

# Identification of polymorphisms in cancer patients that differentially affect survival with age

Aoife Doherty<sup>1</sup>, Yelena Kernogitski<sup>2</sup>, Alexander M. Kulminski<sup>2</sup>, João Pedro de Magalhães<sup>1</sup>

<sup>1</sup>Integrative Genomics of Ageing Group, Institute of Ageing and Chronic Disease, University of Liverpool, Liverpool, L7 8TX, United Kingdom

<sup>2</sup>Biodemography of Aging Research Unit (BARU), Social Science Research Institute, Duke University, Durham, NC 27708, USA

**Correspondence to:** João Pedro de Magalhães; **email:** [jp@senescence.info](mailto:jp@senescence.info)

**Keywords:** ageing, genetics, geriatric oncology, SNP, longevity, WRN

**Received:** May 19, 2017 **Accepted:** October 6, 2017 **Published:** October 20, 2017

**Copyright:** Doherty et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

## ABSTRACT

The World Health Organization predicts that the proportion of the world's population over 60 will almost double from 12% to 22% between 2015 and 2050. Ageing is the biggest risk factor for cancer, which is a leading cause of deaths worldwide. Unfortunately, research describing how genetic variants affect cancer progression commonly neglects to account for the ageing process. Herein is the first systematic analysis that combines a large longitudinal data set with a targeted candidate gene approach to examine the effect of genetic variation on survival as a function of age in cancer patients. Survival was significantly decreased in individuals with heterozygote or rare homozygote (i.e. variant) genotypes compared to those with a common homozygote genotype (i.e. wild type) for two single nucleotide polymorphisms (rs11574358 and rs4147918), one gene (*SIRT3*) and one pathway (FoxO signalling) in an age-dependent manner. All identified genes and pathways have previously been associated with ageing and cancer. These observations demonstrate that there are ageing-related genetic elements that differentially affect mortality in cancer patients in an age-dependent manner. Understanding the genetic determinants affecting prognosis differently with age will be invaluable to develop age-specific prognostic biomarkers and personalized therapies that may improve clinical outcomes for older individuals.

## INTRODUCTION

Cancer is a leading cause of death worldwide; approximately 14.1 million new cancer cases and 8.2 million cancer-related deaths were recorded globally in 2012 [1]. Ageing is the biggest risk factor for cancer, and the majority of tumours are diagnosed in patients older than 60 years [2,3]. The World Health Organization predicts that the proportion of the world's population over the age of 60 will almost double from 12% to 22% between 2015 and 2050 (World Health Organization, 2015). The evolving age demography affects cancer incidence and mortality rates, which has

serious consequences for a country's healthcare system and economy [4,5]. Novel insights into the age-related genetic predisposition of cancer survival would be a major breakthrough in expanding healthy life span in humans.

The last decade has seen extensive efforts to catalogue human genetic variation [6,7] and to correlate variation with phenotypic traits. For example, single nucleotide polymorphisms (SNPs) have been assessed for statistical associations with complex traits such as longevity [8–14] and various common diseases [15–22]. Case-control studies that compare population SNP

frequency to disease characteristics often simply list a set of SNPs statistically significantly associated with a particular condition, not accounting for the ageing process. Unfortunately, this is an oversight, as molecular systems affected by such genetic variants are evolving entities whose interactions change with age [23,24]. For example, the ageing process affects multiple inter-linked molecular systems including the immune [3,25,26], metabolic [27,28] and cardiovascular [29] systems.

Ageing affects cancer incidence rates [30,31], prognosis [32–34] and drug response [35]. Recently, Kulminski et al. (2014) investigated the effects of the e4 allele of the *APOE* gene on human survival in a range of ages from mid-life until extreme old age, and the sensitivity of those effects to cardiovascular disease, cancer and neurodegenerative disease [12]. This allele is thought to have a protective effect against early life infectious disease such as diarrhea and liver damage caused by Hepatitis C virus infection. Their research suggested that, although there is an advantage to the allele in early life, there is a significant adverse effect of the e4 allele on survival that is limited to women with a moderate lifespan (70–95 years). Furthermore, non-skin cancer increased the risk of death of e4 carriers two-fold compared to non-e4 carriers among women of moderate lifespans. These observations suggest the existence of age- and gender-sensitive systemic mechanisms linking the e4 allele to lifespan that can non-additively interfere with cancer-related mechanisms. The research described herein combines the availability of a large longitudinal data set with a targeted candidate gene approach to examine the effect of genetic variation on survival as a function of age from a systematic perspective. Insights obtained from this novel investigation are of high biological importance, as understanding the biomarkers and molecular mechanisms that affect cancer prognosis in an age-dependent manner will provide critical information for age-specific patient outcome and relevant assignment to therapies; an area of research that will only become even more important with a greying population.

## RESULTS

### Data set assembly and quality filtering: phenotypic and genotypic data

The population under study in this analysis is the Framingham Heart Study (FHS). In brief, the FHS comprises >10,000 individuals in different cohorts who have been examined every 2–4 years for up to 60 years (depending on cohort; see Methods). The FHS has previously successfully addressed interesting biological questions related to ageing and disease [11,12,37,38].

Using FHS data, two cancer data sets were assembled based on tumour topography for all patients that were first diagnosed with cancer over the age of 50; all except skin cancer (AESC; n=1,194) and all except skin and sex cancer (AESSC; n=867) (Table S1a). Survival was defined as the length of time (in years) between initial diagnosis and death. In this study, all-cause mortality is being considered. Survival data was organised into: full data set (i.e. those diagnosed aged 50+), and three age categories: individuals diagnosed between the ages of 50 and 64, diagnosed 65 and 79, and those diagnosed aged 80 plus. Herein, we focus on the larger data set, AESC; however, the results for AESSC are also described in the supplementary material. Unless otherwise stated, the results for the two data sets were similar, with the exception that the smaller AESSC data set tended to lose statistical power more quickly. In addition to cancer survival data, we also collected information about tumour grade, comorbidities, sex, cohort and familial relationships for each individual, in order to consider the impact that these factors have on survival (see Methods).

The FHS Affymetrix 50K Human Gene Focused microarray was employed in this study. There were 36,647 SNPs remaining after a quality filtering and linkage disequilibrium (LD) analysis (see Methods). The primary objective of our analysis was to identify polymorphisms and corresponding molecular mechanisms that affect survival in an ageing-related manner. Thus, we used the commonly implemented approach that focuses on SNPs that are in close proximity to known longevity associated genes (LAGs) [42,61–64]. A set of 316 putative longevity associated genes (LAGs) was assembled (Table S1d); and 880 SNPs located near/within 245 of the 316 longevity associated genes were extracted for analysis. The commonly used dominant genotypic model was employed, in which the heterozygous and variant homozygote genotypes were combined into a single risk group (the “variant”/ “Var” group) and compared to the common homozygote (the “wild type”/ “WT” group) [12,68,69].

Subsequently, we conducted a gene-level analysis by examining the 245 genes that the 880 SNPs that had been assigned to in the previous paragraph (Table S1e). A LAG could have (1) One SNP per gene, (2) Two SNPs per gene or (3) More than two SNPs per gene. For the “one SNP per gene” category, each individual was classified as possessing either the wild type or variant genotype at the SNP position. For the “two SNPs per gene” category, each individual was classified depending on the combination of wild type and variant alleles that they possessed at each SNP position; “WT\_WT” (i.e. wild type at first SNP position and wild

type at second SNP position) or “WT\_Var”, “Var\_WT” or “Var\_Var”. For the “more than two SNPs per gene” category, individuals were classed as having a low (<33%), medium (34-65%) or high (>66%) number of wild types. Each gene was assigned to an age-related KEGG pathway, where possible. Individuals were described as either having a low (<50%) or high (>50%) number of wild type SNPs per pathway. In total, 880 SNPs were assigned to 245 aging-related genes and 18 pathways.

To identify differences in survival patterns in different genotypes for a set of ageing-related variants between different diagnosis age brackets, we conducted a survival analysis, comprising (1) Kaplan Meier (KM) Estimator and (2) Cox Proportional Hazards Model. A Kaplan Meier Curve constructs a survival curve to compare the survival patterns of two or more groups of individuals, and a Log Rank Test is subsequently implemented to examine the null hypothesis that there is no difference between the populations in the probability of an event (in this case, death) occurring at a time point. A Kaplan Meier Curve and Log Rank Test cannot account for any other possible confounding factors that may affect survival, and so a Cox Proportional Hazard Regression Model subsequently modelled survival as a function of other variables, including genotype, sex, comorbidity status, cancer grade and cohort. Data sets were then divided by sex, cohort, and equal numbers of genotypes to examine the robustness of both the Kaplan Meier and Cox Model analyses. The reader is directed to the exact number of individuals used in each analysis at relevant times throughout the manuscript.

There is one SNP (rs11574359), two genes (*GPX4* and *SIRT3*) and one pathway (FoxO) that demonstrated age-related patterns of survival in cancer patients and will be discussed in turn. For each SNP, gene and pathway of interest, the results are laid out as a description of: (1) Kaplan Meier Curve, (2) Kaplan Meier Curve using sub-sets of the data (i.e. using equal numbers of genotypes, and dividing the data set by sex and cohort), (3) Cox Model and (4) Cox Model results once the model is adjusted for co-variables of interest rather than stratified, and once the data set is split by sex and cohort.

### **Rs11574358, a non-synonymous SNP in the WRN gene, has an age-dependent impact on mortality in cancer patients**

A Kaplan Meier analysis was conducted for each of the 880 SNPs assigned to LAGs (Table S2a for full output from Kaplan Meier analysis for all SNPs). Four SNPs consistently demonstrate significant (FDR < 0.05) survival differences between the wild type and variant

genotypes in different age categories (Table S2b): rs1794108, rs4147918, rs11574358 and rs317913. In all of these cases, possessing the wild type confers a longer survival time. In this section, we will focus on one SNP, rs11574358, the only SNP whose survival differences between genotypes in an age-dependent manner remained significant after the implementation of the Cox Model. However, it is interesting that the other SNPs significant from the Kaplan Meier analysis (i.e. rs1794108, rs4147918 and rs317913) are all located within well-known cancer and ageing related genes. For example, rs1794108 is a missense deleterious mutation in the proteasome 26S subunit, non-ATPase 13 (*PSMD13*) that is involved in cellular senescence [85], ageing [86] and with the onset of various cancers [87–89]. *ABCA7*, the gene that rs4147918 is located within, is on chromosome 19p13.3, the same chromosomal section as *APOE*, a gene that is well known to be associated with ageing and longevity related traits [21,90,91]. *ABCA7* has also been implicated recently in cancer progression; [92,93]; the SNP itself has also been previously associated with Alzheimer’s Disease [94]. Finally, rs317913 is located within the rat guanine nucleotide dissociation stimulator-like 3 gene (*RGL3*). Both the SNP [95,96] and the gene have previously been identified with cancer-related traits [97].

