top decile	927	293	85.4	84.8	86.5	
Panel D. Top versus bottom decile (Cox model 4 coefficients)						
Respondents in PI_{PGS}	*N*	Events	Median (years)	95% lower CL	95% upper CL	Log-rank *P* (Ho: no difference)
Bottom decile	928	202	89.5	88.5	90.8	**0.000**
Top decile	927	293	85.4	84.8	86.5	

heart conditions (such as heart attack, coronary heart disease, angina, congestive heart failure, or other heart problems; 0.8 and 0.7 years). On the other hand, the $\widehat{PI}_{PGS}$ distinguished a smaller difference than the top decile of BMI (corresponding to BMI > 38.6; 4.4 and 5.3 years), current smoker (9.9 and 11.4 years), and ever diagnosed with chronic lung disease (except asthma such as chronic bronchitis or emphysema; 4.3 and 4.3 years). Notably, these comparisons show that the ability of polygenic scores to classify individuals into groups of different mortality risks is similar to or better than that of several conventional actuarial risk factors when the polygenic scores are combined into our preferred genetic predictor, with the major difference that the prognostic index can be evaluated at a young age.

3.3.2 Results of the cross validation

We performed a cross validation of model (3) to examine whether evaluating the prognostic indices in the full study sample might have introduced overfitting. Across 1,000 iterations, the median *Z* statistic of the test of the calibration slope was only marginally significant (P = 0.045). Similarly, the median R_D^2 of the $\widehat{PI}_{PGS}$ was 0.023 (instead of 0.036) across the iterations. The difference in median lifespan for comparisons (a) and (b) attenuated from 3.5 to 2.7 years and from 4.4 to 3.4 years, respectively. These attenuated differences in median lifespan fall within the confidence intervals of the main estimates and were still strongly significant. Thus, the cross validation suggested that the ability to distinguish lifespan was slightly overestimated

Table 5. Kaplan–Meier survival estimates stratified by the prognostic indices (PIs)

Panel A. Top versus lower nine deciles of the PIs (computed with model 3 coefficients)								
Respondents in PI_{PGS}	Respondents in PI_{PC}	Respondents in PI_{COVAR}	*N*	Events	Median (years)	95% lower CL	95% upper CL	Log-rank *P* (Ho: no difference)
Lower nine deciles	Lower nine deciles	Lower nine deciles	6764	1571	88.9	88.6	89.3	**0.000**
Top decile	Lower nine deciles	Lower nine deciles	736	236	85.8	84.6	86.7	
Top decile	Top decile	Lower nine deciles	82	29	84.2	81.2	89.5	
Top decile	Top decile	Top decile	10	3	80.8	59.4	--	
Lower nine deciles	Top decile	Lower nine deciles	740	204	88.0	86.7	88.6	
Lower nine deciles	Top decile	Top decile	94	28	79.0	77.7	85.2	
Lower nine deciles	Lower nine deciles	Top decile	723	211	85.0	83.2	86.2	
Top decile	Lower nine deciles	Top decile	97	39	80.8	74.0	82.4	
Panel B. Top versus lower nine deciles of the PIs (computed with model 4 coefficients)								
Respondents in PI_{PGS}	Respondents in PI_{PC}	Respondents in PI_{COVAR}	*N*	Events	Median (years)	95% lower CL	95% upper CL	Log-rank *P* (Ho: no difference)
Lower nine deciles	Lower nine deciles	Lower nine deciles	6605	1419	89.2	88.8	89.6	**0.000**
Top decile	Lower nine deciles	Lower nine deciles	694	195	86.8	85.4	87.8	
Top decile	Top decile	Lower nine deciles	86	28	85.9	83.9	89.5	
Top decile	Top decile	Top decile	11	6	70.7	63.6	--	
Lower nine deciles	Top decile	Lower nine deciles	722	170	88.8	87.6	89.6	
Lower nine deciles	Top decile	Top decile	83	45	76.3	75.8	78.7	
Lower nine deciles	Lower nine deciles	Top decile	696	309	78.5	77.3	79.8	
Top decile	Lower nine deciles	Top decile	110	51	74.4	72.7	76.4	

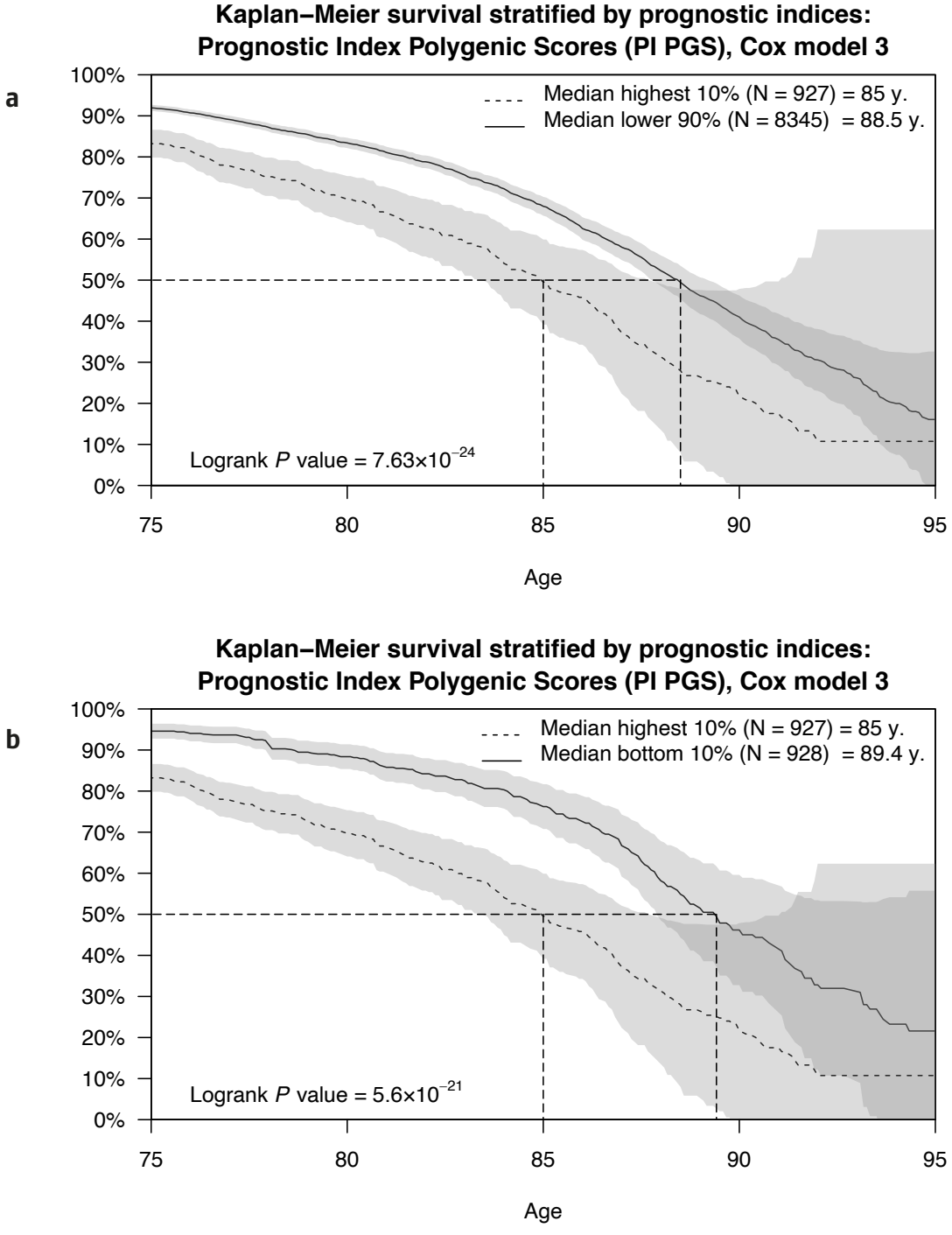
when evaluated in the full sample, but, reassuringly, our conclusions are still supported by these results.

3.4 Results from the analysis of subjective life expectancy, self-rated health, and economic outcomes

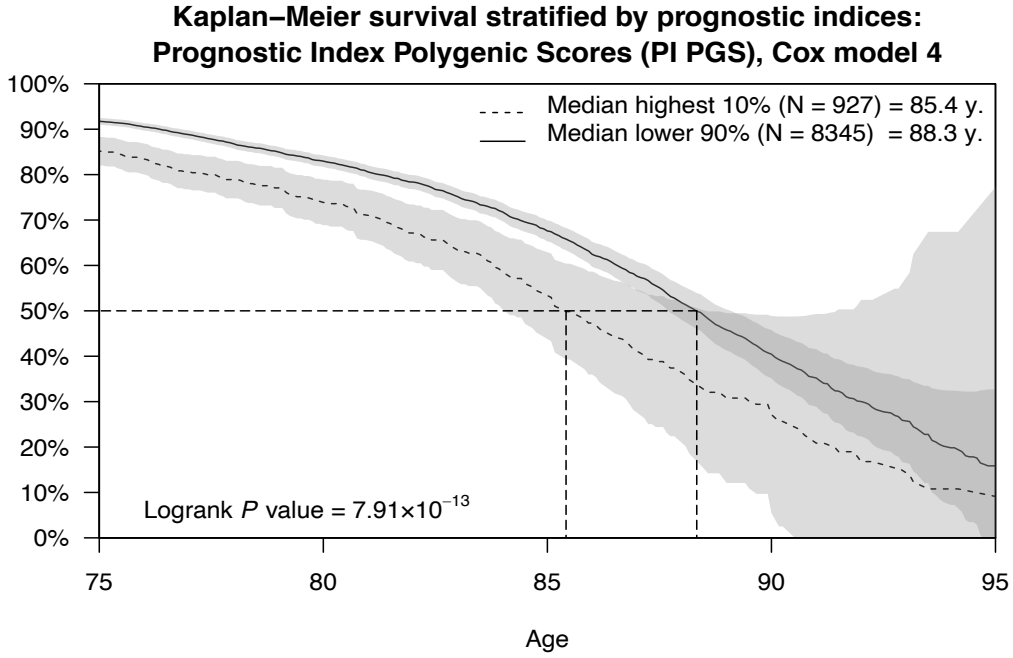
We investigated whether the polygenic scores were associated with subjective life expectancy and self-rated health. The results are reported in **Supplementary Table 5**. Our genetic predictor of interest was the $\widehat{PI}_{PGS}$ based on our preferred model (3), which had been standardized. For subjective life expectancy, the OLS and Tobit estimates were virtually identical, so we discuss only the Tobit results. The coefficient of the $\widehat{PI}_{PGS}$ was estimated to be −0.052 (*SE* = 0.008; $P = 8.03\times10^{-11}$). As expected, greater genetic risk was associated with reporting an expectation of shorter lifespan. However, the effect was small compared to, for example, being female (−2.047; *SE* = 0.523), living in the western part of the United States (0.143; *SE* = 0.030), or being disabled (−0.298; *SE* = 0.101). Interestingly, our results suggest that unobserved genetic risks had indeed been observed and captured by this health measure.