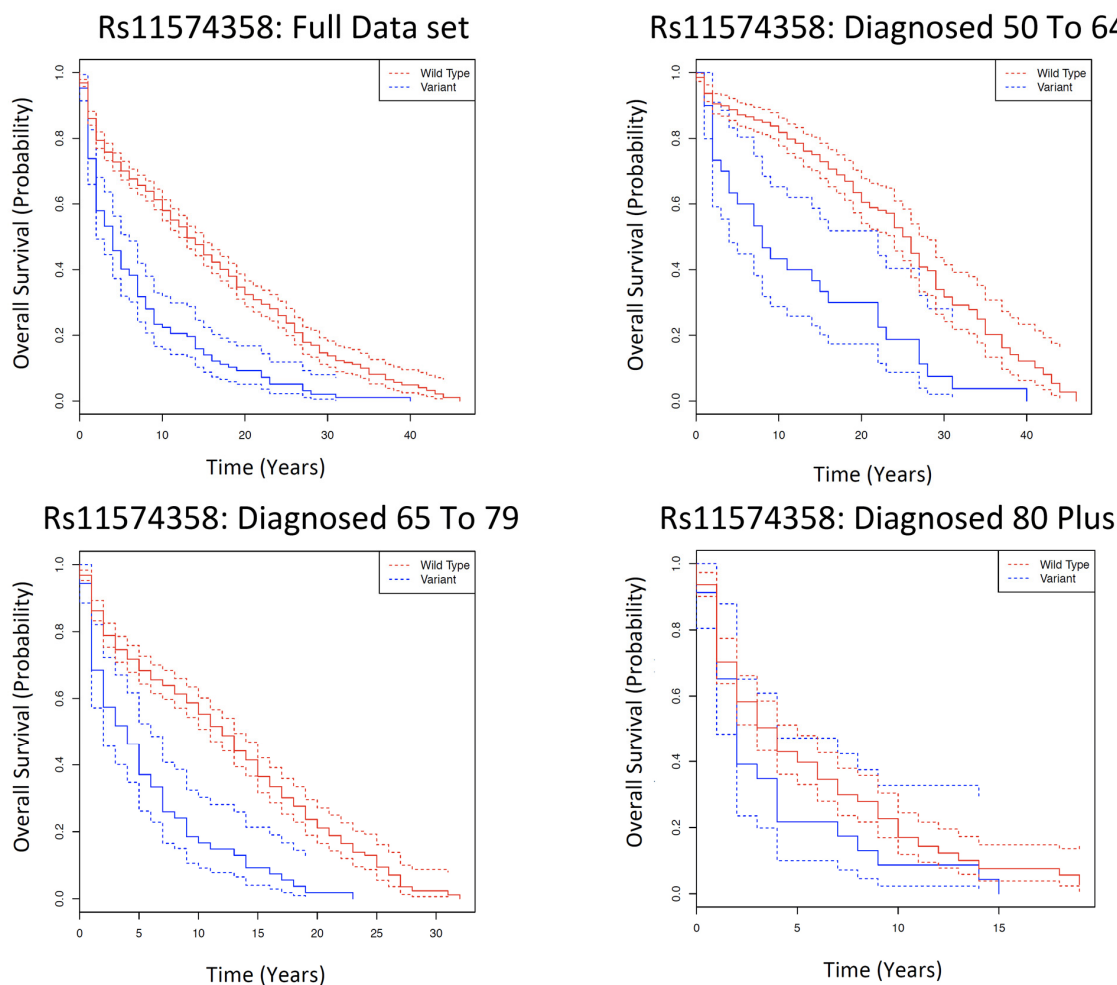
In the Kaplan Meier analysis; for rs11574358, in the full data set (i.e. those diagnosed 50+), the WT genotype patients have a significantly longer 5-year survival rate (SYSR) than those with the variant (Var) genotype (SYSR: WT: 70.1% (95% CI: 67.2-72.9%), Var: 40.1% (95% CI: 31.8-50.6%), FDR= 3.45E-12). Similar significant survival differences are observed in those diagnosed 50 to 64 (FDR= 3.0E-4) and diagnosed 65 to 79 (FDR= 1.16E-7), but not in the diagnosed 80 plus age bracket (FDR= 0.85) (Fig. 1, Table S2b).

A number of analysis repetitions were conducted to ensure the robustness of the observations; these are described in further detail in the Methods. First, in each of the age categories, an equal number of wild type and variant genotypes were randomly extracted, and the analysis was repeated, statistically significant survival patterns (FDR < 0.01) observed in all age groups except those individuals diagnosed over the age of 80 (Table S3a1). However, in this case, although the number of individuals is constant between the genotypes within each age category, there are still different numbers of individuals between the age categories. Thus, for the larger AESC data set, 20 individuals were subsequently elected at random for each genotype for all age categories, and the analysis was repeated; this ensures that the number of individuals remains constant per genotype both within and between all age categories. The Log Rank P Values are significant different in both

those diagnosed 50 to 64 ( $P=3.94E-03$ ) and those diagnosed 65 to 79 ( $P=2.41E-03$ ) (Table S3a). Dividing the AESC data set up by sex and cohort, statistically significant ( $FDR < 0.05$ ) survival differences are still observed in all the age categories except those diagnosed over 80 in the offspring (Table S3c) cohort, and in males (Table S3d) and females (Table S3e); and in the full data set and those diagnosed 65 to 79 in the original cohort (Table 3b). Thus, we suggest that the differences in survival patterns between the wild type and variant genotypes between the different age categories cannot be easily explained by the effects of cohort, sex or sample size.

The genotypic and phenotypic distribution of the data set used in the Cox Model is found in Table S3f. After the Cox Model, that accounted for the effects of sex,

cohort, comorbidities, tumour grade at diagnosis and familial relationships, different survival patterns are still observed in all of the age categories except those diagnosed over the age of 80. The hazard ratio of the risk allele in the full data set (i.e. diagnosed 50+) is 1.95 (95% CI: 1.47-2.58;  $FDR= 2.64E-05$ ) indicating that those individuals with the variant genotype (i.e. heterozygous or rare homozygote) have a 1.95-fold increased risk of death compared to those that possess the wild type (common homozygote). The HR increases to 4.65 (95% CI: 2.56-8.43;  $FDR= 6.01E-06$ ) in the diagnosed 50 to 64 age category, and decreases to 2.81 (95% CI: 1.77-4.46;  $FDR= 2.57E-04$ ) for those diagnosed 65 to 79 (Table 1, Table S3g). There are no significant survival differences in those individuals diagnosed  $>80$  ( $FDR= 0.98$ ). The Cox Model analysis was repeated using slightly varying data sets to ensure



**Figure 1.** Kaplan Meier survival estimates of overall cancer survival for rs11574358 among the Framingham Heart Study according to a dominant genotype model for different age categories, in which the wild type is the dominant homozygote, and the variant is the heterozygote and the minor homozygote. The full data set indicates all individuals diagnosed with cancer over the age of 50; and subsequently each age category is the individuals diagnosed with cancer in that particular age category. Solid lines indicate survival curve, dashed line indicates 95% confidence interval.



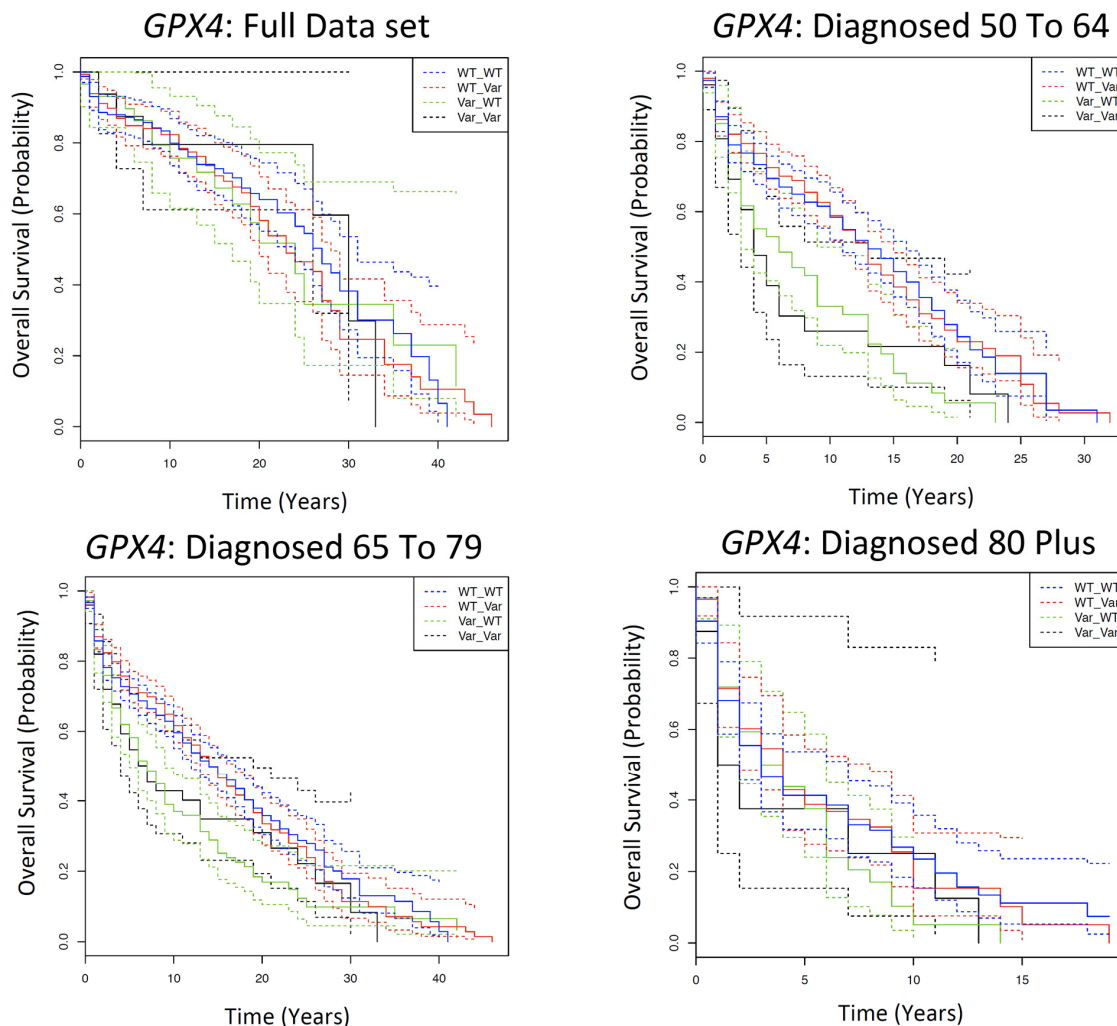
robustness of observations; as described in the Methods. First, the same significant age-dependent effects on survival as described are observed by adjusting the Cox model co-variables instead of stratifying (Table S3h). Second, significant or marginally significant survival differences in the same age categories are still observed once the data set is separated by sex (Females: Table S3i; Males: Table S3j) and offspring cohort (Table S3l) and in those diagnosed 65 to 79 in the original cohort (Table S3k), or if an equal number of individuals (AESC:  $N=176$ ; AESSC:  $N=119$ ) are randomly selected from each age category and the analysis is repeated (Table S8a).

Rs11574358 is a non-synonymous SNP in the Werner (*WRN*) gene converting a serine to an alanine. SIFT, a sequence homology-based tool that predicts deleterious

substitutions based on the degree of conservation of amino acid residues based on alignments of closely related sequences [98] predicts that this variant is deleterious to protein function.

### Rs4147918 in *GPX4*, and *SIRT3*, display patterns of age-dependent differences in mortality

Each of the 880 SNPs in the analysis was assigned to a gene  $\pm 60$ kb of the SNP. There are 60 genes with one SNP assigned to the gene, 54 genes with two SNPs assigned to the gene, and 131 genes with more than two SNPs assigned to the gene (Table S1f). The full set of Kaplan Meier results for the gene analysis is in Table S2c. There was one gene of interest with two SNPs per gene; glutathione peroxidase 4 (*GPX4*) that had two SNPs assigned to it: rs4147918 and rs757232. In the KM



**Figure 2.** Kaplan Meier survival estimates of overall survival for *GPX4* among the Framingham Heart Study according to a dominant genotype model, in which the wild type is the dominant homozygote, and the variant is the heterozygote and the minor homozygote. The full data set indicates all individuals diagnosed with cancer over the age of 50; and subsequently each age category is the individuals diagnosed with cancer in that particular age category. Solid lines indicate survival curve, dashed line indicates 95% confidence interval.

analysis, there are significantly different survival differences between the genotypes in both the full data set (FDR= 0.046) and in those diagnosed 65 to 79 (FDR= 0.013). In these two age categories, the average 5YSR for those with a WT allele in the first SNP position (i.e. a WT\_WT or WT\_Var genotype) is at least 17% higher than those with a variant allele in the first position (i.e. a Var\_WT or Var\_Var genotype; Fig. 2, Table S2d). There are no statistically significant differences in survival between genotypes for those diagnosed 50 to 64, or those diagnosed 80 plus (FDR=> 0.05). To examine the effects of sample size on the observations, 20 individuals were subsequently randomly selected from each of the genotypes per age category and the analysis was repeated (Table S4a). Similar to earlier observations, there are no statistically different survival patterns between the genotypes in any of the age categories except those diagnosed 65 to 79 (P

Val = 0.03). Similar observations were made once the data set is divided by sex (Males: Table S4b, Females: Table S4c) and in the original cohort once the data is divided by cohort (Table S4d) but not in the offspring cohort (Table S4e).