Next, we performed an ordinal logit regression of self-rated health (a higher value represents poorer health). In alignment with subjective life expectancy, we found that

Figure 1. Kaplan-Meier survival curves stratified by the prognostic index of the polygenic scores. *Using the prognostic index of the polygenic scores ($\widehat{PI}_{PGS}$) computed with the coefficient estimates of Cox models (3) and (4), we performed two comparisons of stratified survival functions. The first comparison (panels* **a** *and* **c***) was stratified by the top versus lower nine deciles of the PI distribution, while the second comparison (panels* **b** *and* **d***) was stratified by the top versus bottom decile. The dashed lines display the median survival in the two strata. The log-rank P value indicates whether the survival functions (not the median) of the two strata are different.*



c


Kaplan–Meier survival stratified by prognostic indices:
Prognostic Index Polygenic Scores (PI PGS), Cox model 4
Median highest 10% (N = 927) = 85.4 y.
Median lower 90% (N = 8345) = 88.3 y.
Logrank P value = 7.91×10−13
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
75
80
85
90
95
Age

d

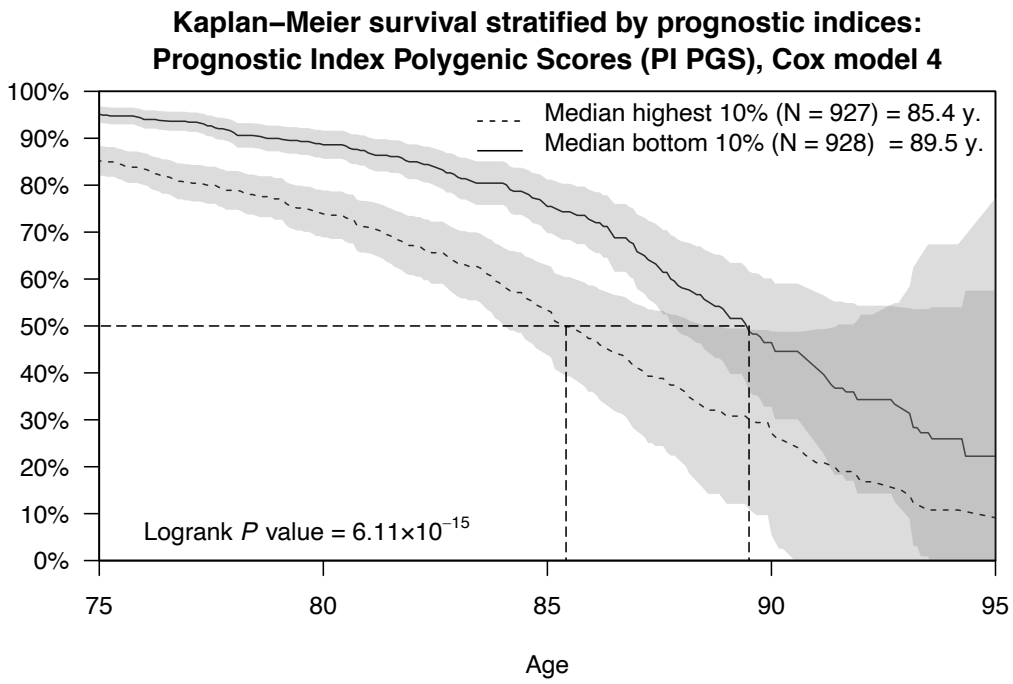
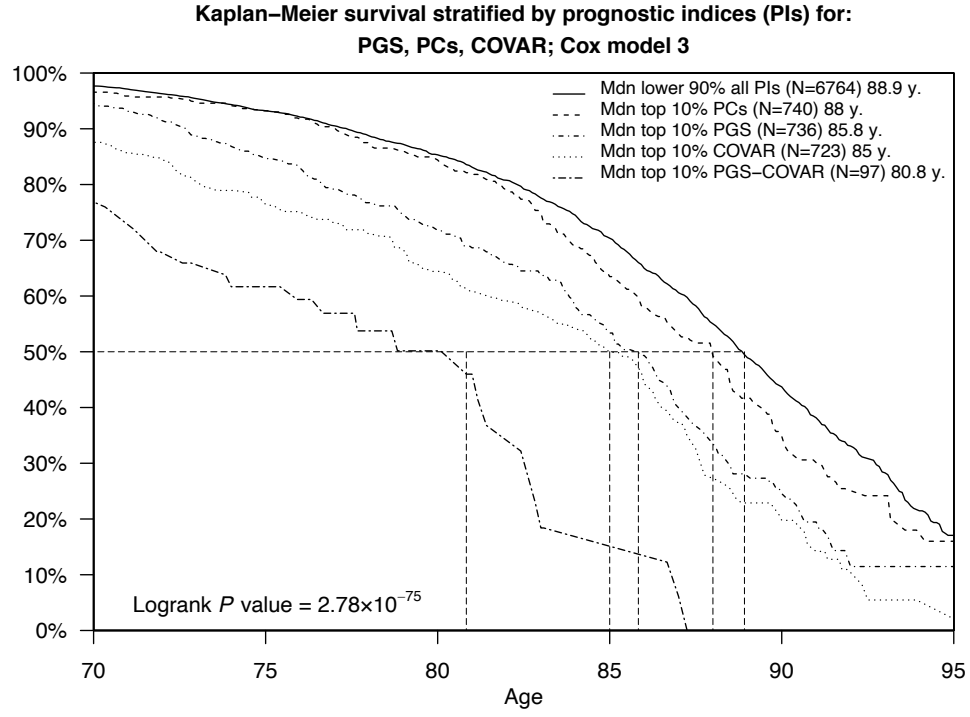
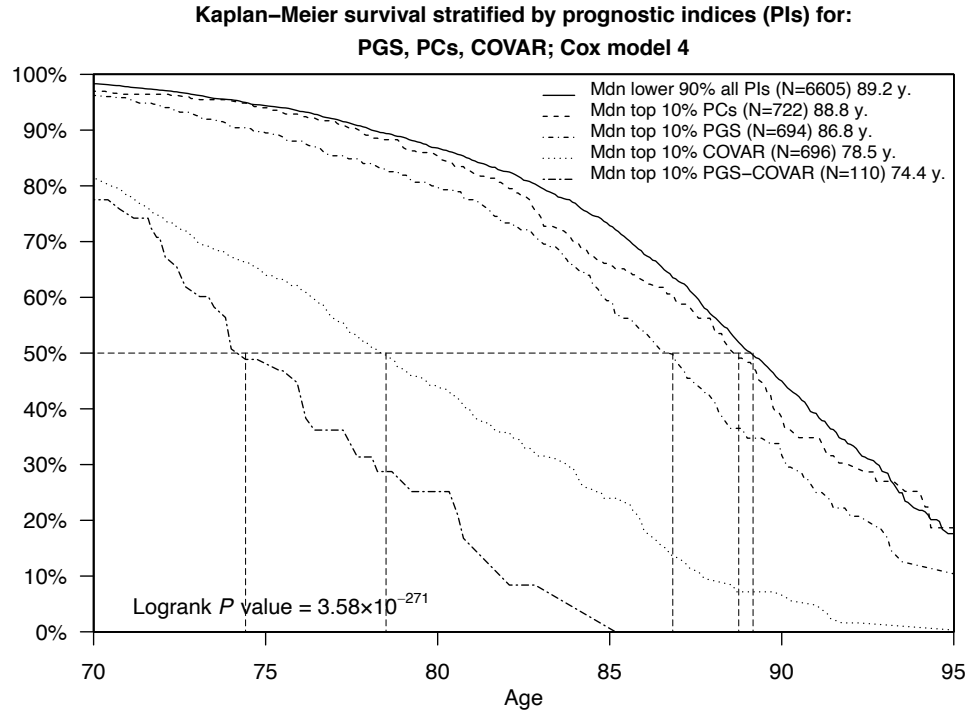

Kaplan–Meier survival stratified by prognostic indices:
Prognostic Index Polygenic Scores (PI PGS), Cox model 4
Median highest 10% (N = 927) = 85.4 y.
Median bottom 10% (N = 928) = 89.5 y.
Logrank P value = 6.11×10−15
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
75
80
85
90
95
Age

Figure 2 Kaplan-Meier survival curves stratified by the three prognostic indices. *Using the prognostic indices of the polygenic scores ($\widehat{PI}_{PGS}$), the genetic PCs ($\widehat{PI}_{PC}$), and the covariates ($\widehat{PI}_{COVAR}$), computed with the coefficient estimates of Cox models 3 (panel* **a***) and 4 (panel* **b***), we compared the simultaneously stratified survival functions by the top versus lower nine deciles of the distribution of the three PIs. The dashed lines display the median survival in the strata. The log-rank P value indicates whether the survival functions (not the median) of the strata are different. Three strata are not displayed in the figure but instead in* **Table 5**.

a



b



greater genetic risk was associated with reporting poorer health ($\hat{\beta}$ = 0.195; *SE* = 0.019; *P* = 1.03×10^{-24}). Thus, assuming a proportional influence across the response categories, the odds ratio for being in a higher category was 1.215 per standard deviation of the $\widehat{PI}_{PGS}$. The effect was similar to that of being female (OR = 1.213), but much smaller than being disabled (OR = 8.315). Overall, our results indicate that the genetic risks had indeed manifested and been observed via the respondent's health.

Finally, we investigated whether the polygenic scores were associated with eleven economic variables. The results are reported in **Supplementary Table 6**. Our genetic predictor of interest, the $\widehat{PI}_{PGS}$ based on model (3), was associated at marginal significance with five of the outcomes (*P* < 0.05) and Bonferroni-significant with three (*P* < 0.0045), in the expected direction. That is, increased genetic risk was associated with greater probability of having or expecting a work-limiting health problem (OR = 1.202; *P* = 1.08×10^{-13} and $\hat{\beta}$ = 1.079; *P* = 8.66×10^{-4}, respectively), less retirement satisfaction[e] (OR = 1.132; *P* = 1.08×10^{-6}), fewer waves of long-term care insurance coverage ($\hat{\beta}$ = −0.007; *P* = 0.0068), and shorter financial planning horizon[f] (OR = 0.953; *P* = 0.016). However, the estimates were relatively small. For example, per standard deviation, the $\widehat{PI}_{PGS}$ was associated with reporting a one percentage point greater probability of having a work limiting health problem in the next ten years and with two months shorter long-term care insurance coverage. Nonetheless, the effects were comparable with those of, for example, years of schooling, which was also associated with these five outcomes. Across the five outcomes, the difference compared to years of schooling was the greatest for long-term care insurance, where a standard deviation increase in years of schooling was associated with ~10 months longer coverage. We did not find evidence that greater genetic risk was associated with life insurance. Overall, our results show that polygenic scores for common diseases and mortality risks are indeed associated with economic outcomes, including insurance coverage.

e Retirement satisfaction was coded as "1. very; 2. moderately; 3. not at all".

f Financial planning horizon was coded as "1. next few months; 2. next year; 3. next few years; 4. next 5–10 years; 5. longer than 10 years".