The genotypic and phenotypic distribution of the data set used in the Cox Model is in Table S4f. In the “diagnosed 65 to 79” age category, possessing the “Var\_WT” genotype led 1.85-fold increased risk of death compared to the “WT\_WT” genotype (95% CI=1.24-2.74; FDR = 0.02). In all other age categories of the AESC data set, there was no statistically significant difference in survival between genotypes (Table S4g, Table 1). A similar pattern is observed when variables are adjusted for rather than stratified in the Cox model (Table S4h) and pre-FDR significance is observed if the data set is divided by sex (Males: Table

**Table 1. Summary of the significant SNP, genes and pathways of interest after the Cox model analysis.**

Analysis	WT Allele (HR = 1.0)	Risk Allele	Gene	Age Category	N in KM analysis	HR	95% CI	P Value	FDR
<b>SNP Analysis</b>									
<b>rs11574358</b>	<b>TT</b>	<b>GT+GG</b>	<b>WRN</b>	<b>50+</b>	1,133	1.95	1.47-2.58	2.83E-06	2.64E-05
				<b>50-64</b>	378	4.65	2.56-8.43	6.40E-02	6.01E-06
				<b>65-79</b>	558	2.81	1.77-4.46	8.65E-01	2.57E-04
				<b>80+</b>	197	1.16	0.56-2.17	0.27	0.98
<b>Gene Analysis</b>									
<b>GPX4</b>	<b>WT1_WT 2</b>	<b>Var1_W T2</b>	<b>GPX4</b>	<b>50+</b>	1,035	1.25	0.94-1.47	0.12	0.48
				<b>50-64</b>	351	1.32	0.54-3.22	0.53	0.82
				<b>65-79</b>	504	1.85	1.85 (1.24- 2.74)	2.3E-3	0.02
				<b>80+</b>	180	1.40	0.72-2.69	0.32	0.74
<b>SIRT3</b>	<b>High WT</b>	<b># Low WT</b>	<b># SIRT3</b>	<b>50+</b>	1,037	1.49	1.06-2.03	0.02	0.09
				<b>50-64</b>	353	1.23	0.59-2.56	0.57	0.81
				<b>65-79</b>	501	1.98	1.21-3.21	0.007	0.06
				<b>80+</b>	183	1.02	0.51-2.01	0.96	0.96
<b>Pathway Analysis</b>									
<b>FoxO Pathway</b>	<b>High WT</b>	<b># Low WT</b>	<b># -</b>	<b>50+</b>	1,146	1.13	0.85-1.51	0.39	0.39
				<b>50-64</b>	382	1.31	0.72-2.38	0.37	0.39
				<b>65-79</b>	564	1.90	1.13-3.11	0.014	0.058
				<b>80+</b>	200	1.52	0.69-3.37	0.29	0.39

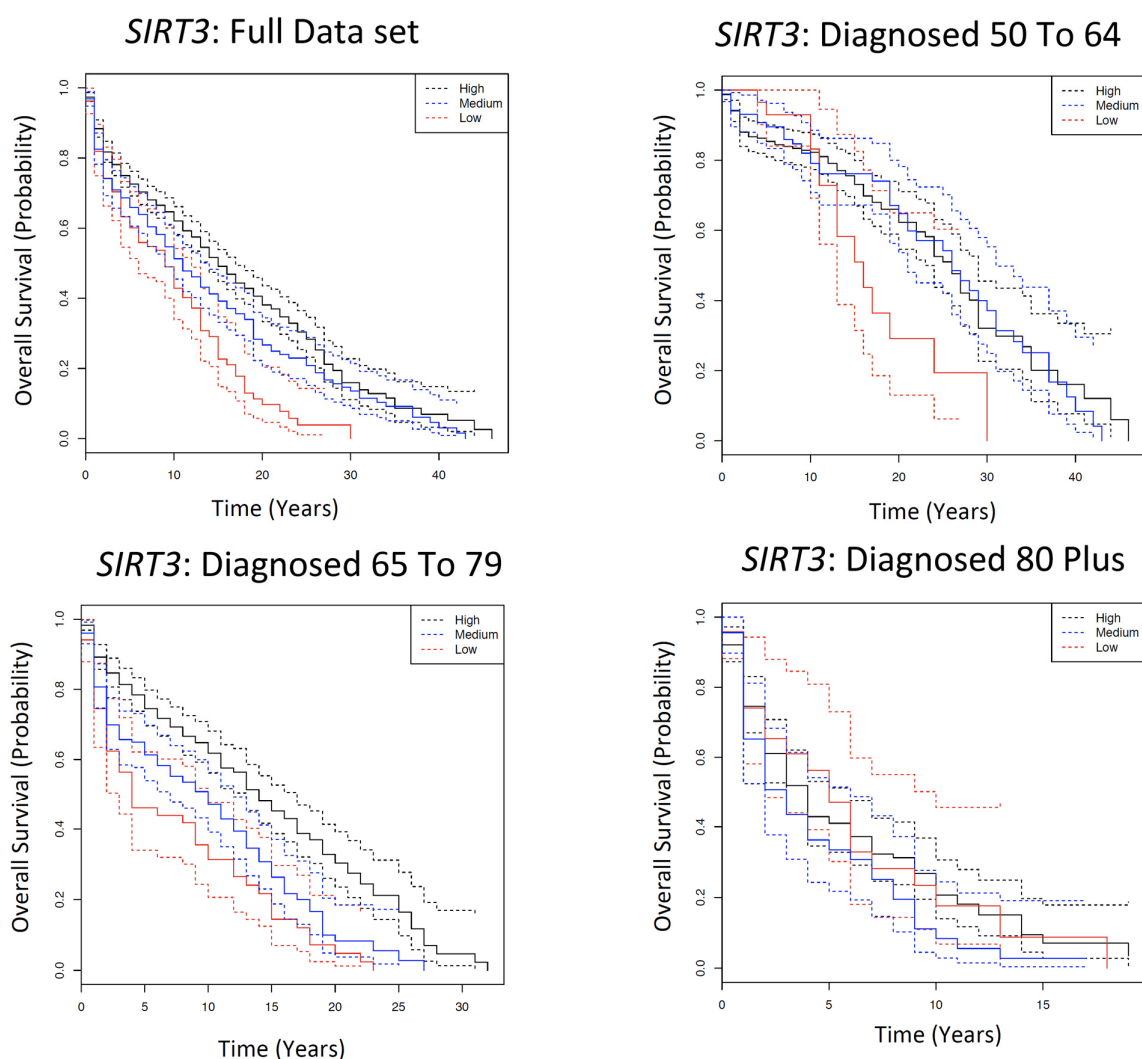
Column “N in KM Analysis” is the number of individuals in the original Kaplan Meier Curve. The number of individuals per genotype in the Cox Model is found in supplementary material. Column 5, “HR”, is the hazard ratio of the risk allele relative to the wild type. 95% CI is the 95% confidence interval of this HR. P value represents the significance of the HR and FDR is the FDR (Benjamini-Hochberg)-corrected P value.

S4i, Females: Table S4j) and in the Original cohort (Table S4k), but not the Offspring cohort (Table S4l). A similar observation of age-dependent survival differences between allele combinations with marginal significance is made in both the AESC (N=347) and the AESSC (N=287) data sets if an equal number of individuals is randomly selected from all age categories; Table S8b). Thus, the age dependent effects of this gene are not easily explainable by sample size, cohort, sex, familial relationship or tumour characteristic effects.

The data suggests that having a variant allele at the first position (rs4147918) and a wild type allele at the second position (rs757232) leads to an increased risk of death for patients with cancer in an age-dependent manner. SNPs were allowed to be  $\pm 60\text{kb}$  of a gene in

this data set. Thus, although the SNPs were technically assigned to *GPX4*, rs4147918 (i.e. the variant) is located in the nearby *ABCA7* gene, while rs757232 (the wild type) is in Histocompatibility (Minor) HA-1 gene (*HMHA1*). Rs757232 did not display significantly different survival patterns after the KM analysis in the individual SNP analysis (Table S2a). However, rs4147918 significantly affects survival in an age-dependent pattern in the SNP-level KM analysis (Full Data set FDR =  $1.54\text{E-}04$ , Diagnosed 65 to 79 =  $9.51\text{E-}06$ ; Table S2b) and is located within a known cancer and ageing related gene, as described in the previous section.

A second gene of interest is NAD-dependent deacetylase sirtuin-3, mitochondrial (*SIRT3*) that was



**Figure 3.** Kaplan Meier survival estimates of overall survival for *SIRT3* among the Framingham Heart Study according to a dominant genotype model, comparing patients with a high number of wild types to those with a low number of wild types. The full data set indicates all individuals diagnosed with cancer over the age of 50; and subsequently each age category is the individuals diagnosed with cancer in that particular age category. Solid lines indicate survival curve, dashed line indicates 95% confidence interval.

assigned four SNPs: rs11246007, rs11246020, rs1794108 and rs2280544. Each individual is characterised depending on the number of wild type alleles: “low” (one), “medium” (two) or “high” (three or four) number of wild type alleles. In *SIRT3*, possessing a high number of wild types led to at least 20% longer 5YSR in both the full data set (FDR=8.8E-07) and those diagnosed 65 to 79 (FDR=8.8E-07; Fig. 3, Table S2e).

Similar to previous sections, the effects of sample size were considered by randomly selecting 20 individuals from both wild type and variant genotypes for each of the age categories and repeating the analysis, thus ensuring that the number of individuals per genotype is constant both within and between all age categories. In agreement with our observations, there were no statistically significant survival differences in any of the age categories except marginal significance for those diagnosed between the age of 65 and 79 (Log Rank P Val = 0.09; Table S5a). Similar age-dependent survival differences between genotypes for this gene are observed if the data set was divided by cohort (Original Cohort: Table S5b, Offspring Cohort: Table S5c) or sex (Males: Table S5d, Females: Table S5e).

The genotypic and phenotypic distribution of the data set used in the Cox Model is found in Table S5e. After conducting a Cox Model analysis, individuals with a low number of wild types that were diagnosed 65 to 79 have a 1.98-fold (95% CI=1.21-3.21; FDR= 0.06) increased risk of death compared to those with a high number of wild type alleles (Table S5g). A similar observation are consistently observed if one adjusts for the co-variables rather than stratifies (Table S5h), and similar significant or pre-FDR significant or marginally significant (in the case of original cohort) pattern is observed once the data is split by cohort or sex (Tables S5i-S5l). Similar patterns are found with marginal significance in the AESSC data set if an equal number of individuals ( $N= 119$ ) is selected from each age category and the analysis is repeated; and in both AESC ( $N=347$ ) and AESSC ( $N=287$ ) data sets if equal numbers of individuals are randomly selected from the full data set, those diagnosed 50 to 64 and those diagnosed 65 to 79 age categories (Table S8).

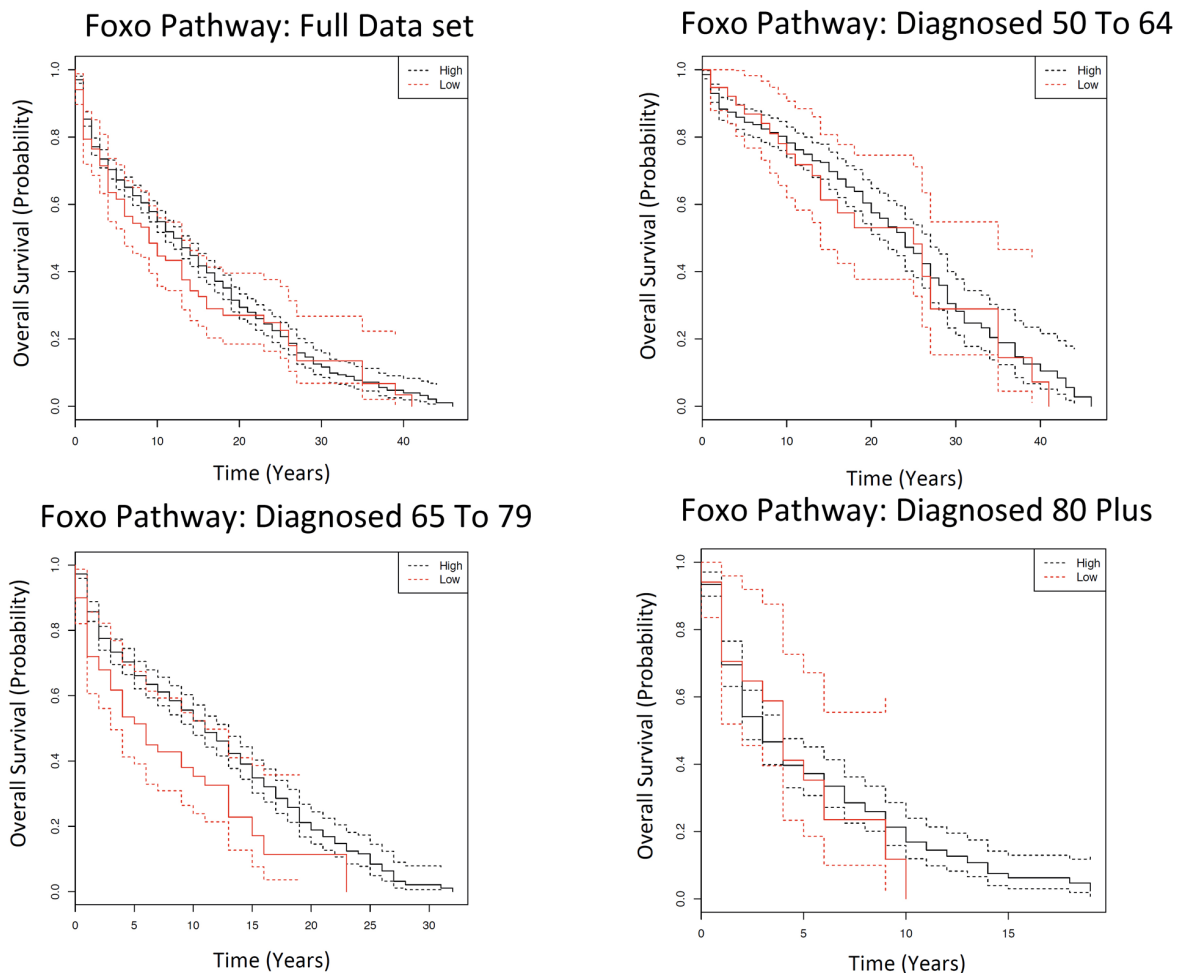
SNPs were allowed to be  $\pm 60$ kb of a gene in this data set. Rs11246007 is an intron variant in *SIRT3*. Rs11246020 is a missense mutation in *SIRT3* that converts a valine to an isoleucine. Rs2280544 is a UTR Variant 3' SNP in the BET1 Golgi Vesicular Membrane Trafficking Protein Like (*BET1L*) gene. Interestingly, rs1794108 is one of four SNPs that consistently demonstrated survival differences in the initial Kaplan Meier SNP analysis, as described earlier (Table S2a).

## SNPs in the FoxO pathway display age-dependent patterns of mortality

Six of the eighteen putative ageing-related pathways (Rap1 signalling, FoxO signalling, Cell Cycle, p53 signalling, Fc epsilon signalling and TNF signalling) displayed marginally significant (FDR < 0.10) survival differences in the KM analysis between two genotype groups (i.e. having a low and high number of wild types) in different age groups (Table S2f). Unsurprisingly, these pathways are known to be involved in cancer-related processes; e.g. p53 signalling, Cell Cycle and Rap1 signalling (Table S2f). There is one pathway of interest after the Cox Model that will be discussed in detail: FoxO signalling. There 29 genes (Table S6a) and 108 SNPs (Table S6b) assigned to the FoxO pathway. In the KM analysis, there is significantly different survival patterns observed in those diagnosed 65 to 79 (FDR=0.058; Fig. 4, Table S2g). In this case, having a high number of wild type alleles confers a protective effect on survival. Unlike all of the other SNPs and genes, the same pattern is not replicated once equal numbers individuals per genotype per age category are extracted and the Kaplan Meier analysis is repeated (Table S6c). In the diagnosed 65 to 79 age category, significantly different survival patterns between genotypes are observed in the original cohort (Table S6d) and in females (Table S6g) but not in the offspring cohort (Table S6e) or in males (Table S6f).

The genotypic and phenotypic distribution of the data set used in the Cox Model is in Table S6h. After the Cox model, possessing a low number of wild types leads to 1.90 (95% CI: 1.13-3.11; FDR=0.058) increased risk of death compared to possessing a high number of wild types in those diagnosed 65 to 79 (Table S6i; Table 1). Subsequently, a similarly significant increased hazard ratio is observed for those with a low number of wild type alleles in the diagnosed 65 to 79 age category if the Cox model is adjusted for co-variables rather than stratified (Table S6j), and similar pre-FDR significance is observed if the analysis is conducted on each sex separately (Table S6k, S6l) and in original cohort (Table S6m), but not in the offspring cohort (Table S6n). Post-FDR significantly different survival is also observed in the AESC ( $N=349$ ) and AESSC ( $N=139$ ) data sets if an equal number of individuals are selected from each age category and the analysis is repeated (Table S8d). The three most common types of the effects for the SNPs in the FoxO pathway are intron variants (38%), non-coding transcript variants (17%) and mis-sense variants (11%).





**Figure 4.** Kaplan Meier survival estimates of overall survival for FoxO pathway among the Framingham Heart Study according to a dominant genotype model, comparing patients with a high number of wild types to those with a low number of wild types. The full data set indicates all individuals diagnosed with cancer over the age of 50; and subsequently each age category is the individuals diagnosed with cancer in that particular age category. Solid lines indicate survival curve, dashed line indicates 95% confidence interval.

### Replication of results in an independent population

A replication population was selected based on the combination of two cohorts; Atherosclerosis Risk in Communities Study and Cardiovascular Health Study; the assembly of these data sets is described in detail in the methods. We attempted to replicate our results in this combined replication population; however the Framingham Heart Study is quite unique; both in terms of the large amount of data collected over a long period of time and the SNPs that were genotyped on its unique custom Human Gene Focused Affy 50K array. Unfortunately, although this replication population has been invaluable in the past in investigating various traits associated with ageing and longevity (for example, [12,99]), there was a low overlap in the SNPs (16 SNPs) out of all of the SNPs of interest in the SNP/gene/path-

way analysis in this study and the ARIC/CHS arrays. Only two SNPs of interest could be considered further in our analysis; however, rs11574358; a key SNP of interest in our analysis had a minor allele frequency of 0 in both the ARIC and CHS data sets and so could not be considered further. Thus, the replication study purely focussed on rs4147918 (MAF of 0.036 and 0.044 in CHS and ARIC, respectively).

For rs4147918, 92% (N=1,808) of the data set was assigned as wild type (i.e. homozygous dominant for this polymorphism), while 8% (N=140) were considered as variants. The cancer phenotype available for this study in ARIC/CHS included cancers at all sites because there was no information on cancer without skin in ARIC. Comorbidity phenotype was defined as the score counting presence of the following diseases in

an individual (each disease was coded as 1; no disease was coded as 0): heart failure, diabetes, stroke, and/or myocardial infarction. A Kaplan Meier analysis indicated that there were significantly different survival patterns only in the Full Data set ( $P=0.03$ ) and not in any of the specific age categories, although the difference was marginally significant in those diagnosed 65 to 79 ( $P=0.06$ ; Table S7). Thus, there were no significant results of note after the Cox Model and the relationship between survival between the different age categories could not be considered using this population (Table S7).

## DISCUSSION

The investigation described herein is the first systematic study to address how the ageing process impacts the effect that single nucleotide polymorphisms have on cancer survival; a field that could greatly affect the possibility of individualising cancer prognoses and treatments in the post-genomic era. There were two SNPs (rs11574358 and rs4147918), one gene (*SIRT3*) and one pathway (FoxO signalling) that may be of interest for further consideration.

Rs11574358 is a non-synonymous SNP in the Werner (*WRN*) gene. *WRN* is responsible for the progeroid Werner Syndrome, characterised by the accelerated appearance of features associated with ageing [100–102]. This syndrome is well established to be associated with an elevated risk of cancer [103,104]. Some of the most common co-occurring chronic co-morbidities among cancer patients include ischemic heart disease, hypertension and hyperlipidemia [58]. Rs11574358 was recently identified to be associated with traits related to ageing, including cardiovascular disease prevalence, systolic blood pressure, cancer prevalence, total cholesterol and cystatin C in serum (chronic kidney disease) [42,105]. Since we know that rs11574358 is associated with traits such as cardiovascular disease prevalence and systolic blood pressure and the polymorphism converts a serine to an alanine residue, and has been suggested to abolish the phosphoserine structure of the protein and potentially affect protein function [106]; it is possible that this SNP may exert an age-dependent effect on mortality by contributing to the effect or severity of the most common cancer co-morbidities. Alternative hypotheses could be that the polymorphism may be associated with cancer aggressiveness, tissue susceptibility to cancer invasion or to changes in other ageing processes that contribute to mortality.

The second SNP of interest was rs4147918; this SNP was assigned to the *GPX4* LAG in this analysis, but as SNPs were allowed to be  $\pm 60$ kb of a LAG, the SNP is

actually located within the nearby *ABCA7* (Entrez ID: 10347), a member of the superfamily of ATP-binding cassette transporters that transport various molecules across extra- and intra-cellular membranes. Rs4147918 induces a glutamine-to-arginine change in an exon of the *ABCA7* gene. Although this SNP is considered tolerated (i.e. not considered deleterious to function), even a variant that can be tolerated, or a synonymous change that appears not to affect amino acid selection at all, could still affect protein function and disease susceptibility, particularly when combined with the ageing process [107–110]. *ABCA7* is on chromosome 19p13.3, the same chromosomal section as *APOE*, a gene that is well known to be associated with ageing and longevity related traits [12,21,91,111,112]. In *ABCA7*, seven SNPs (including rs4147918) are cholesterol-related and showed a significant association with Late Onset Alzheimer's Disease [21], although little else is reported on the clinical effects caused by this polymorphism. Thus, it is possible that similar to rs11574358, if some of the most common chronic co-morbidities among cancer patients include ischemic heart disease, hypertension and hyperlipidemia [58]; and rs4147918 is known to be cholesterol-related and associated with the ageing process, then it is possible that this polymorphism may exert an effect on age-dependent mortality by changing the severity of particular co-morbidities.

There was one gene of interest for further analysis; *SIRT3*, with four SNPs (although the SNPs are in *SIRT3*, and two closely-related genes *PSMD13* and *BET1L*). Interestingly, *SIRT3* is known to be involved in both the cancer and ageing processes [113–115]. In addition, *PSMD13* and *SIRT3* share a promoter [116], and *PSMD13*, a proteasome subunit, is involved in the degradation of abnormal proteins, cellular senescence [85] and ageing [86] and variants in this gene have previously been associated with the onset of various cancers [87–89]. Thus, it is plausible that having a high number of variants in closely related genes could impact cancer progression in an age-dependent manner.

There was one pathway of interest for potential further investigation; FoxO signalling. The ability of FoxO factors to induce cell cycle arrest, DNA repair and apoptosis makes them attractive candidates as tumour suppressors [117,118]. In addition, members of the FoxO pathway are known to affect ageing and longevity [117–119]. Thus, since having a high number of variants appears to negatively affect cancer survival and the pathway is known to play an important role in the aetiology of cancer and the ageing process, perhaps a high level of variation in this pathway affects mortality in an age-dependent manner; which should be considered further.