4. Discussion

We investigated how well a broad set of polygenic scores for common medical conditions and mortality risks could distinguish differences in lifespan, and we benchmarked their performance in this regard to conventional actuarial risk factors. Our main finding is that polygenic scores have a joint capacity to classify people into groups of different mortality risk that is non-negligible, and this capacity is comparable to that of some conventional risk factors, including sex, former smoking, and years of schooling. Moreover, we found that the classification was even stronger than that of some of the diagnosed medical conditions that have been ascertained in the HRS, including ever being diagnosed with cancer. With this comparison, our results extend the literature on this topic, also by adjusting more extensively for conventional risk factors in multiple regression. Importantly, the polygenic scores were able to explain a non-trivial part of the variation in survival even when statistically adjusted for a range of confounders. We emphasize that our results represent only a lower bound of the predictive accuracy that polygenic scores will reach once larger GWAS studies become available[134,169].

4.1.1 Implications for insurance markets

Given these results, it is reasonable to expect that commercial interest in offering genetic tests for disease, mortality risks, and longevity to consumers is bound to increase further. Similarly, our results imply that polygenic scores already contain information that could be valuable to the insurance industry (e.g., in underwriting) if insurers were able to obtain genetic data from applicants and customers. Depending on whether genetic health information will remain legally or voluntarily exempt from underwriting, adverse selection will be more or less likely to occur as more people acquire knowledge of their genetic risks.

We did not find that our preferred genetic predictor was associated with life-insurance coverage. That is, we found no evidence that individuals with increased *unobserved* genetic risk would have already self-selected into purchasing life insurance. This finding aligns with previous studies that have found little evidence of adverse selection in that market[61,170,171]. Nonetheless, since the genetic risks were assumed to be unobserved by the respondents, we should perhaps not expect an already extensive self-selection, in particular since life insurance is typically purchased at middle age, before many heritable medical conditions have had time to manifest[30].

Contrary to life insurance, we did find a significant but small negative association between our preferred genetic predictor and long-term care insurance. A conceivable mechanism could be that elderly individuals who have observed a decline in their health (partly due to their unobserved genetic risks) chose not to purchase this insurance, or to let it lapse, as they may not expect to reach an age that will require long-term care. The probability of requiring assistance with activities of daily living becomes more substantial after age 85[109], and we found that the top decile of genetic risk survived to just about that age. The associations between our preferred genetic predictor with subjective life expectancy and self-rated health support the idea that unobserved genetic risks had indeed manifested into observable medical conditions. Thus, our results could imply that long-term care insurance is subject to weak adverse selection. However, an alternative explanation could be the already high premium markups reported particularly for this insurance type[109,172], which could have rendered premiums unfair and less attractive for people who consider themselves to be at risk of dying early. Lastly, we believe that the difference between life insurance and long-term care insurance could be explained by the fact that the latter is more often purchased at an older age[173], leaving more time for genetic risks to manifest.

Similarly, our results imply that consumers will at some point have knowledge of genetic risks that they may be incentivized to disclose when purchasing insurance products tied to their survival. This would apply, for example, in the market for "enhanced annuities" (life annuities underwritten with not only demographic but also medical information), which is growing in some countries. Applicants at greater risk could potentially benefit from lower premiums (or higher benefits) if those risks were underwritten[174,175]. It has been reported that standard-rate life annuities may be actuarially unfair in the United States because premiums are determined using low mortality assumptions to counter adverse selection (as it is assumed that this product is bought primarily by the healthiest and wealthiest)[176]. In such a market, the possibility to reveal individual risks could benefit people with a reduced life expectancy and who consider the standard rate expensive. At the same time, some experts argue that an increasing demand for enhanced annuities could crowd out the standard rate product, and genetic testing might exacerbate that development[175]. Overall, we consider further investigation of consumer behavior under conditions of private knowledge of genetic risks to be an interesting avenue of future research.

We think that as the accuracy of genetic predictors matures and more consumers acquire private knowledge of their genetic risk, genetic health information may eventually have to be treated just like any other kind of medical information that can already be requested by insurance providers. There are many scholars who argue that

providers are already entrusted with handling very private and sensitive information, such as medical journals and tests, and that there is no reason to expect them to handle genetic information with any less prudence[31]. But some of them consider it unethical to charge more for genetic factors that are outside the control of the carrier (although it can be argued that many conventional risks are also outside the control of the affected), and that there is a real risk that insurance coverage may be denied to persons who received an outcome in the "genetic lottery" that resulted in greater need of solidarity[60]. Importantly, it could be detrimental to public health if people avoid genetic testing for medical or research purposes out of fear of discrimination[177]. Thus, strong arguments can be made that actuarial discrimination of any kind based on genetic factors should be restricted. For now, the insurance industry has often chosen to self-regulate by imposing moratoriums on the use of genetic information, while at the same time there are reports of ongoing genetic discrimination in countries without legal restrictions or moratoriums[60,64,66,178]. Overall, we encourage policymakers, the insurance industry, and other stakeholders to monitor this development closely and to have a scientifically informed discussion about the potential consequences of price and other discrimination based on genetic information on the one hand and adverse selection on the other.

4.1.2 Limitations

Our results should be considered in light of a few limitations. First, it is likely that mortality selection has led to underestimation of the current performance of polygenic scores. This would have foremost affected the scores for conditions that manifest at younger age. For example, contrary to our expectations, we did not find an association with the score for coronary artery disease, which is a major cause of NCDs[179]. Another important limitation, which is endemic to the GWAS literature, is that we did not study individuals of non-European ancestry. Thus, we do not know whether our findings apply to persons of other ancestry. Unfortunately, it will take many years before we can thoroughly answer that question. Lastly, we acknowledge that the medical conditions ascertained in the HRS are based on self-reports and lack specificity, and it could be that their true impact on survival is understated in our analyses. Future studies, with access to richer medical data, will have to determine more precisely how much information polygenic scores can add above and beyond the various biomarkers for disease and already acquired medical conditions.

4.1.3 Implications for the consumer genetics market

Our results bear on the ongoing debate on whether and, if so, how to regulate the commercialization of genetic health information and the market for third-party interpretation services[27,28,44,180–182]. Many companies sell genetic tests with limited guidance and lengthy disclaimers[32], leaving their customers puzzled[29,43]. For example, it could be financially detrimental for persons dependent on a life insurance policy if this is voluntarily terminated out of a false belief of low genetic risk. Therefore, we agree with others who call for more extensive consumer protection[183]. Overall, more research is needed to determine how vulnerable types of consumers, for example those with particularly high genetic risk or those with weak genetic literacy, react after exploring their DNA for health information. In particular, genetic tests with low accuracy are easily subject to misinterpretation[183]. We therefore think that advertisement of genetic predictions of longevity is for now ethically questionable at many levels and should only be done with great care. We say this even though the appropriateness of such services obviously depends on how they are marketed and how the results are presented. At this moment, however, many genetic interpretation services hide behind extensive disclaimers to void them of responsibility, and that may well be considered a questionable practice[183].

Therefore, as a policy recommendation, we encourage regulatory authorities to consider prognostic genetic testing for disease and longevity to be a form of genetic counseling. As such, we would consider it reasonable to limit the practice to licensed or accredited institutions, whether public or private. But standards for genetic counseling are still maturing globally[184,185], and until regulatory measures are taken and an industry standard has been established, it is likely that appropriate consumer protection will lag behind technological developments. At the same time, since many people appear eager to purchase genetic health information, we consider it undesirable to completely restrict an individual person's right to explore their DNA, with or without the assistance of a certified counselor. Also, the borderless nature of genetic testing, where consumers can send their genetic data to services located in different jurisdictions, makes it practically impossible to effectively regulate this market at national level only. Appropriate regulation of consumer genetics will thus require international agreements in order to be effective.

4.1.4 Conclusions

In conclusion, the estimates presented here clearly show the relevance of polygenic scores in the context of insurance. However, much research is required before it can be determined which polygenic scores may potentially meet the criteria for evidence-based underwriting and before accurate and fair premiums could be developed. Ultimately, policymakers and regulatory agencies will have to strike a difficult balance between keeping private insurance fair and viable on the one hand, while on the other hand ensuring satisfactory consumer protection against genetic discrimination and privacy violations. In the meantime, we see a tangible risk of both genetic discrimination and an informational advantage in favor of the consumer that could lead to adverse selection.

Author contributions

This study was conceived by R.K.L. and designed by R.K.L together with P.D.K. P.D.K oversaw and supervised the study. The preregistered analysis plan was written by R.K.L. with critical input by P.D.K. R.K.L performed the analyses. R.K.L. and P.D.K together interpreted the results. R.K.L wrote the manuscript with critical input from P.D.K. R.K.L. produced the tables and figures.

Acknowledgments

The research was approved by the Research Ethics Review Board (RERB) at the School of Business and Economics (SBE) of Vrije Universiteit Amsterdam (20180927.1.pkr730). We gratefully acknowledge Netspar for financially supporting this research project with a topicality grant (BD2018.01) and thank Bas Werker and Anja de Waegenaere for their valuable feedback on the research questions, study design, and draft manuscript. We are thankful to Casper Burik for constructing the genetic PCs and Aysu Okbay for a providing a meta-analysis of educational attainment that excluded the Health and Retirement Study (HRS). This work was carried out on the Dutch national e-infrastructure (grant 17148) with support of the SURF Cooperative. Philipp Koellinger was financially supported by an ERC Consolidator Grant (647648 EdGe). HRS is supported by the National Institute on Aging (NIA U01AG009740). The genotyping was funded separately by the National Institute on Aging (RC2 AG036495, RC4 AG039029). Genotyping was conducted by the NIH Center for Inherited Disease Research (CIDR) at Johns Hopkins University in Baltimore, Maryland. Genotyping quality control and final preparation of the data were performed by the Genetics Coordinating Center at the University of Washington in Seattle, Washington. Genotype data can be accessed via the database of Genotypes and Phenotypes (dbGaP, http://www.ncbi.nlm.nih.gov/gap, accession number phs000428.v1.p1). Researchers who wish to link genetic data with other HRS measures that are not in dbGaP, such as educational attainment, must apply to the HRS for permission. See the HRS website (http://hrsonline.isr.umich.edu/gwas) for details.