There are a number of limitations to this investigation. First, it was not possible to identify an ideal replication population, as the Framingham Human Gene Focused Affy 50K array is a unique custom array. We did not consider cancer-specific deaths in this instance; due to lack of data (In the AESC full data set, there were 663 causes of deaths recorded, 47 of these had cancer as a primary cause of death). Although all-cause mortality is a common end-point for many successful survival analyses [18,49,50,55], we acknowledge that the absence of data on cancer-specific deaths versus all other causes of death limits our ability to interpret our observations, and leads to more questions. It would be interesting in the future to do similar investigations with cancer-specific death data available, in order to tease apart the specific contribution that variants such as those described in this research are contributing to specifically cancer survival.

As with all polymorphism-trait studies, it is possible that our SNPs of interest do not exert effects on mortality themselves, but are in high linkage disequilibrium with the ungenotyped SNPs of interest. Some of the effects observed in this analysis would appear to be quite modest; for example, the gene and pathway analyses exhibit modest effects (~2 fold increased risk of death in particular age groups). However, similarly-sized effects have been reported in other analyses which have provided fascinating insight into the impact of genetic variants on various survival patterns [12,18,74,120,121]. In addition, given the complex nature of longevity and ageing as traits and the known difficulties in identifying SNPs and genes associated with these traits in even much larger human studies that do not account for the ageing process [122], the identification of SNPs that display even modest effects on cancer survival differently with age warrant further attention. Finally, it is not clear why age-dependent survival effects are generally observed in those diagnosed 65 to 79, rarely in those diagnosed 50 to 64 and never in those diagnosed over the age of 80. To eliminate potential bias caused by unequal sample size, the Kaplan Meier curves were repeated, using the same number of individuals per genotype between and within age categories, and the Cox Model analyses were divided by sex and cohort. One potential reason for a lack of observation in diagnosed >80 group could be that the overall mortality rate in the older population is such that more individuals die of unrelated causes, and so die with cancer but not due to cancer. However, this does not explain the lack of observation in the diagnosed 50 to 64 group. In addition, Kulminski et al. (2014) similarly demonstrated that there is a significant adverse effect of the e4 allele on survival that is limited to women with a moderate lifespan (70-95 years); i.e. an effect was also not observed in the young (<70 years)

or extremely old (>95 years) age categories. They suggested that possible reasons for these observations could include the buffering mechanisms by other genes [123] and/or the environmental modulations of genetic effects [124]. Therefore, an interesting open question arising from this work is why some variants affect survival in certain age groups, but not in others.

In summary, this investigation suggests that the *APOE* variant identified by Kulminski et al. [12] may not be the sole variant that affects cancer survival in an age-dependent manner. This study is the first exploratory systematic investigation to identify SNPs, genes and pathways that differentially affect mortality depending on the age of diagnosis, whose findings need to be independently validated in a suitable population, once such a population arises in the future. If corroborated, such information would provide potential targets for further exploration as prognostic biomarkers and individualised therapies in the post-genomic era. Given that we live in a greying population, and the majority of tumours are diagnosed in aged patients, such knowledge will be an invaluable tool advancing the field of geriatric oncology.

## MATERIALS AND METHODS

### Study population: the Framingham Heart Study

The Framingham Heart Study (FHS) [36] comprises 5,209 respondents aged 28-62 at baseline who have been biennially examined for almost 60 years. The Framingham Heart Study Offspring (FHSO) respondents (N=5,124) aged 5-70 at baseline were biological descendants (N=3,514), their spouses (N=1,576) and adopted offspring (N=34) of the FHS participants who were examined about every four years at nine visits. The study design has been detailed previously [39,40]. Phenotypic data was collected through FHS clinic examinations, hospital admission surveillances and monitoring death registries. Biospecimens were mostly collected in the late 1980s and through the 1990s from surviving participants [12]. The FHS data are available from the NIH SHARe through dbGaP [41] (accession number phs000007.v29.p10).

### Assembly of cancer and comorbidity phenotypic data set

As described in the Results, two cancer data sets were assembled based on tumour topography for all patients that were first diagnosed with cancer over the age of 50 and were not diagnosed solely via death cert: (1) All cancer except skin cancer (AESC; n=1,194) and (2) All cancer except skin and sex cancer (AESSC; n=867).

Skin cancer is commonly not considered in similar analyses (for example, [12,42]) due to accurate diagnosis difficulties that may affect survival times [43–48]. Sex-related cancers were subsequently also removed to examine survival differences from cancers that are common to both sexes, as conducted in [12]. All-cause mortality is being considered in this analysis, this has been a common end-point for many successful survival analyses [17,49,50]. The data was right-censored for the Kaplan Meier analysis. If date of death was not available, the date of last contact was used.

Factors that may affect survival were considered. Comorbidities (i.e. additional diseases that co-occur with a primary disease of interest) can affect cancer diagnosis [51], treatment [52–54] and prognosis [55–57]. The co-morbidities included in our analysis were: cerebrovascular disease, diabetes, congestive heart failure, dementia, myocardial infarction and subsequent cancer diagnoses; some of which are the most common co-morbidities in cancer [2,58,59]. Each individual was classified as having a low (0-2), medium (3-4) or high (5-6) number of co-morbidities. A tumour grade (1-4) was also assigned to each tumour.

#### **Assembly and quality filtering of the genetic data set**

The FHS Affymetrix 50K Human Gene Focused microarray was employed in this study. Quality control filtering was conducted as in [42] using PLINK v. 1.9 [60]: SNPs were removed if they exhibited Hardy Weinberg  $P$  values  $<10^{-2}$ ,  $>10\%$  missingness,  $>2\%$  Mendel errors,  $<2\%$  Minor Allele Frequency (MAF) or were located on a sex chromosome. Using PLINK, a LD analysis was also conducted between SNPs by calculating pair-wise  $r^2$  statistics for founders only, SNPs in high LD (i.e.  $r^2 > 0.9$ ) were removed.

A set of 298 human LAGs (<http://genomics.senescence.info/genes/human.html>; Table S1b) and 83 human homolog to mouse genes (<http://genomics.senescence.info/genes/models.html>; Table S1c) were obtained from GenAge Build 17 [65]. Once the two data sets (i.e. human, and human homolog to mouse) were combined and redundant genes were removed, 316 putative LAGs remained (Table S1d). The location of each gene was retrieved from Ensembl [66], and SNPs located within  $\pm 60\text{kb}$  of each putative LAG were extracted, similar to [42,67]. 880 SNPs were putatively identified as ageing-related using this method, covering 245 of the 316 genes in the data set (Table S1e). SNPs were subsequently assigned to the 245 LAGs; as described in the Results, a LAG could have one, two or more than two SNPs assigned to it. Third, SNPs were assigned to LAG pathways. A set of 69 pathways considered related to ageing process given their known involvement in

ageing related processes were extracted from the KEGG database v. 74 [70] and 18 pathways that had at least 10 genes and 10% overlap between the pathway genes and the set of longevity associated genes in this analysis were extracted [70] (Table S1g).

#### **Statistical analyses: the Kaplan Meier Estimator and Cox Proportional Hazards Model**

A KM curve was constructed for each SNP/gene/pathway of interest using the “survival” package v. 2.38 [71] in R v. 3.2.2 [72]. A Log Rank Test [73] subsequently examined the null hypothesis that there is no difference between the populations in the probability of an event (in this case, death) occurring at a time point. Robustness of the constructed Kaplan Meier curves were examined by repeating the analysis twice: (1) Using equal number of wild type and variant genotypes and (2) Once the data set was divided by cohort (i.e. original and offspring cohorts) and sex (i.e. males and females).

The KM Estimator and Log Rank Test do not allow other explanatory variables to be considered when estimating the survival differences between two groups. Thus, for SNPs, genes and pathways that displayed initial survival differences using the KM approach, a Cox Proportional Hazard Regression Model modelled survival as a function of genotype, sex, comorbidity status, cancer grade and cohort, implemented in the “coxme” package v. 2.2.3 [71] in R v. 3.2.2 [72]. Schoenfeld residuals were used to test whether the proportional hazards assumption of the model were met, and only models with  $P$  values  $>0.05$  were considered valid. As a result of non-proportionality, the models were stratified by sex, comorbidity status, cancer grade and cohort, similar to what has been conducted in other studies [74–76]. Genotypes were treated as categorical variables. In addition, a multivariate kinship frailty model implemented in the kinship package (<http://cran.r-project.org/web/packages/kinship/>) in R was incorporated to account for any familial relationships within the FHS [39,77]. Missing grade and genotypic information was imputed with the widely used multiple imputation method [78–81], implemented in MICE v. 2.25 [82]. For the Cox Model, in the SNP analysis, the hazard ratio (HR) describes the increase (or decrease) risk of death in the Var allele, compared to the WT allele. In the gene-level analysis, the risk of each genotype group is being compared to either the WT\_WT allele (for those genes with two SNPs) or possessing a high number of WT alleles (for those genes with  $>2$  SNPs). In the pathway analysis, the risk of each genotype group is being compared to those individuals with a high number of WT alleles.



To examine the robustness of the Cox Model, we repeated the analysis twice: (1) By adjusting the model for all of the factors instead of stratification (i.e. cohort, sex, grade, co-morbidities and, for the SNP analysis, other SNPs of interest from the Kaplan Meier analysis) and (2) Once the data was split by sex (i.e. male and female) and cohort (i.e. original and offspring), and using an equal number of individuals per genotype and age category. A Benjamini-Hochberg False Discovery Rate (FDR) was calculated across the AESC and AESSC data sets, all age brackets and all SNPs, genes or pathways of interest. In addition, once the data was divided based on cohort (i.e. original and offspring) or sex, the FDR was conducted across all conditions (i.e. male, female, original and offspring). A FDR <0.05 was considered statistically significant.

### Replication analysis in an independent population

A replication population was selected based on the combination of two cohorts. In the Atherosclerosis Risk in Communities Study (ARIC) [83], the study participants (aged 45-64 at baseline in 1987) were randomly selected and recruited at four field centres across the U.S. We used data from four available examinations. Genotyping in 12,771 ARIC participants (N=9,633 Caucasians) was conducted using Affymetrix 6.0 array (1,000K SNPs). For the Cardiovascular Health Study (CHS) [84], the main cohort of the CHS participants (N=5,201 Caucasians) aged 65+ years at baseline in 1989 was examined annually through 1999. The CHS clinic exams ended in June 1999. After June 1999 two phone calls per year to participants collected information on incidence of diseases and death. Deaths also were ascertained through surveillance and at semi-annual contacts. SNPs for the present study were selected from the Candidate Gene Association Resource (CARE) that included records for 5,531 CHS participants. This investigation combined the two studies to carry out our analyses, restricting to participants of European descent.

### Availability of data and materials

The data generated during this study (for example survival study results) is available in the supplementary material. The Framingham SHARe data used for the analyses described in this manuscript were obtained through dbGaP (accession number phs000007.v29.p10). The URL for this study is: [https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study\\_id=phs000007.v29.p10&phv=159482&phd=1105&pha=4313&pht=1415&phvf=&phdf=&phaf=&phtf=&dssp=1&consent=&temp=1](https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000007.v29.p10&phv=159482&phd=1105&pha=4313&pht=1415&phvf=&phdf=&phaf=&phtf=&dssp=1&consent=&temp=1).

### Ethics approval and consent to participate

FHS data was obtained from the Framingham Heart Study of the National Heart Lung and Blood Institute of the National Institutes of Health and Boston University School of Medicine. This work was supported by the National Heart, Lung and Blood Institute's Framingham Heart Study (Contract No.N01-HC-25195). For this type of study, formal consent is not required.

### AUTHOR CONTRIBUTIONS

JPdM conceived experiment. AD performed analysis; YK performed analysis in replication study. AMK supervised analysis in replication study. AD drafted manuscript. All authors edited and finalized manuscript.

### ACKNOWLEDGEMENTS

The authors thank members of the Integrative Genomics of Ageing Group for useful discussion. This work was supported by the use of data from the National Heart, Lung and Blood Institute's Framingham Heart Study (Contract No. HHSN268201500001I).

### CONFLICTS OF INTEREST

The authors declare no competing interests.

### FUNDING

This project was supported by a European Molecular Biology Organization (EMBO) fellowship to A.D. A.D.'s work was also supported by a Wellcome Trust grant (104978/Z/14/Z) to J.P.M. AMK and YK acknowledge support by Grants No P01 AG043352 and R01AG047310 from the National Institute on Aging to AMK. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

### REFERENCES

1. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA Cancer J Clin.* 2015; 65:87–108. <https://doi.org/10.3322/caac.21262>
2. Edwards BK, Noone AM, Mariotto AB, Simard EP, Boscoe FP, Henley SJ, Jemal A, Cho H, Anderson RN, Kohler BA, Ehemann CR, Ward EM. Annual Report to the Nation on the status of cancer, 1975-2010, featuring prevalence of comorbidity and impact on survival among persons with lung, colorectal, breast, or prostate cancer. *Cancer.* 2014; 120:1290–314. <https://doi.org/10.1002/cncr.28509>

3. de Magalhães JP. How ageing processes influence cancer. *Nat Rev Cancer*. 2013; 13:357–65. <https://doi.org/10.1038/nrc3497>
4. Meropol NJ, Schulman KA. Cost of cancer care: issues and implications. *J Clin Oncol*. 2007; 25:180–86. <https://doi.org/10.1200/JCO.2006.09.6081>
5. Luengo-Fernandez R, Leal J, Gray A, Sullivan R. Economic burden of cancer across the European Union: a population-based cost analysis. *Lancet Oncol*. 2013; 14:1165–74. [https://doi.org/10.1016/S1470-2045\(13\)70442-X](https://doi.org/10.1016/S1470-2045(13)70442-X)
6. International HapMap Consortium. A haplotype map of the human genome. *Nature*. 2005; 437:1299–320. <https://doi.org/10.1038/nature04226>
7. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res*. 2001; 29:308–11. <https://doi.org/10.1093/nar/29.1.308>
8. van Heemst D, Mooijaart SP, Beekman M, Schreuder J, de Craen AJ, Brandt BW, Slagboom PE, Westendorp RG, and Long Life study group. Variation in the human TP53 gene affects old age survival and cancer mortality. *Exp Gerontol*. 2005; 40:11–15. <https://doi.org/10.1016/j.exger.2004.10.001>
9. Atzmon G, Rincon M, Schechter CB, Shuldiner AR, Lipton RB, Bergman A, Barzilai N. Lipoprotein genotype and conserved pathway for exceptional longevity in humans. *PLoS Biol. Public Library of Science*. 2006; 4:e113.
10. Bonafè M, Olivieri F. Genetic polymorphism in long-lived people: cues for the presence of an insulin/IGF-pathway-dependent network affecting human longevity. *Mol Cell Endocrinol*. 2009; 299:118–23. <https://doi.org/10.1016/j.mce.2008.10.038>
11. Yashin AI, Arbeev KG, Wu D, Arbeeve LS, Kulminski A, Akushevich I, Culminkaya I, Stallard E, Ukraintseva SV. How lifespan associated genes modulate aging changes: lessons from analysis of longitudinal data. *Front Genet*. 2013; 4:3. <https://doi.org/10.3389/fgene.2013.00003>
12. Kulminski AM, Arbeev KG, Culminkaya I, Arbeeve L, Ukraintseva SV, Stallard E, Christensen K, Schupf N, Province MA, Yashin AI. Age, gender, and cancer but not neurodegenerative and cardiovascular diseases strongly modulate systemic effect of the Apolipoprotein E4 allele on lifespan. *PLoS Genet*. 2014; 10:e1004141. <https://doi.org/10.1371/journal.pgen.1004141>
13. Frazer KA, Murray SS, Schork NJ, Topol EJ. Human genetic variation and its contribution to complex traits. *Nat Rev Genet*. 2009; 10:241–51. <https://doi.org/10.1038/nrg2554>
14. Budovsky A, Craig T, Wang J, Tacutu R, Csordas A, Lourenço J, Fraifeld VE, de Magalhães JP. LongevityMap: a database of human genetic variants associated with longevity. *Trends Genet*. 2013; 29:559–60. <https://doi.org/10.1016/j.tig.2013.08.003>
15. Atanasovska B, Kumar V, Fu J, Wijmenga C, Hofker MH. GWAS as a Driver of Gene Discovery in Cardiometabolic Diseases. *Trends Endocrinol Metab*. 2015; 26:722–32. <https://doi.org/10.1016/j.tem.2015.10.004>
16. Welter D, MacArthur J, Morales J, Burdett T, Hall P, Junkins H, Klemm A, Flicek P, Manolio T, Hindorf L, Parkinson H. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res*. 2014; 42:D1001–06. <https://doi.org/10.1093/nar/gkt1229>
17. Bijl MJ, van Schaik RH, Lammers LA, Hofman A, Vulto AG, van Gelder T, Stricker BH, Visser LE. The CYP2D6\*4 polymorphism affects breast cancer survival in tamoxifen users. *Breast Cancer Res Treat*. 2009; 118:125–30. <https://doi.org/10.1007/s10549-008-0272-2>
18. Van Blarigan EL, Ma J, Kenfield SA, Stampfer MJ, Sesso HD, Giovannucci EL, Witte JS, Erdman JW Jr, Chan JM, Penney KL. Plasma antioxidants, genetic variation in SOD2, CAT, GPX1, GPX4, and prostate cancer survival. *Cancer Epidemiol Biomarkers Prev*. 2014; 23:1037–46. <https://doi.org/10.1158/1055-9965.EPI-13-0670>
19. Ertekin-Taner N. Genetics of Alzheimer disease in the pre- and post-GWAS era. *Alzheimers Res Ther*. 2010; 2:3. <https://doi.org/10.1186/alzrt26>
20. Jeck WR, Siebold AP, Sharpless NE. Review: a meta-analysis of GWAS and age-associated diseases. *Aging Cell*. 2012; 11:727–31. <https://doi.org/10.1111/j.1474-9726.2012.00871.x>
21. Bertram L, Tanzi RE. Genome-wide association studies in Alzheimer's disease. *Hum Mol Genet*. 2009; 18:R137–45. <https://doi.org/10.1093/hmg/ddp406>
22. Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet*. 2005; 6:95–108. <https://doi.org/10.1038/nrg1521>
23. Burke SN, Barnes CA. Neural plasticity in the ageing brain. *Nat Rev Neurosci*. 2006; 7:30–40. <https://doi.org/10.1038/nrn1809>
24. Vasto S, Candore G, Balistreri CR, Caruso M, Colonna-Romano G, Grimaldi MP, Listi F, Nuzzo D, Lio D, Caruso C. Inflammatory networks in ageing, age-related diseases and longevity. *Mech Ageing Dev*.

- 2007; 128:83–91.  
<https://doi.org/10.1016/j.mad.2006.11.015>
25. Webster RG. Immunity to influenza in the elderly. *Vaccine*. 2000; 18:1686–89.  
[https://doi.org/10.1016/S0264-410X\(99\)00507-1](https://doi.org/10.1016/S0264-410X(99)00507-1)
  26. Effros RB. Role of T lymphocyte replicative senescence in vaccine efficacy. *Vaccine*. 2007; 25:599–604.  
<https://doi.org/10.1016/j.vaccine.2006.08.032>
  27. Chahal HS, Drake WM. The endocrine system and ageing. *J Pathol*. 2007; 211:173–80.  
<https://doi.org/10.1002/path.2110>
  28. Barzilai N, Huffman DM, Muzumdar RH, Bartke A. The critical role of metabolic pathways in aging. *Diabetes*. 2012; 61:1315–22. <https://doi.org/10.2337/db11-1300>
  29. Ferrari AU, Radaelli A, Centola M. Invited review: aging and the cardiovascular system. *J Appl Physiol* (1985). 2003; 95:2591–97.  
<https://doi.org/10.1152/jappphysiol.00601.2003>
  30. Parkin DM. Global cancer statistics in the year 2000. *Lancet Oncol*. 2001; 2:533–43.  
[https://doi.org/10.1016/S1470-2045\(01\)00486-7](https://doi.org/10.1016/S1470-2045(01)00486-7)
  31. Anisimov VN. Carcinogenesis and aging 20 years after: escaping horizon. *Mech Ageing Dev*. 2009; 130:105–21.  
<https://doi.org/10.1016/j.mad.2008.02.004>
  32. Federico M, Vitolo U, Zinzani PL, Chisesi T, Clò V, Bellesi G, Magagnoli M, Liberati M, Boccomini C, Niscola P, Pavone V, Cuneo A, Santini G, et al. Prognosis of follicular lymphoma: a predictive model based on a retrospective analysis of 987 cases. *Intergruppo Italiano Linfomi. Blood*. 2000; 95:783–89.
  33. Lancet JE, Willman CL, Bennett JM. Acute myelogenous leukemia and aging. *Clinical interactions. Hematol Oncol Clin North Am*. 2000; 14:251–67. [https://doi.org/10.1016/S0889-8588\(05\)70287-2](https://doi.org/10.1016/S0889-8588(05)70287-2)
  34. Maas HA, Kruitwagen RF, Lemmens VE, Goey SH, Janssen-Heijnen ML. The influence of age and comorbidity on treatment and prognosis of ovarian cancer: a population-based study. *Gynecol Oncol*. 2005; 97:104–09.  
<https://doi.org/10.1016/j.ygyno.2004.12.026>
  35. Balducci L. Geriatric oncology: challenges for the new century. *Eur J Cancer*. 2000; 36:1741–54.  
[https://doi.org/10.1016/S0959-8049\(00\)00169-6](https://doi.org/10.1016/S0959-8049(00)00169-6)
  36. Dawber TR, Meadors GF, Moore FE Jr. Epidemiological approaches to heart disease: the Framingham Study. *Am J Public Health Nations Health*. 1951; 41:279–81.  
<https://doi.org/10.2105/AJPH.41.3.279>
  37. Driver JA, Beiser A, Au R, Kreger BE, Splansky GL, Kurth T, Kiel DP, Lu KP, Seshadri S, Wolf PA. Inverse association between cancer and Alzheimer’s disease: results from the Framingham Heart Study. *BMJ*. 2012; 344:e1442. <https://doi.org/10.1136/bmj.e1442>
  38. He C, Chasman DI, Dreyfus J, Hwang SJ, Ruitter R, Sanna S, Buring JE, Fernández-Rhodes L, Franceschini N, Hankinson SE, Hofman A, Lunetta KL, Palmieri G, et al. Reproductive aging-associated common genetic variants and the risk of breast cancer. *Breast Cancer Res*. 2012; 14:R54. <https://doi.org/10.1186/bcr3155>
  39. Govindaraju DR, Cupples LA, Kannel WB, O’Donnell CJ, Atwood LD, D’Agostino RB Sr, Fox CS, Larson M, Levy D, Murabito J, Vasani RS, Splansky GL, Wolf PA, Benjamin EJ. Genetics of the Framingham Heart Study population. *Adv Genet*. 2008; 62:33–65.  
[https://doi.org/10.1016/S0065-2660\(08\)00602-0](https://doi.org/10.1016/S0065-2660(08)00602-0)
  40. Cupples LA, Heard-Costa N, Lee M, Atwood LD, and Framingham Heart Study Investigators. Genetics Analysis Workshop 16 Problem 2: the Framingham Heart Study data. *BMC Proc*. 2009 (Suppl 7); 3:S3. <https://doi.org/10.1186/1753-6561-3-s7-s3>
  41. Mailman MD, Feolo M, Jin Y, Kimura M, Tryka K, Bagoutdinov R, Hao L, Kiang A, Paschall J, Phan L, Popova N, Pretel S, Ziyabari L, et al. The NCBI dbGaP database of genotypes and phenotypes. *Nat Genet*. 2007; 39:1181–86. <https://doi.org/10.1038/ng1007-1181>
  42. Kulminski AM, Culminskaya I. Genomics of human health and aging. *Age (Dordr)*. 2013; 35:455–69.  
<https://doi.org/10.1007/s11357-011-9362-x>
  43. Singh H, Sethi S, Raber M, Petersen LA. Errors in cancer diagnosis: current understanding and future directions. *J Clin Oncol*. 2007; 25:5009–18.  
<https://doi.org/10.1200/JCO.2007.13.2142>
  44. Abbasi NR, Shaw HM, Rigel DS, Friedman RJ, McCarthy WH, Osman I, Kopf AW, Polsky D. Early diagnosis of cutaneous melanoma: revisiting the ABCD criteria. *JAMA*. 2004; 292:2771–76.  
<https://doi.org/10.1001/jama.292.22.2771>
  45. Heath M, Jaimes N, Lemos B, Mostaghimi A, Wang LC, Peñas PF, Nghiem P. Clinical characteristics of Merkel cell carcinoma at diagnosis in 195 patients: the AEIOU features. *J Am Acad Dermatol*. 2008; 58:375–81.  
<https://doi.org/10.1016/j.jaad.2007.11.020>
  46. Cormier JN, Xing Y, Ding M, Lee JE, Mansfield PF, Gershenwald JE, Ross MI, Du XL. Ethnic differences among patients with cutaneous melanoma. *Arch Intern Med*. 2006; 166:1907–14.  
<https://doi.org/10.1001/archinte.166.17.1907>