References

1. Hamermesh, D. S. Expectations, Life Expectancy, and Economic Behavior. *Q. J. Econ.* **100**, 389–408 (1985).
2. Bloom, D. E., Canning, D. & Graham, B. Longevity and Life-cycle Savings. *Scand. J. Econ.* **105**, 319–338 (2003).
3. Benítez-Silva, H. & Ni, H. Health status and health dynamics in an empirical model of expected longevity. *J. Health Econ.* **27**, 564–584 (2008).
4. Bloom, D. E., Canning, D. & Moore, M. Optimal Retirement with Increasing Longevity. *Scand. J. Econ.* **116**, 838–858 (2014).
5. van Solinge, H. & Henkens, K. Living longer, working longer? The impact of subjective life expectancy on retirement intentions and behaviour. *Eur. J. Public Health* **20**, 47–51 (2009).
6. Boyer, M., De Donder, P., Fluet, C., Leroux, M-L. & Michaud, P-C. Long-term care risk misperceptions. *Geneva Pap. Risk Insur. Issues Pract.* **44**, 183–215 (2019).
7. Dormont, B., Samson, A. L., Fleurbaey, M., Luchini, S. & Schokkaert, E. Individual Uncertainty About Longevity. *Demography* **55**, 1829–1854 (2018).
8. Griffin, B., Loh, V. & Hesketh, B. A mental model of factors associated with subjective life expectancy. *Soc. Sci. Med.* **82**, 79–86 (2013).
9. Dwyer, D. S. & Mitchell, O. S. Health problems as determinants of retirement: Are self-rated measures endogenous? *J. Health Econ.* **18**, 173–193 (1999).
10. Peracchi, F. & Perotti, V. *Subjective survival probabilities and life tables: Evidence from Europe*. (Einaudi Institute for Economics and Finance (EIEF), 2011).
11. Palloni, A. & Novak, B. Subjective Survival Expectations and Observed Survival: How Consistent Are They? *Vienna Yearb. Popul. Res.* **14**, 187–227 (2016).
12. Benjamins, M. R., Hummer, R. A., Eberstein, I. W. & Nam, C. B. Self-reported health and adult mortality risk: An analysis of cause-specific mortality. *Soc. Sci. Med.* **59**, 1297–1306 (2004).
13. Stenholm, S. *et al.* Self-Rated Health in the Last 12 Years of Life Compared to Matched Surviving Controls: The Health and Retirement Study. *PLoS One* **9**, e107879 (2014).
14. Idler, E. L. & Benyamini, Y. Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies. *J. Health Soc. Behav.* **38**, 21–37 (1997).
15. Perozek, M. Using subjective expectations to forecast longevity: do survey respondents know something we don't know? *Demography* **45**, 95–113 (2008).
16. Deeg, D. *et al.* The rise in life expectancy – corresponding rise in subjective life expectancy? *Netspar Industry Design Paper Series 122* (2015).
17. Lee, L. O. *et al.* Optimism is associated with exceptional longevity in 2 epidemiologic cohorts of men and women. *Proc. Natl. Acad. Sci.* 201900712 (2019).
18. Binder, M. & Buenstorf, G. Smile or die: Can subjective well-being increase survival in the face of substantive health impairments? *Econ. Hum. Biol.* **31**, 209–227 (2018).
19. Kim, E. S. *et al.* Optimism and cause-specific mortality: A prospective cohort study. *Am. J. Epidemiol.* **185**, 21–29 (2017).
20. Hoogwegt, M. T. *et al.* Exercise mediates the association between positive affect and 5-year mortality in patients with ischemic heart disease. *Circ. Cardiovasc. Qual. Outcomes* **6**, 559–566 (2013).
21. Rappange, D. R., Brouwer, W. B. F. & van Exel, J. Rational expectations? An explorative study of subjective survival probabilities and lifestyle across Europe. *Heal. Expect. an Int. J. public Particip. Heal. care Heal. policy* **19**, 121–137 (2016).

22. Kutlu Koç, V. Essays on Subjective Survival Probabilities, Consumption, and Retirement Decisions. (Universiteit Utrecht, 2015).
23. Hurd, M. D. & McGarry, K. The Predictive Validity of Subjective Probabilities of Survival. *Econ. J.* **112**, 966–985 (2002).
24. Gordon, S. Deciphering direct-to-consumer genetic testing. *Medical Economics* (2017).
25. Smart, A., Bolnick, D. A. & Tutton, R. Health and genetic ancestry testing: Time to bridge the gap. *BMC Med. Genomics* **10**, 1–9 (2017).
26. Khan, R. & Mittelman, D. Consumer genomics will change your life, whether you get tested or not. *Genome Biol.* **19**, 120 (2018).
27. Kricka, L. J., Fortina, P., Mai, Y. & Patrinos, G. P. Ethics watch: Direct-access genetic testing: The view from Europe. *Nat. Rev. Genet.* **12**, 670 (2011).
28. Moscarello, T., Murray, B., Reuter, C. M. & Demo, E. Direct-to-consumer raw genetic data and third-party interpretation services: more burden than bargain? *Genet. Med.* **21**, 539–541 (2019).
29. Nelson, S. C., Bowen, D. J. & Fullerton, S. M. Third-Party Genetic Interpretation Tools: A Mixed-Methods Study of Consumer Motivation and Behavior. *Am. J. Hum. Genet.* (2019).
30. Nabholz, C. & Rechfeld, F. *Seeing the future? How genetic testing will impact life insurance.* (2017).
31. Rothstein, M. *Genetics and Life Insurance: Medical Underwriting and Social Policy.* (MIT Press, 2004).
32. Blell, M. & Hunter, M. A. Direct-to-Consumer Genetic Testing's Red Herring: 'Genetic Ancestry' and Personalized Medicine. *Front. Med.* **6**, 48 (2019).
33. Horton, R. *et al.* Direct-to-consumer genetic testing. *BMJ* **367**, l5688 (2019).
34. MyHeritage. MyHeritage DNA Health+Ancestry Test – Discover how your genes may affects your health with comprehensive reports. (2019). Available at: https://blog.myheritage.com/2019/05/introducing-the-myheritage-dna-health-ancestry-test/. (Accessed: 14th June 2019)
35. Chatterjee, N., Shi, J. & García-Closas, M. Developing and evaluating polygenic risk prediction models for stratified disease prevention. *Nat. Rev. Genet.* **17**, 392–406 (2016).
36. Khera, A. V. *et al.* Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat. Genet.* (2018).
37. Desikan, R. S. *et al.* Genetic assessment of age-associated Alzheimer disease risk: Development and validation of a polygenic hazard score. *PLoS Med.* **14**, 1–17 (2017).
38. Khera, A. V *et al.* Polygenic Prediction of Weight and Obesity Trajectories from Birth to Adulthood. *Cell* **177**, 587-596.e9 (2019).
39. Torkamani, A. & Topol, E. Polygenic Risk Scores Expand to Obesity. *Cell* **177**, 518–520 (2019).
40. Abraham, G. *et al.* Genomic risk score offers predictive performance comparable to clinical risk factors for ischaemic stroke. *Nat. Commun.* **10**, 5819 (2019).
41. Larmuseau, M. H. D. Growth of ancestry DNA testing risks huge increase in paternity issues. *Nat. Hum. Behav.* **3**, 5 (2019).
42. Rechfeld, F., Rischatsch, M., Nemoto, T. & Nabholz, C. Can life insurance pass the genetic test? (2019).
43. Wang, C. *et al.* Consumer use and response to online third-party raw DNA interpretation services. *Mol. Genet. Genomic Med.* **6**, 35–43 (2018).
44. Badalato, L., Kalokairinou, L. & Borry, P. Third party interpretation of raw genetic data: an ethical exploration. *Eur. J. Hum. Genet.* **25**, 1189 (2017).
45. Genomelink. Upload DNA Data and Know More About Yourself. *Genomelink.com* (2019).
46. Eling, M. & Ghavibazoo, O. Research on long-term care insurance: status quo and directions for future research. *Geneva Pap. Risk Insur. - Issues Pract.* **44**, 303–356 (2019).

47. Oster, E., Shoulson, I., Quaid, K. & Dorsey, E. R. Genetic adverse selection: Evidence from long-term care insurance and Huntington disease. *J. Public Econ.* **94**, 1041–1050 (2010).
48. Peter, R., Richter, A. & Thistle, P. Endogenous information, adverse selection, and prevention: Implications for genetic testing policy. *J. Health Econ.* **55**, 95–107 (2017).
49. Finkelstein, A. & Poterba, J. Adverse Selection in Insurance Markets: Policyholder Evidence from the U.K. Annuity Market. *J. Polit. Econ.* **112**, 183–208 (2004).
50. Nielsen, R. B., Zick, C. D., Mayer, R. N. & Smith, K. R. Genetic testing and the demand for cancer insurance. *J. Consum. Policy* **24**, 1–21 (2001).
51. Zick, C. D., Smith, K. R., Mayer, R. N. & Botkin, J. R. Genetic testing, adverse selection, and the demand for life insurance. *Am. J. Med. Genet.* **93**, 29–39 (2000).
52. Aktan-Collan, K., Haukkala, A. & Kääriäinen, H. Life and health insurance behaviour of individuals having undergone a predictive genetic testing programme for hereditary non-polyposis colorectal cancer. *Community Genet.* **4**, 219–224 (2001).
53. Zick, C. D. *et al*. Genetic testing for Alzheimer's disease and its impact on insurance purchasing behavior. *Health Aff. (Millwood).* **24**, 483–90 (2005).
54. Gericke, N. *et al*. Exploring Relationships Among Belief in Genetic Determinism, Genetics Knowledge, and Social Factors. *Sci. Educ.* **26**, 1223–1259 (2017).
55. Chapman, R. *et al*. New literacy challenge for the twenty-first century: genetic knowledge is poor even among well educated. *J. Community Genet.* **10**, 73–84 (2019).
56. Carver, R. B., Castéra, J., Gericke, N., Evangelista, N. A. M. & El-Hani, C. N. Young Adults' Belief in Genetic Determinism, and Knowledge and Attitudes towards Modern Genetics and Genomics: The PUGGS Questionnaire. *PLoS One* **12**, e0169808 (2017).
57. Tandy-Connor, S. *et al*. False-positive results released by direct-to-consumer genetic tests highlight the importance of clinical confirmation testing for appropriate patient care. *Genet. Med.* **00**, 1–7 (2018).
58. Lea, D. H., Kaphingst, K. A., Bowen, D., Lipkus, I. & Hadley, D. W. Communicating genetic and genomic information: health literacy and numeracy considerations. *Public Health Genomics* **14**, 279–289 (2011).
59. Hodgson, R. & Haddow, G. *Self-diagnosis – Does the Consumer Have the Advantage Over the Insurer?* (2016).
60. Prince, A. E. R. Comparative perspectives: regulating insurer use of genetic information. *Eur. J. Hum. Genet.* **27**, 340–348 (2019).
61. Sijbrands, E. J. G., Tornij, E. & Homsma, S. J. Mortality risk prediction by an insurance company and long-term follow-up of 62,000 men. *PLoS One* **4**, e5457–e5457 (2009).
62. Nabholz, C. & Somerville, K. *Fair risk assessment in life & health insurance.* (2011).
63. Borry, P. *et al*. Legislation on direct-to-consumer genetic testing in seven European countries. *Eur. J. Hum. Genet.* **20**, 715–721 (2012).
64. Tiller, J. *et al*. Genetic discrimination by Australian insurance companies: a survey of consumer experiences. *Eur. J. Hum. Genet.* (2019).
65. Harper, P. Insurance and genetic testing. *Lancet* **341**, 224–227 (1993).
66. Newson, A. J., Tiller, J., Keogh, L. A., Otlowski, M. & Lacaze, P. Genetics and Insurance in Australia: Concerns around a Self-Regulated Industry. *Public Health Genomics* **20**, 247–256 (2017).
67. Joly, Y. *et al*. Life insurance: Genomic stratification and risk classification. *Eur. J. Hum. Genet.* **22**, 575–579 (2014).
68. Klitzman, R., Appelbaum, P. S. & Chung, W. K. Should Life Insurers Have Access to Genetic Test Results? *JAMA* **312**, 1855–1856 (2014).