47. Gloster HM Jr, Neal K. Skin cancer in skin of color. *J Am Acad Dermatol*. 2006; 55:741–60. <https://doi.org/10.1016/j.jaad.2005.08.063>
48. Bradford PT. Skin cancer in skin of color. *Dermatol Nurs*. 2009; 21:170–77.
49. Backus LI, Boothroyd DB, Phillips BR, Belperio P, Halloran J, Mole LA. A sustained virologic response reduces risk of all-cause mortality in patients with hepatitis C. *Clin Gastroenterol Hepatol*. 2011; 9:509–516.e1. <https://doi.org/10.1016/j.cgh.2011.03.004>
50. Perna L, Hoffmeister M, Schöttker B, Arndt V, Haug U, Holleczeck B, Burwinkel B, Ordóñez-Mena JM, Brenner H. Vitamin D receptor polymorphism and colorectal cancer-specific and all-cause mortality. *Cancer Epidemiol*. 2013; 37:905–07. <https://doi.org/10.1016/j.canep.2013.09.007>
51. Bjerager M, Palshof T, Dahl R, Vedsted P, Olesen F. Delay in diagnosis of lung cancer in general practice. *Br J Gen Pract*. 2006; 56:863–68.
52. Janssen-Heijnen ML, Maas HA, Houterman S, Lemmens VE, Rutten HJ, Coebergh JW. Comorbidity in older surgical cancer patients: influence on patient care and outcome. *Eur J Cancer*. 2007; 43:2179–93. <https://doi.org/10.1016/j.ejca.2007.06.008>
53. Goede V, Cramer P, Busch R, Bergmann M, Stauch M, Hopfinger G, Stilgenbauer S, Döhner H, Westermann A, Wendtner CM, Eichhorst B, Hallek M, and German CLL Study Group. Interactions between comorbidity and treatment of chronic lymphocytic leukemia: results of German Chronic Lymphocytic Leukemia Study Group trials. *Haematologica*. 2014; 99:1095–100. <https://doi.org/10.3324/haematol.2013.096792>
54. Cardenal F, Nadal E, Jové M, Faivre-Finn C. Concurrent systemic therapy with radiotherapy for the treatment of poor-risk patients with unresectable stage III non-small-cell lung cancer: a review of the literature. *Ann Oncol*. 2015; 26:278–88. <https://doi.org/10.1093/annonc/mdu229>
55. Patnaik JL, Byers T, Diguseppi C, Denberg TD, Dabelea D. The influence of comorbidities on overall survival among older women diagnosed with breast cancer. *J Natl Cancer Inst*. 2011; 103:1101–11. <https://doi.org/10.1093/jnci/djr188>
56. Albertsen PC, Moore DF, Shih W, Lin Y, Li H, Lu-Yao GL. Impact of comorbidity on survival among men with localized prostate cancer. *J Clin Oncol*. 2011; 29:1335–41. <https://doi.org/10.1200/JCO.2010.31.2330>
57. Piccirillo JF, Tierney RM, Costas I, Grove L, Spitznagel EL Jr. Prognostic importance of comorbidity in a hospital-based cancer registry. *JAMA*. 2004; 291:2441–47. <https://doi.org/10.1001/jama.291.20.2441>
58. Williams GR, Mackenzie A, Magnuson A, Olin R, Chapman A, Mohile S, Allore H, Somerfield MR, Targia V, Extermann M, Cohen HJ, Hurria A, Holmes H. Comorbidity in older adults with cancer. *J Geriatr Oncol*. 2016; 7:249–57. <https://doi.org/10.1016/j.jgo.2015.12.002>
59. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis*. 1987; 40:373–83. [https://doi.org/10.1016/0021-9681\(87\)90171-8](https://doi.org/10.1016/0021-9681(87)90171-8)
60. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007; 81:559–75. <https://doi.org/10.1086/519795>
61. Webster MT, Kamgari N, Perloski M, Hoepfner MP, Axelsson E, Hedhammar Å, Pielberg G, Lindblad-Toh K. Linked genetic variants on chromosome 10 control ear morphology and body mass among dog breeds. *BMC Genomics*. BioMed Central. 2015; 16:474. <https://doi.org/10.1186/s12864-015-1702-2>
62. Willcox BJ, Donlon TA, He Q, Chen R, Grove JS, Yano K, Masaki KH, Willcox DC, Rodriguez B, Curb JD. FOXO3A genotype is strongly associated with human longevity. *Proc Natl Acad Sci USA*. 2008; 105:13987–92. <https://doi.org/10.1073/pnas.0801030105>
63. Yuan HY, Chiou JJ, Tseng WH, Liu CH, Liu CK, Lin YJ, Wang HH, Yao A, Chen YT, Hsu CN. FASTSNP: an always up-to-date and extendable service for SNP function analysis and prioritization. *Nucleic Acids Res*. 2006; 34:W635–41. <https://doi.org/10.1093/nar/gkl236>
64. Yue P, Melamud E, Moulton J. SNPs3D: candidate gene and SNP selection for association studies. *BMC Bioinformatics*. 2006; 7:166. <https://doi.org/10.1186/1471-2105-7-166>
65. Tacutu R, Craig T, Budovsky A, Wuttke D, Lehmann G, Taranukha D, Costa J, Fraifeld VE, de Magalhães JP. Human Ageing Genomic Resources: integrated databases and tools for the biology and genetics of ageing. *Nucleic Acids Res*. 2013; 41:D1027–33. <https://doi.org/10.1093/nar/gks1155>
66. Cunningham F, Amode MR, Barrell D, Beal K, Billis K, Brent S, Carvalho-Silva D, Clapham P, Coates G, Fitzgerald S, Gil L, Girón CG, Gordon L, et al. Ensembl 2015. *Nucleic Acids Res*. 2015; 43:D662–69. <https://doi.org/10.1093/nar/gku1010>



67. Lunetta KL, D'Agostino RB Sr, Karasik D, Benjamin EJ, Guo CY, Govindaraju R, Kiel DP, Kelly-Hayes M, Massaro JM, Pencina MJ, Seshadri S, Murabito JM. Genetic correlates of longevity and selected age-related phenotypes: a genome-wide association study in the Framingham Study. *BMC Med Genet.* 2007 (Suppl 1); 8:S13. <https://doi.org/10.1186/1471-2350-8-S1-S13>
68. Kang M, Ding X, Xu M, Zhu H, Liu S, Wang M, Wu D, Tong N, Gong W, Zhou J, Zhang Z. Genetic variation rs10484761 on 6p21.1 derived from a genome-wide association study is associated with gastric cancer survival in a Chinese population. *Gene.* 2014; 536:59–64. <https://doi.org/10.1016/j.gene.2013.11.087>
69. Jang MJ, Kim JW, Jeon YJ, Chong SY, Hong SP, Hwang SG, Oh D, Cho YK, Ji YG, Kim NK. Polymorphisms of folate metabolism-related genes and survival of patients with colorectal cancer in the Korean population. *Gene.* 2014; 533:558–64. <https://doi.org/10.1016/j.gene.2013.09.056>
70. Kanehisa M. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* 2000; 28: 27–30.
71. Therneau TM, Grambsch PM. *Modeling Survival Data: Extending the Cox Model.* New York: Springer; 2000.
72. Ihaka R, Gentleman R. A Language for Data Analysis and Graphics. *J Comput Graph Stat.* 1996; 5:299-314.
73. Harrington D, Fleming T. A class of rank test procedures for censored survival data. *Biometrika.* 1982; 69:553–66. <https://doi.org/10.1093/biomet/69.3.553>
74. Ambrosone CB, Sweeney C, Coles BF, Thompson PA, McClure GY, Korourian S, Fares MY, Stone A, Kadlubar FF, Hutchins LF. Polymorphisms in glutathione S-transferases (GSTM1 and GSTT1) and survival after treatment for breast cancer. *Cancer Res.* 2001; 61:7130–35.
75. Kabbinavar FF, Hambleton J, Mass RD, Hurwitz HI, Bergsland E, Sarkar S. Combined analysis of efficacy: the addition of bevacizumab to fluorouracil/leucovorin improves survival for patients with metastatic colorectal cancer. *J Clin Oncol.* 2005; 23:3706–12. <https://doi.org/10.1200/JCO.2005.00.232>
76. Krag DN, Anderson SJ, Julian TB, Brown AM, Harlow SP, Costantino JP, Ashikaga T, Weaver DL, Mamounas EP, Jalovec LM, Frazier TG, Noyes RD, Robidoux A, et al. Sentinel-lymph-node resection compared with conventional axillary-lymph-node dissection in clinically node-negative patients with breast cancer: overall survival findings from the NSABP B-32 randomised phase 3 trial. *Lancet Oncol.* 2010; 11:927–33. [https://doi.org/10.1016/S1470-2045\(10\)70207-2](https://doi.org/10.1016/S1470-2045(10)70207-2)
77. Govindarajulu US, Lin H, Lunetta KL, D'Agostino RB Sr. Frailty models: applications to biomedical and genetic studies. *Stat Med.* 2011; 30:2754–64. <https://doi.org/10.1002/sim.4277>
78. Kang T, Kraft P, Gauderman WJ, Thomas D, and Framingham Heart Study. Multiple imputation methods for longitudinal blood pressure measurements from the Framingham Heart Study. *BMC Genet.* 2003 (Suppl 1); 4:S43. <https://doi.org/10.1186/1471-2156-4-S1-S43>
79. Moons KG, Donders RA, Stijnen T, Harrell FE Jr. Using the outcome for imputation of missing predictor values was preferred. *J Clin Epidemiol.* 2006; 59:1092–101. <https://doi.org/10.1016/j.jclinepi.2006.01.009>
80. Efstathiou JA, Bae K, Shipley WU, Hanks GE, Pilepich MV, Sandler HM, Smith MR. Cardiovascular mortality and duration of androgen deprivation for locally advanced prostate cancer: analysis of RTOG 92-02. *Eur Urol.* 2008; 54:816–23. <https://doi.org/10.1016/j.eururo.2008.01.021>
81. Jung SY, Sereika SM, Linkov F, Brufsky A, Weissfeld JL, Rosenzweig M. The effect of delays in treatment for breast cancer metastasis on survival. *Breast Cancer Res Treat.* 2011; 130:953–64. <https://doi.org/10.1007/s10549-011-1662-4>
82. van Buuren S, Groothuis-Oudshoorn K. Mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software.* 2011; 45:1-67.
83. The Aric Investigators. The Atherosclerosis Risk in Community(ARIC) Study: design and Objectives. *Am J Epidemiol.* 1989; 129:687–702. <https://doi.org/10.1093/oxfordjournals.aje.a115184>
84. Fried LP, Borhani NO, Enright P, Furberg CD, Gardin JM, Kronmal RA, Kuller LH, Manolio TA, Mittelmark MB, Newman A, O'Leary DH, Psaty B, Rautaharju P, et al. The cardiovascular health study: design and rationale. *Ann Epidemiol.* 1991; 1:263–76. [https://doi.org/10.1016/1047-2797\(91\)90005-W](https://doi.org/10.1016/1047-2797(91)90005-W)
85. Farout L, Friguet B. Proteasome function in aging and oxidative stress: implications in protein maintenance failure. *Antioxid Redox Signal.* 2006; 8:205–16. <https://doi.org/10.1089/ars.2006.8.205>
86. Chondrogianni N, Gonos ES. Proteasome dysfunction in mammalian aging: steps and factors involved. *Exp Gerontol.* 2005; 40:931–38. <https://doi.org/10.1016/j.exger.2005.09.004>
87. Han W, Han MR, Kang JJ, Bae JY, Lee JH, Bae YJ, Lee JE, Shin HJ, Hwang KT, Hwang SE, Kim SW, Noh DY. Genomic alterations identified by array comparative genomic hybridization as prognostic markers in