69. Van Hoyweghen, I. & Horstman, K. Evidence-based underwriting in the molecular age: the politics of reinsurance companies towards the genetics issue. *New Genet. Soc.* **28**, 317–337 (2009).
70. Prince, A. E. R. Non-Disclosure of Genetic Information in Life Insurance Applications as Grounds for Policy Rescission. *Health Matrix* **26**, 255–307 (2016).
71. Tenenbaum, J. D. & Goodman, K. W. Beyond the Genetic Information Nondiscrimination Act: ethical and economic implications of the exclusion of disability, long-term care and life insurance. *Per. Med.* **14**, 153–157 (2017).
72. Visscher, P. M. *et al.* 10 Years of GWAS Discovery: Biology, Function, and Translation. *Am. J. Hum. Genet.* **101**, 5–22 (2017).
73. Tam, V. *et al.* Benefits and limitations of genome-wide association studies. *Nat. Rev. Genet.* (2019).
74. Polderman, T. J. C. *et al.* Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nat. Genet.* **47**, 702–709 (2015).
75. Pilling, L. C. *et al.* Human longevity: 25 genetic loci associated in 389,166 UK biobank participants. *Aging* **9**, 2504–2520 (2017).
76. Brooks-Wilson, A. R. Genetics of healthy aging and longevity. *Hum. Genet.* **132**, 1323–1338 (2013).
77. Ganna, A. *et al.* Genetic determinants of mortality. Can findings from genome-wide association studies explain variation in human mortality? *Hum. Genet.* **132**, 553–561 (2013).
78. Pilling, L. C. *et al.* Human longevity is influenced by many genetic variants: evidence from 75,000 UK Biobank participants. *Ageing* **8**, 547–560 (2016).
79. Timmers, P. R. H. J. *et al.* Genomics of 1 million parent lifespans implicates novel pathways and common diseases and distinguishes survival chances. *Elife* **8**, e39856 (2019).
80. Phillips, K. A., Deverka, P. A., Hooker, G. W. & Douglas, M. P. Genetic Test Availability And Spending: Where Are We Now? Where Are We Going? *Health Aff.* **37**, 710–716 (2018).
81. Godard, B., ten Kate, L., Evers-Kiebooms, G. & Aymé, S. Population genetic screening programmes: Principles, techniques, practices, and policies. *Eur. J. Hum. Genet.* **11**, 49–87 (2003).
82. Kaplan, W. *et al. Priority Medicines for Europe and the World 2013 Update.* (2013).
83. Jan, S. *et al.* Action to address the household economic burden of non-communicable diseases. *Lancet* **391**, 2047–2058 (2018).
84. Bloom, D. E. *et al. The Global Economic Burden of Non-communicable Diseases.* (2011).
85. Dagenais, G. R. *et al.* Variations in common diseases, hospital admissions, and deaths in middle-aged adults in 21 countries from five continents (PURE): a prospective cohort study. *Lancet* (2019).
86. Daar, A. S. *et al.* Grand challenges in chronic non-communicable diseases. *Nature* **450**, 20–22 (2007).
87. Kranzler, H. R. *et al.* Genome-wide association study of alcohol consumption and use disorder in 274,424 individuals from multiple populations. *Nat. Commun.* **10**, 1499 (2019).
88. Liu, M. *et al.* Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nat. Genet.* (2019).
89. Yusuf, S. *et al.* Modifiable risk factors, cardiovascular disease, and mortality in 155 722 individuals from 21 high-income, middle-income, and low-income countries (PURE): a prospective cohort study. *Lancet* (2019).
90. Dudbridge, F. Polygenic Epidemiology. *Genet. Epidemiol.* **40**, 268–272 (2016).
91. Willer, C. J. *et al.* Discovery and refinement of loci associated with lipid levels. *Nat. Genet.* **45**, 1274–83 (2013).

92. Karlsson Linnér, R. *et al*. Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences. *Nat. Genet.* **51**, 404–413 (2019).
93. Witte, J. S., Visscher, P. M. & Wray, N. R. The contribution of genetic variants to disease depends on the ruler. *Nat. Rev. Genet.* **15**, 765–776 (2014).
94. Mills, M. C. & Rahal, C. A scientometric review of genome-wide association studies. *Commun. Biol.* **2**, 9 (2019).
95. Young, A. I., Benonisdottir, S., Przeworski, M. & Kong, A. Deconstructing the sources of genotype-phenotype associations in humans. *Science* **365**, 1396–1400 (2019).
96. Pasaniuc, B. & Price, A. L. Dissecting the genetics of complex traits using summary association statistics. *Nat. Rev. Genet.* **18**, 117–127 (2017).
97. Saunders, G. *et al*. Leveraging European infrastructures to access 1 million human genomes by 2022. *Nat. Rev. Genet.* (2019).
98. Buniello, A. *et al*. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.* **47**, D1005–D1012 (2018).
99. Torkamani, A., Wineinger, N. E. & Topol, E. J. The personal and clinical utility of polygenic risk scores. *Nat. Rev. Genet.* **19**, 1–10 (2018).
100. Lee, J. J. *et al*. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat. Genet.* **50**, 1112–1121 (2018).
101. Mavaddat, N. *et al*. Polygenic Risk Scores for Prediction of Breast Cancer and Breast Cancer Subtypes. *Am. J. Hum. Genet.* **104**, 21–34 (2019).
102. Escott-Price, V., Shoai, M., Pither, R., Williams, J. & Hardy, J. Polygenic score prediction captures nearly all common genetic risk for Alzheimer's disease. *Neurobiol. Aging* **49**, 214 (2017).
103. Krapohl, E. *et al*. Multi-polygenic score approach to trait prediction. *Mol. Psychiatry* 1–7 (2017).
104. Marioni, R. E. *et al*. Genetic variants linked to education predict longevity. *Proc. Natl. Acad. Sci.* (2016).
105. McDaid, A. F. *et al*. Bayesian association scan reveals loci associated with human lifespan and linked biomarkers. *Nat. Commun.* **8**, 15842 (2017).
106. Joshi, P. K. *et al*. Genome-wide meta-analysis associates HLA-DQA1/DRB1 and LPA and lifestyle factors with human longevity. *Nat. Commun.* **8**, 1–13 (2017).
107. Shin, S. H., Lillard, D. R. & Bhattacharya, J. *UNDERSTANDING THE CORRELATION BETWEEN ALZHEIMER'S DISEASE POLYGENIC RISK, WEALTH, AND THE COMPOSITION OF WEALTH HOLDINGS. NBER Working Paper Series* (2019).
108. Baker, M. & Adams, J. *Approaching the endgame: The future of Defined Benefit pension schemes in the UK*. (2019).
109. *Handbook of Insurance*. (Springer Science+Business Media, 2013).
110. Petkantchin, V. EU anti-discrimination policy's impact on insurance risk management: A parallel with the US sub-prime crisis. *Pensions An Int. J.* **15**, 155–160 (2010).
111. Sonnega, A. *et al*. Cohort Profile: The Health and Retirement Study (HRS). *Int. J. Epidemiol.* **43**, 576–585 (2014).
112. Domingue, B. W. *et al*. Mortality selection in a genetic sample and implications for association studies. *Int. J. Epidemiol.* **46**, 1285–1294 (2017).
113. Mailman, M. D. *et al*. The NCBI dbGaP database of genotypes and phenotypes. *Nat. Genet.* **39**, 1181 (2007).

114. Duncan, L. *et al*. Analysis of polygenic risk score usage and performance in diverse human populations. *Nat. Commun.* **10**, 3328 (2019).
115. Martin, A. R. *et al*. Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations. *Am. J. Hum. Genet.* **100**, 635–649 (2017).
116. Martin, A. R. *et al*. Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat. Genet.* **51**, 584–591 (2019).
117. Auton, A. *et al*. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
118. Price, A. L. *et al*. Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006).
119. Price, A. L., Zaitlen, N. A., Reich, D. & Patterson, N. New approaches to population stratification in genome-wide association studies. *Nat. Rev. Genet.* **11**, 459 (2010).
120. Hamer, D. & Sirota, L. Beware the chopsticks gene. *Mol. Psychiatry* **5**, 11–13 (2000).
121. Evangelou, E. *et al*. Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. *Nat. Genet.* **50**, 1412–1425 (2018).
122. Wang, Z. J. *et al*. Association of body mass index with mortality and cardiovascular events for patients with coronary artery disease: A systematic review and meta-analysis. *Heart* **101**, 1631–1638 (2015).
123. Hoffmann, T. J. *et al*. A large electronic-health-record-based genome-wide study of serum lipids. *Nat. Genet.* **50**, 401–413 (2018).
124. Lleras-Muney, A. The Relationship Between Education and Adult Mortality in the United States. *Rev. Econ. Stud. Mon Rev. Econ. Stud.* **723710723**, 189–221 (2005).
125. Smith, G. D. *et al*. Height and risk of death among men and women: aetiological implications of associations with cardiorespiratory disease and cancer mortality. *J. Epidemiol. Community Health* **54**, 97 LP – 103 (2000).
126. Rehm, J. & Shield, K. D. Global alcohol-attributable deaths from cancer, liver cirrhosis, and injury in 2010. *Alcohol Res.* **35**, 174–183 (2013).
127. Smyth, A. *et al*. Alcohol consumption and cardiovascular disease, cancer, injury, admission to hospital, and mortality: a prospective cohort study. *Lancet* **386**, 1945–1954 (2015).
128. Nielsen, J. B. *et al*. Biobank-driven genomic discovery yields new insight into atrial fibrillation biology. *Nat. Genet.* **50**, 1234–1239 (2018).
129. Winkler, T. W. *et al*. Quality control and conduct of genome-wide association meta-analyses. *Nat. Protoc.* **9**, 1192–212 (2014).
130. Kuchenbaecker, K. B. *et al*. Evaluation of Polygenic Risk Scores for Breast and Ovarian Cancer Risk Prediction in BRCA1 and BRCA2 Mutation Carriers. *JNCI J. Natl. Cancer Inst.* **109**, (2017).
131. Bulik-Sullivan, B. K. *et al*. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
132. Bulik-Sullivan, B. K. *et al*. An Atlas of Genetic Correlations across Human Diseases and Traits. *Nat. Genet.* **47**, 1236–1241 (2015).
133. Lee, J. J., McGue, M., Iacono, W. G. & Chow, C. C. The accuracy of LD Score regression as an estimator of confounding and genetic correlations in genome- wide association studies. *Genet. Epidemiol.* **42**, 783–795 (2018).
134. Daetwyler, H. D., Villanueva, B. & Woolliams, J. a. Accuracy of predicting the genetic risk of disease using a genome-wide approach. *PLoS One* **3**, (2008).
135. Dudbridge, F. Power and Predictive Accuracy of Polygenic Risk Scores. *PLoS Genet.* **9**, e1003348 (2013).
136. Choi, S. W. & O'Reilly, P. F. PRSice-2: Polygenic Risk Score software for biobank-scale data. *Gigascience* **8**, (2019).
137. Altshuler, D. M. *et al*. Integrating common and rare genetic variation in diverse human populations. *Nature* **467**, 52–58 (2010).

138. Chang, C. C. *et al*. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
139. Jansen, I. E. *et al*. Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nat. Genet.* (2019).
140. Malik, R. *et al*. Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nat. Genet.* **50**, 524–537 (2018).
141. Yengo, L. *et al*. Meta-analysis of genome-wide association studies for height and body mass index in - 700000 individuals of European ancestry. *Hum. Mol. Genet.* **27**, 3641–3649 (2018).
142. Michailidou, K. *et al*. Association analysis identifies 65 new breast cancer risk loci. *Nature* **551**, 92 (2017).
143. Wuttke, M. *et al*. A catalog of genetic loci associated with kidney function from analyses of a million individuals. *Nat. Genet.* **51**, 957–972 (2019).
144. Abbott, L. *et al*. RAPID GWAS OF THOUSANDS OF PHENOTYPES FOR 337,000 SAMPLES IN THE UK BIOBANK. *nealelab.is* (2017). Available at: http://www.nealelab.is/blog/2017/7/19/rapid-gwas-of-thousands-of-phenotypes-for-337000-samples-in-the-uk-biobank.
145. van der Harst, P. & Verweij, N. Identification of 64 Novel Genetic Loci Provides an Expanded View on the Genetic Architecture of Coronary Artery Disease. *Circ. Res.* **122**, 433–443 (2018).
146. Howard, D. M. *et al*. Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nat. Neurosci.* (2019).
147. Jansen, P. R. *et al*. Genome-wide analysis of insomnia in 1,331,010 individuals identifies new risk loci and functional pathways. *Nat. Genet.* (2019).
148. Schumacher, F. R. *et al*. Association analyses of more than 140,000 men identify 63 new prostate cancer susceptibility loci. *Nat. Genet.* **50**, 928–936 (2018).
149. Pardiñas, A. F. *et al*. Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nat. Genet.* (2018).
150. Mahajan, A. *et al*. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nat. Genet.* (2018).
151. Kaplan, E. L. & Meier, P. Nonparametric Estimation from Incomplete Observations. *J. Am. Stat. Assoc.* **53**, 457–481 (1958).
152. Clark, T. G., Bradburn, M. J., Love, S. B. & Altman, D. G. Survival Analysis Part I: Basic concepts and first analyses. *Br. J. Cancer* **89**, 232 (2003).
153. Wright, K. M. *et al*. A Prospective Analysis of Genetic Variants Associated with Human Lifespan. *G3 Genes|Genomes|Genetics* **9**, 2863 LP – 2878 (2019).
154. R: A Language and Environment for Statistical Computing. (2019).
155. Therneau, T. M. & Lumley, T. Package 'survival'. (2019).
156. Kassambara, A., Kosinski, M., Biecek, P. & Fabian, S. Package 'survminer'. (2019).
157. Potapov, A. S., Adler, W. & Schmid, M. Package ' survAUC '. (2015).
158. Haibe-kains, A. B. *et al*. Package ' survcomp '. (2019).
159. Hosmer, D. W. J. & Lemeshow, S. *Applied Survival Analysis: Regression Modeling of Time to Event Data*. (Wiley & Sons, Ltd, 1998).
160. Bradburn, M. J., Clark, T. G., Love, S. B. & Altman, D. G. Survival Analysis Part II: Multivariate data analysis – an introduction to concepts and methods. *Br. J. Cancer* **89**, 431 (2003).
161. Royston, P. & Altman, D. G. External validation of a Cox prognostic model: principles and methods. *BMC Med. Res. Methodol.* **13**, 33 (2013).
162. *Applied survival analysis using R*. (Springer, 2016).
163. Bradburn, M. J., Clark, T. G., Love, S. B. & Altman, D. G. Survival Analysis Part III: Multivariate data analysis – choosing a model and assessing its adequacy and fit. *Br. J. Cancer* **89**, 605 (2003).

164. Klein, J. P. & Moeschberger, M. L. *Survival analysis: techniques for censored and truncated data*. (Springer, 2003).
165. Harrell Jr, F. E., Califf, R. M., Pryor, D. B., Lee, K. L. & Rosati, R. A. Evaluating the Yield of Medical Tests. *JAMA* **247**, 2543–2546 (1982).
166. Royston, P. & Sauerbrei, W. A new measure of prognostic separation in survival data. *Stat. Med.* **23**, 723–748 (2004).
167. Preacher, K. J., Rucker, D. D., MacCallum, R. C. & Nicewander, W. A. Use of the extreme groups approach: a critical reexamination and new recommendations. *Psychol. Methods* **10**, 178–192 (2005).
168. Noone, A. *et al. SEER Cancer Statistics Review, 1975-2015, National Cancer Institute*. (American Cancer Society, 2018).
169. de Vlaming, R. *et al*. Meta-GWAS Accuracy and Power (MetaGAP) Calculator Shows that Hiding Heritability Is Partially Due to Imperfect Genetic Correlations across Studies. *PLOS Genet.* **13**, e1006495 (2017).
170. Harris, T. & Yelowitz, A. Is there adverse selection in the life insurance market? Evidence from a representative sample of purchasers. *Econ. Lett.* **124**, 520–522 (2014).
171. McCarthy, D. & Mitchell, O. S. International Adverse Selection in Life Insurance and Annuities. *Aging Adv. Ind. States* 119–135 (2010).
172. Glenzer, F. & Achou, B. Annuities, long-term care insurance, and insurer solvency. *Geneva Pap. Risk Insur. - Issues Pract.* **44**, 252–276 (2019).
173. Cornell, P. Y., Grabowski, D. C., Cohen, M., Shi, X. & Stevenson, D. G. Medical Underwriting In Long-Term Care Insurance: Market Conditions Limit Options For Higher-Risk Consumers. *Health Aff. (Millwood).* **35**, 1494–1503 (2016).
174. Veyssiere, L. *et al. Study on the performance and adequacy of pension decumulation practices in four EU countries*. (2017).
175. Steinorth, P. The demand for enhanced annuities. *J. Public Econ.* **96**, 973–980 (2012).
176. Brown, R. L. & Scahill, P. *Issues in the Issuance of Enhanced Annuities*. (2010).
177. Keogh, L. *et al*. Choosing not to undergo predictive genetic testing for hereditary colorectal cancer syndromes: expanding our understanding of decliners and declining. *J. Behav. Med.* **40**, 583–594 (2017).
178. Barlow-Stewart, K., Liepins, M., Doble, A. & Otlowski, M. How are genetic test results being used by Australian life insurers? *Eur. J. Hum. Genet.* **26**, 1248–1256 (2018).
179. Lloyd-Jones, D. M. *et al*. Prediction of lifetime risk for cardiovascular disease by risk factor burden at 50 years of age. *Circulation* **113**, 791–798 (2006).
180. Allyse, M. A., Robinson, D. H., Ferber, M. J. & Sharp, R. R. Direct-to-Consumer Testing 2.0: Emerging Models of Direct-to-Consumer Genetic Testing. *Mayo Clin. Proc.* **93**, 113–120 (2018).
181. Van Hellemondt, R. E., Hendriks, A. & Breuning, M. *Regulating the Use of Genetic Tests: Is Dutch Law an Example for Other Countries with Regard to DTC Genetic Testing. Amsterdam Law Forum* (2011).
182. Howard, H. C. & Borry, P. Europe and direct-to-consumer genetic tests. *Nat. Rev. Genet.* **13**, 146 (2012).
183. Schleit, J., Naylor, L. V & Hisama, F. M. First, do no harm: direct-to-consumer genetic testing. *Genet. Med.* **21**, 510–511 (2019).
184. Ormond, K. E. *et al*. Genetic counseling globally: Where are we now? *Am. J. Med. Genet. Part C Semin. Med. Genet.* **178**, 98–107 (2018).
185. Abacan, M. *et al*. The Global State of the Genetic Counseling Profession. *Eur. J. Hum. Genet.* **27**, 183–197 (2019).