- tamoxifen-treated estrogen receptor-positive breast cancer. *BMC Cancer*. BioMed Central. 2006; 6:1.
88. Stronach EA, Sellar GC, Blenkinsop C, Rabiasz GJ, Taylor KJ, Miller EP, Massie CE, Al-Nafussi A, Smyth JF, Porteous DJ, Gabra H. Identification of clinically relevant genes on chromosome 11 in a functional model of ovarian cancer tumor suppression. *Cancer Res*. 2003; 63:8648–55.
89. Torresan C, Santos SC, Camps J, Pereira SF, Ribeiro EM, Cavalli IJ, Haddad BR, Cavalli LR. Abstract 327: genomic profiling of sentinel lymph node breast cancer metastasis. *Cancer Res*. 2014; 70:327–327. <https://doi.org/10.1158/1538-7445.AM10-327>
90. Kulminski AM, Arbeev KG, Culminskaya I, Ukraintseva SV, Stallard E, Province MA, Yashin AI. Trade-offs in the effects of the apolipoprotein E polymorphism on risks of diseases of the heart, cancer, and neurodegenerative disorders: insights on mechanisms from the Long Life Family Study. *Rejuvenation Res*. 2015; 18:128–35. <https://doi.org/10.1089/rej.2014.1616>
91. Christensen K, Johnson TE, Vaupel JW. The quest for genetic determinants of human longevity: challenges and insights. *Nat Rev Genet*. 2006; 7:436–48. <https://doi.org/10.1038/nrg1871>
92. Sakashita K, Tanaka F, Zhang X, Mimori K, Kamohara Y, Inoue H, Sawada T, Hirakawa K, Mori M. Clinical significance of ApoE expression in human gastric cancer. *Oncol Rep*. 2008; 20:1313–19.
93. Chen YC, Pohl G, Wang TL, Morin PJ, Risberg B, Kristensen GB, Yu A, Davidson B, Shih IM, Apolipoprotein E. Apolipoprotein E is required for cell proliferation and survival in ovarian cancer. *Cancer Res*. 2005; 65:331–37.
94. Kamboh MI, Demirci FY, Wang X, Minster RL, Carrasquillo MM, Pankratz VS, Younkin SG, Saykin AJ, Jun G, Baldwin C, Logue MW, Buross J, Farrer L, et al, and Alzheimer's Disease Neuroimaging Initiative. Genome-wide association study of Alzheimer's disease. *Transl Psychiatry*. 2012; 2:e117. <https://doi.org/10.1038/tp.2012.45>
95. Brisson GD, Alves LR, Pombo-de-Oliveira MS. Genetic susceptibility in childhood acute leukaemias: a systematic review. *Eccancermedalscience*. 2015; 9:539. <https://doi.org/10.3332/ecancer.2015.539>
96. Hosking FJ, Papaemmanuil E, Sheridan E, Kinsey SE, Lightfoot T, Roman E, Irving JA, Allan JM, Taylor M, Tomlinson IP, Greaves M, Houlston RS. Genome-wide homozygosity signatures and childhood acute lymphoblastic leukemia risk. *Blood*. 2010; 115:4472–77. <https://doi.org/10.1182/blood-2009-09-244483>
97. Repasky GA, Chenette EJ, Der CJ. Renewing the conspiracy theory debate: does Raf function alone to mediate Ras oncogenesis? *Trends Cell Biol*. 2004; 14:639–47. <https://doi.org/10.1016/j.tcb.2004.09.014>
98. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc*. 2009; 4:1073–81. <https://doi.org/10.1038/nprot.2009.86>
99. Newman AB, Glynn NW, Taylor CA, Sebastiani P, Perls TT, Mayeux R, Christensen K, Zmuda JM, Barral S, Lee JH, Simonsick EM, Walston JD, Yashin AI, Hadley E. Health and function of participants in the Long Life Family Study: A comparison with other cohorts. *Aging (Albany NY)*. 2011; 3:63–76. <https://doi.org/10.18632/aging.100242>
100. Ikeda Y, Kumagai H, Motozawa Y, Suzuki J, Akazawa H, Komuro I. Understanding Vascular Diseases: Lessons From Premature Aging Syndromes. *Can J Cardiol*. 2016; 32:650–58. <https://doi.org/10.1016/j.cjca.2015.12.003>
101. Sugimoto M. A cascade leading to premature aging phenotypes including abnormal tumor profiles in Werner syndrome (review). *Int J Mol Med*. 2014; 33:247–53. <https://doi.org/10.3892/ijmm.2013.1592>
102. Coppède F. The epidemiology of premature aging and associated comorbidities. *Clin Interv Aging*. 2013; 8:1023–32. <https://doi.org/10.2147/CIA.S37213>
103. Goto M, Miller RW, Ishikawa Y, Sugano H. Excess of rare cancers in Werner syndrome (adult progeria). *Cancer Epidemiol Biomarkers Prev*. 1996; 5:239–46.
104. Lauper JM, Krause A, Vaughan TL, Monnat RJ Jr. Spectrum and risk of neoplasia in Werner syndrome: a systematic review. *PLoS One*. 2013; 8:e59709. <https://doi.org/10.1371/journal.pone.0059709>
105. Köttgen A, Pattaro C, Böger CA, Fuchsberger C, Olden M, Glazer NL, Parsa A, Gao X, Yang Q, Smith AV, O'Connell JR, Li M, Schmidt H, et al. New loci associated with kidney function and chronic kidney disease. *Nat Genet*. 2010; 42:376–84. <https://doi.org/10.1038/ng.568>
106. Leslie R, O'Donnell CJ, Johnson AD. GRASP: analysis of genotype-phenotype results from 1390 genome-wide association studies and corresponding open access database. *Bioinformatics*. 2014; 30:i185–94. <https://doi.org/10.1093/bioinformatics/btu273>
107. Zhou T, Weems M, Wilke CO. Translationally optimal codons associate with structurally sensitive sites in proteins. *Mol Biol Evol*. Oxford University Press; 2009; 26: 1571–80.

108. Hudson NJ, Gu Q, Nagaraj SH, Ding YS, Dalrymple BP, Reverter A. Eukaryotic evolutionary transitions are associated with extreme codon bias in functionally-related proteins. *PLoS One*. 2011; 6:e25457. <https://doi.org/10.1371/journal.pone.0025457>
109. Angov E, Hillier CJ, Kincaid RL, Lyon JA. Heterologous protein expression is enhanced by harmonizing the codon usage frequencies of the target gene with those of the expression host. *PLoS One*. 2008; 3:e2189. <https://doi.org/10.1371/journal.pone.0002189>
110. Sauna ZE, Kimchi-Sarfaty C. Understanding the contribution of synonymous mutations to human disease. *Nat Rev Genet*. 2011; 12:683–91. <https://doi.org/10.1038/nrg3051>
111. Waring SC, Rosenberg RN. Genome-wide association studies in Alzheimer disease. *Arch Neurol*. 2008; 65:329–34. <https://doi.org/10.1001/archneur.65.3.329>
112. Perls TT. The different paths to 100. *Am J Clin Nutr*. 2006; 83:484S–87S.
113. Guarente L, Franklin H. Epstein Lecture: Sirtuins, aging, and medicine. *N Engl J Med*. 2011; 364:2235–44. <https://doi.org/10.1056/NEJMra1100831>
114. Kim HS, Patel K, Muldoon-Jacobs K, Bisht KS, Aykin-Burns N, Pennington JD, van der Meer R, Nguyen P, Savage J, Owens KM, Vassilopoulos A, Ozden O, Park SH, et al. SIRT3 is a mitochondria-localized tumor suppressor required for maintenance of mitochondrial integrity and metabolism during stress. *Cancer Cell*. 2010; 17:41–52. <https://doi.org/10.1016/j.ccr.2009.11.023>
115. Finley LW, Carracedo A, Lee J, Souza A, Egia A, Zhang J, Teruya-Feldstein J, Moreira PI, Cardoso SM, Clish CB, Pandolfi PP, Haigis MC. SIRT3 opposes reprogramming of cancer cell metabolism through HIF1 $\alpha$  destabilization. *Cancer Cell*. 2011; 19:416–28. <https://doi.org/10.1016/j.ccr.2011.02.014>
116. Bellizzi D, Rose G, Cavalcante P, Covello G, Dato S, De Rango F, Greco V, Maggolini M, Feraco E, Mari V, Franceschi C, Passarino G, De Benedictis G. A novel VNTR enhancer within the SIRT3 gene, a human homologue of SIR2, is associated with survival at oldest ages. *Genomics*. 2005; 85:258–63. <https://doi.org/10.1016/j.ygeno.2004.11.003>
117. Greer EL, Brunet A. FOXO transcription factors in ageing and cancer. *Acta Physiol (Oxf)*. 2008; 192:19–28. <https://doi.org/10.1111/j.1748-1716.2007.01780.x>
118. Mueller MM, Castells-Roca L, Babu V, Ermolaeva MA, Müller RU, Frommolt P, Williams AB, Greiss S, Schneider JI, Benzing T, Schermer B, Schumacher B. DAF-16/FOXO and EGL-27/GATA promote developmental growth in response to persistent somatic DNA damage. *Nat Cell Biol*. 2014; 16:1168–79. <https://doi.org/10.1038/ncb3071>
119. McCormick M, Chen K, Ramaswamy P, Kenyon C. New genes that extend *Caenorhabditis elegans*' lifespan in response to reproductive signals. *Aging Cell*. 2012; 11:192–202. <https://doi.org/10.1111/j.1474-9726.2011.00768.x>
120. Tommiska J, Eerola H, Heinonen M, Salonen L, Kaare M, Tallila J, Ristimäki A, von Smitten K, Aittomäki K, Heikkilä P, Blomqvist C, Nevanlinna H. Breast cancer patients with p53 Pro72 homozygous genotype have a poorer survival. *Clin Cancer Res*. 2005; 11:5098–103. <https://doi.org/10.1158/1078-0432.CCR-05-0173>
121. Lu H, Shu XO, Cui Y, Kataoka N, Wen W, Cai Q, Ruan ZX, Gao YT, Zheng W. Association of genetic polymorphisms in the VEGF gene with breast cancer survival. *Cancer Res*. 2005; 65:5015–19. <https://doi.org/10.1158/0008-5472.CAN-04-2786>
122. de Magalhães JP. Why genes extending lifespan in model organisms have not been consistently associated with human longevity and what it means to translation research. *Cell Cycle*. 2014; 13:2671–73. <https://doi.org/10.4161/15384101.2014.950151>
123. Bergman A, Atzmon G, Ye K, MacCarthy T, Barzilai N. Buffering mechanisms in aging: a systems approach toward uncovering the genetic component of aging. *PLOS Comput Biol*. 2007; 3:e170. <https://doi.org/10.1371/journal.pcbi.0030170>
124. Kulminski AM. Unraveling genetic origin of aging-related traits: evolving concepts. *Rejuvenation Res*. 2013; 16:304–12. <https://doi.org/10.1089/rej.2013.1441>

## **SUPPLEMENTARY MATERIAL**

Please browse the Full text version of this manuscript to see the Supplementary Tables S1-S8.