OVERZICHT UITGAVEN IN DE DESIGN PAPER SERIE

1 Naar een nieuw pensioencontract (2011)
Lans Bovenberg en Casper van Ewijk

2 Langlevenrisico in collectieve pensioencontracten (2011)
Anja De Waegenaere, Alexander Paulis en Job Stigter

3 Bouwstenen voor nieuwe pensioencontracten en uitdagingen voor het toezicht daarop (2011)
Theo Nijman en Lans Bovenberg

4 European supervision of pension funds: purpose, scope and design (2011)
Niels Kortleve, Wilfried Mulder and Antoon Pelsser

5 Regulating pensions: Why the European Union matters (2011)
Ton van den Brink, Hans van Meerten and Sybe de Vries

6 The design of European supervision of pension funds (2012)
Dirk Broeders, Niels Kortleve, Antoon Pelsser and Jan-Willem Wijckmans

7 Hoe gevoelig is de uittredeleeftijd voor veranderingen in het pensioenstelsel? (2012)
Didier Fouarge, Andries de Grip en Raymond Montizaan

8 De inkomensverdeling en levensverwachting van ouderen (2012)
Marike Knoef, Rob Alessie en Adriaan Kalwij

9 Marktconsistente waardering van zachte pensioenrechten (2012)
Theo Nijman en Bas Werker

10 De RAM in het nieuwe pensioenakkoord (2012)
Frank de Jong en Peter Schotman

11 The longevity risk of the Dutch Actuarial Association's projection model (2012)
Frederik Peters, Wilma Nusselder and Johan Mackenbach

12 Het koppelen van pensioenleeftijd en pensioenaanspraken aan de levensverwachting (2012)
Anja De Waegenaere, Bertrand Melenberg en Tim Boonen

13 Impliciete en expliciete leeftijdsdifferentiatie in pensioencontracten (2013)
Roel Mehlkopf, Jan Bonenkamp, Casper van Ewijk, Harry ter Rele en Ed Westerhout

14 Hoofdlijnen Pensioenakkoord, juridisch begrepen (2013)
Mark Heemskerk, Bas de Jong en René Maatman

15 Different people, different choices: The influence of visual stimuli in communication on pension choice (2013)
Elisabeth Brüggen, Ingrid Rohde and Mijke van den Broeke

16 Herverdeling door pensioenregelingen (2013)
Jan Bonenkamp, Wilma Nusselder, Johan Mackenbach, Frederik Peters en Harry ter Rele

17 Guarantees and habit formation in pension schemes: A critical analysis of the floor-leverage rule (2013)
Frank de Jong and Yang Zhou

18 The holistic balance sheet as a building block in pension fund supervision (2013)
Erwin Fransen, Niels Kortleve, Hans Schumacher, Hans Staring and Jan-Willem Wijckmans

19 Collective pension schemes and individual choice (2013)
Jules van Binsbergen, Dirk Broeders, Myrthe de Jong and Ralph Koijen

20 Building a distribution builder: Design considerations for financial investment and pension decisions (2013)
Bas Donkers, Carlos Lourenço, Daniel Goldstein and Benedict Dellaert

21 Escalerende garantietoezeggingen: een alternatief voor het StAr RAM-contract (2013)
Servaas van Bilsen, Roger Laeven en Theo Nijman
22 A reporting standard for defined contribution pension plans (2013)
Kees de Vaan, Daniele Fano, Herialt Mens and Giovanna Nicodano
23 Op naar actieve pensioenconsumenten: Inhoudelijke kenmerken en randvoorwaarden van effectieve pensioencommunicatie (2013)
Niels Kortleve, Guido Verbaal en Charlotte Kuiper
24 Naar een nieuw deelnemergericht UPO (2013)
Charlotte Kuiper, Arthur van Soest en Cees Dert
25 Measuring retirement savings adequacy; developing a multi-pillar approach in the Netherlands (2013)
Marike Knoef, Jim Been, Rob Alessie, Koen Caminada, Kees Goudswaard, and Adriaan Kalwij
26 Illiquiditeit voor pensioenfondsen en verzekeraars: Rendement versus risico (2014)
Joost Driessen
27 De doorsneesystematiek in aanvullende pensioenregelingen: effecten, alternatieven en transitiepaden (2014)
Jan Bonenkamp, Ryanne Cox en Marcel Lever
28 EIOPA: bevoegdheden en rechtsbescherming (2014)
Ivor Witte
29 Een institutionele beleggersblik op de Nederlandse woningmarkt (2013)
Dirk Brounen en Ronald Mahieu
30 Verzekeraar en het reële pensioencontract (2014)
Jolanda van den Brink, Erik Lutjens en Ivor Witte
31 Pensioen, consumptiebehoeften en ouderenzorg (2014)
Marike Knoef, Arjen Hussem, Arjan Soede en Jochem de Bresser
32 Habit formation: implications for pension plans (2014)
Frank de Jong and Yang Zhou
33 Het Algemeen pensioenfonds en de taakafbakening (2014)
Ivor Witte
34 Intergenerational Risk Trading (2014)
Jiajia Cui and Eduard Ponds
35 Beëindiging van de doorsneesystematiek: juridisch navigeren naar alternatieven (2015)
Dick Boeijen, Mark Heemskerk en René Maatman
36 Purchasing an annuity: now or later? The role of interest rates (2015)
Thijs Markwat, Roderick Molenaar and Juan Carlos Rodriguez
37 Entrepreneurs without wealth? An overview of their portfolio using different data sources for the Netherlands (2015)
Mauro Mastrogiacomo, Yue Li and Rik Dillingh
38 The psychology and economics of reverse mortgage attitudes. Evidence from the Netherlands (2015)
Rik Dillingh, Henriëtte Prast, Mariacristina Rossi and Cesira Urzì Brancati
39 Keuzevrijheid in de uittreedleeftijd (2015)
Arthur van Soest
40 Afschaffing doorsneesystematiek: verkenning van varianten (2015)
Jan Bonenkamp en Marcel Lever
41 Nederlandse pensioenopbouw in internationaal perspectief (2015)
Marike Knoef, Kees Goudswaard, Jim Been en Koen Caminada
42 Intergenerationele risicodeling in collectieve en individuele pensioencontracten (2015)
Jan Bonenkamp, Peter Broer en Ed Westerhout
43 Inflation Experiences of Retirees (2015)
Adriaan Kalwij, Rob Alessie, Jonathan Gardner and Ashik Anwar Ali
44 Financial fairness and conditional indexation (2015)
Torsten Kleinow and Hans Schumacher
45 Lessons from the Swedish occupational pension system (2015)
Lans Bovenberg, Ryanne Cox and Stefan Lundbergh

46 Heldere en harde pensioenrechten onder een PPR (2016)
Mark Heemskerk, René Maatman en Bas Werker

47 Segmentation of pension plan participants: Identifying dimensions of heterogeneity (2016)
Wiebke Eberhardt, Elisabeth Brüggen, Thomas Post and Chantal Hoet

48 How do people spend their time before and after retirement? (2016)
Johannes Binswanger

49 Naar een nieuwe aanpak voor risicoprofielmeting voor deelnemers in pensioenregelingen (2016)
Benedict Dellaert, Bas Donkers, Marc Turlings, Tom Steenkamp en Ed Vermeulen

50 Individueel defined contribution in de uitkeringsfase (2016)
Tom Steenkamp

51 Wat vinden en verwachten Nederlanders van het pensioen? (2016)
Arthur van Soest

52 Do life expectancy projections need to account for the impact of smoking? (2016)
Frederik Peters, Johan Mackenbach en Wilma Nusselder

53 Effecten van gelaagdheid in pensioendocumenten: een gebruikersstudie (2016)
Louise Nell, Leo Lentz en Henk Pander Maat

54 Term Structures with Converging Forward Rates (2016)
Michel Vellekoop and Jan de Kort

55 Participation and choice in funded pension plans (2016)
Manuel García-Huitrón and Eduard Ponds

56 Interest rate models for pension and insurance regulation (2016)
Dirk Broeders, Frank de Jong and Peter Schotman

57 An evaluation of the nFTK (2016)
Lei Shu, Bertrand Melenberg and Hans Schumacher

58 Pensioenen en inkomensongelijkheid onder ouderen in Europa (2016)
Koen Caminada, Kees Goudswaard, Jim Been en Marike Knoef

59 Towards a practical and scientifically sound tool for measuring time and risk preferences in pension savings decisions (2016)
Jan Potters, Arno Riedl and Paul Smeets

60 Save more or retire later? Retirement planning heterogeneity and perceptions of savings adequacy and income constraints (2016)
Ron van Schie, Benedict Dellaert and Bas Donkers

61 Uitstroom van oudere werknemers bij overheid en onderwijs. Selectie uit de poort (2016)
Frank Cörvers en Janneke Wilschut

62 Pension risk preferences. A personalized elicitation method and its impact on asset allocation (2016)
Gosse Alserda, Benedict Dellaert, Laurens Swinkels and Fieke van der Lecq

63 Market-consistent valuation of pension liabilities (2016)
Antoon Pelsser, Ahmad Salahnejhad and Ramon van den Akker

64 Will we repay our debts before retirement? Or did we already, but nobody noticed? (2016)
Mauro Mastrogiacomo

65 Effectieve ondersteuning van zelfmanagement voor de consument (2016)
Peter Lapperre, Alwin Oerlemans en Benedict Dellaert

66 Risk sharing rules for longevity risk: impact and wealth transfers (2017)
Anja De Waegenaere, Bertrand Melenberg and Thijs Markwat

67 Heterogeniteit in doorsneeproblematiek. Hoe pakt de transitie naar degressieve opbouw uit voor verschillende pensioenfondsen? (2017)
Loes Frehen, Wouter van Wel, Casper van Ewijk, Johan Bonekamp, Joost van Valkengoed en Dick Boeijen

68 De toereikendheid van pensioenopbouw na de crisis en pensioenhervormingen (2017)
Marike Knoef, Jim Been, Koen Caminada, Kees Goudswaard en Jason Rhuggenaath

69 De combinatie van betaald en onbetaald werk in de jaren voor pensioen (2017)
Marleen Damman en Hanna van Solinge

70 Default life-cycles for retirement savings (2017)
Anna Grebenchtchikova, Roderick Molenaar, Peter Schotman en Bas Werker

71 Welke keuzemogelijkheden zijn wenselijk vanuit het perspectief van de deelnemer? (2017)
Casper van Ewijk, Roel Mehlkopf, Sara van den Bleeken en Chantal Hoet

72 Activating pension plan participants: investment and assurance frames (2017)
Wiebke Eberhardt, Elisabeth Brüggen, Thomas Post en Chantal Hoet

73 Zerotopia – bounded and unbounded pension adventures (2017)
Samuel Sender

74 Keuzemogelijkheden en maatwerk binnen pensioenregelingen (2017)
Saskia Bakels, Agnes Joseph, Niels Kortleve en Theo Nijman

75 Polderen over het pensioenstelsel. Het debat tussen de sociale partners en de overheid over de oudedagvoorzieningen in Nederland, 1945-2000 (2017)
Paul Brusse

76 Van uitkeringsovereenkomst naar PPR (2017)
Mark Heemskerk, Kees Kamminga, René Maatman en Bas Werker

77 Pensioenresultaat bij degressieve opbouw en progressieve premie (2017)
Marcel Lever en Sander Muns

78 Bestedingsbehoeften bij een afnemende gezondheid na pensionering (2017)
Lieke Kools en Marike Knoef

79 Model Risk in the Pricing of Reverse Mortgage Products (2017)
Anja De Waegenaere, Bertrand Melenberg, Hans Schumacher, Lei Shu and Lieke Werner

80 Expected Shortfall voor toezicht op verzekeraars: is het relevant? (2017)
Tim Boonen

81 The Effect of the Assumed Interest Rate and Smoothing on Variable Annuities (2017)
Anne G. Balter and Bas J.M. Werker

82 Consumer acceptance of online pension investment advice (2017)
Benedict Dellaert, Bas Donkers and Carlos Lourenço

83 Individualized life-cycle investing (2017)
Gréta Oleár, Frank de Jong and Ingmar Minderhoud

84 The value and risk of intergenerational risk sharing (2017)
Bas Werker

85 Pensioenwensen voor en na de crisis (2017)
Jochem de Bresser, Marike Knoef en Lieke Kools

86 Welke vaste dalingen en welk beleggingsbeleid passen bij gewenste uitkeringsprofielen in verbeterde premieregelingen? (2017)
Johan Bonekamp, Lans Bovenberg, Theo Nijman en Bas Werker

87 Inkomens- en vermogensafhankelijke eigen bijdragen in de langdurige ouderenzorg: een levensloopperspectief (2017)
Arjen Hussem, Harry ter Rele en Bram Wouterse

88 Creating good choice environments – Insights from research and industry practice (2017)
Elisabeth Brüggen, Thomas Post and Kimberley van der Heijden

89 Two decades of working beyond age 65 in the Netherlands. Health trends and changes in socio-economic and work factors to determine the feasibility of extending working lives beyond age 65 (2017)
Dorly Deeg, Maaike van der Noordt and Suzan van der Pas

90 Cardiovascular disease in older workers. How can workforce participation be maintained in light of changes over time in determinants of cardiovascular disease? (2017)
Dorly Deeg, E. Burgers and Maaike van der Noordt

91 Zicht op zzp-pensioen (2017)
Wim Zwinkels, Marike Knoef, Jim Been, Koen Caminada en Kees Goudswaard

92 Return, risk, and the preferred mix of PAYG and funded pensions (2017)
Marcel Lever, Thomas Michielsen and Sander Muns

93 Life events and participant engagement in pension plans (2017)
Matthew Blakstad, Elisabeth Brüggen and Thomas Post

94 Parttime pensioneren en de arbeidsparticipatie (2017)
Raymond Montizaan

95 Keuzevrijheid in pensioen: ons brein wil niet kiezen, maar wel gekozen hebben (2018)
Walter Limpens en Joyce Vonken

96 Employability after age 65? Trends over 23 years in life expectancy in good and in poor physical and cognitive health of 65-74-year-olds in the Netherlands (2018)
Dorly Deeg, Maaike van der Noordt, Emiel Hoogendijk, Hannie Comijs and Martijn Huisman

97 Loslaten van de verplichte pensioenleeftijd en het organisatieklimaat rondom langer doorwerken (2018)
Jaap Oude Mulders, Kène Henkens en Harry van Dalen

98 Overgangseffecten bij introductie degressieve opbouw (2018)
Bas Werker

99 You're invited – RSVP! The role of tailoring in incentivising people to delve into their pension situation (2018)
Milena Dinkova, Sanne Elling, Adriaan Kalwij en Leo Lentz

100 Geleidelijke uittreding en de rol van deeltijdpensioen (2018)
Jonneke Bolhaar en Daniël van Vuuren

101 Naar een model voor pensioencommunicatie (2018)
Leo Lentz, Louise Nell en Henk Pander Maat

102 Tien jaar UPO. Een terugblik en vooruitblik op inhoud, doelen en effectiviteit (2018)
Sanne Elling en Leo Lentz

103 Health and household expenditures (2018)
Raun van Ooijen, Jochem de Bresser en Marike Knoef

104 Keuzevrijheid in de uitkeringsfase: internationale ervaringen (2018)
Marcel Lever, Eduard Ponds, Rik Dillingh en Ralph Stevens

105 The move towards riskier pension products in the world's best pension systems (2018)
Anne G. Balter, Malene Kallestrup-Lamb and Jesper Rangvid

106 Life Cycle Option Value: The value of consumer flexibility in planning for retirement (2018)
Sonja Wendel, Benedict Dellaert and Bas Donkers

107 Naar een duidelijk eigendomsbegrip (2018)
Jop Tangelder

108 Effect van stijging AOW-leeftijd op arbeidsongeschiktheid (2018)
Rik Dillingh, Jonneke Bolhaar, Marcel Lever, Harry ter Rele, Lisette Swart en Koen van der Ven

109 Is de toekomst gearriveerd? Data science en individuele keuzemogelijkheden in pensioen (2018)
Wesley Kaufmann, Bastiaan Starink en Bas Werker

110 De woontevredenheid van ouderen in Nederland (2018)
Jan Rouwendal

111 Towards better prediction of individual longevity (2018)
Dorly Deeg, Jan Kardaun, Maaike van der Noordt, Emiel Hoogendijk en Natasja van Schoor

112 Framing in pensioenkeuzes. Het effect van framing in de keuze voor beleggingsprofiel in DC-plannen naar aanleiding van de Wet verbeterde premieregeling (2018)
Marijke van Putten, Rogier Potter van Loon, Marc Turlings en Eric van Dijk

113 Working life expectancy in good and poor self-perceived health among Dutch workers aged 55–65 years with a chronic disease over the period 1992–2016 (2019)
Astrid de Wind, Maaike van der Noordt, Dorly Deeg and Cécile Boot

114 Working conditions in post-retirement jobs: A European comparison (2019)
Ellen Dingemans and Kène Henkens

115 Is additional indebtedness the way to increase mortgage-default insurance coverage? (2019)
Yeorim Kim, Mauro Mastrogiacomo, Stefan Hochguertel and Hans Bloemen

116 Appreciated but complicated pension Choices? Insights from the Swedish Premium Pension System (2019)
Monika Böhnke, Elisabeth Brüggen and Thomas Post

117 Towards integrated personal financial planning. Information barriers and design propositions (2019)
Nitesh Bharosa and Marijn Janssen

118 The effect of tailoring pension information on navigation behavior (2019)
Milena Dinkova, Sanne Elling, Adriaan Kalwij and Leo Lentz

119 Opleiding, levensverwachting en pensioenleeftijd: een vergelijking van Nederland met andere Europese landen (2019)
Johan Mackenbach, José Rubio Valverde en Wilma Nusselder

120 Giving with a warm hand: Evidence on estate planning and bequests (2019)
Eduard Suari-Andreu, Raun van Ooijen, Rob J.M. Alessie and Viola Angelini

121 Investeren in menselijk kapitaal: een gecombineerd werknemers- en werkgeversperspectief (2019)
Raymond Montizaan, Merlin Nieste en Davey Poulissen

122 The rise in life expectancy – corresponding rise in subjective life expectancy? Changes over the period 1999-2016 (2019)
Dorly Deeg, Maaike van der Noordt, Noëlle Sant, Henrike Galenkamp, Fanny Janssen and Martijn Huisman

123 Pensioenaanvullingen uit het eigen woningbezit (2019)
Dirk Brounen, Niels Kortleve en Eduard Ponds

124 Personal and work-related predictors of early exit from paid work among older workers with health limitations (2019)
Nils Plomp, Sascha de Breij and Dorly Deeg

125 Het delen van langlevenrisico (2019)
Anja De Waegenaere, Agnes Joseph, Pascal Janssen en Michel Vellekoop

126 Maatwerk in pensioencommunicatie (2019)
S.K. Elling en L.R. Lentz

127 Dutch Employers' Responses to an Aging Workforce: Evidence from Surveys, 2009-2017 (2019)
Jaap Oude Mulders, Kène Henkens and Hendrik P. van Dalen

128 Preferences for solidarity and attitudes towards the Dutch pension system – Evidence from a representative sample (2019)
Arno Riedl, Hans Schmeets and Peter Werner

129 Deeltijdpensioen geen wondermiddel voor langer doorwerken (2019)
Henk-Wim de Boer, Tunga Kantarcı, Daniel van Vuuren en Ed Westerhout

130 Spaarmotieven en consumptiegedrag (2019)
Johan Bonekamp en Arthur van Soest

131 Substitute services: a barrier to controlling long-term care expenditures (2019)
Mark Kattenberg and Pieter Bakx

132 Voorstel keuzearchitectuur pensioensparen voor zelfstandigen (2019)
Jona Linde

133 The impact of the virtual integration of assets on pension risk preferences of individuals (2019)
Sesil Lim, Bas Donkers en Benedict Dellaert

134 Reforming the statutory retirement age: Policy preferences of employers (2019)
Hendrik P. van Dalen, Kène Henkens and Jaap Oude Mulders

135 Compensatie bij afschaffing doorsnee-systematiek (2019)
Dick Boeijen, Chantal de Groot, Mark Heemskerk, Niels Kortleve en René Maatman

136 Debt affordability after retirement, interest rate shocks and voluntary repayments (2019)
Mauro Mastrogiacomo

137 Using social norms to activate pension plan members: insights from practice (2019)
Joyce Augustus-Vonken, Pieter Verhallen, Lisa Brüggen and Thomas Post

138 Alternatieven voor de huidige verplichtstelling van bedrijfstakpensioenfondsen (2020)
Erik Lutjens en Fieke van der Lecq

139 Eigen bijdrage aan ouderenzorg (2020)
Pieter Bakx, Judith Bom, Marianne Tenand en Bram Wouterse

140 Inrichting fiscaal kader bij afschaffing doorsneesystematiek (2020)
Bastiaan Starink en Michael Visser

141 Hervorming langdurige zorg: trends in het gebruik van verpleging en verzorging (2020)
Pieter Bakx, Pilar Garcia-Gomez, Sara Rellstab, Erik Schut en Eddy van Doorslaer

142 Genetic health risks, insurance, and retirement (2020)
Richard Karlsson Linnér and Philipp D. Koellinger

Netspar

Network for Studies on Pensions, Aging and Retirement

This is a publication of:
Netspar
Phone +31 13 466 2109
E-mail info@netspar.nl
www.netspar.nl

April 2020