

## Research

### Adolescent Cardiorespiratory Fitness and Risk of Cancer in Late Adulthood: Nationwide Sibling-Controlled Cohort Study

Marcel Ballin, associated researcher<sup>1,2,a</sup>, Daniel Berglind, associate professor<sup>2,3,4</sup>, Pontus Henriksson, associate professor<sup>5</sup>, Martin Neovius, professor<sup>6</sup>, Anna Nordström, researcher<sup>7,8</sup>, Francisco B. Ortega, professor<sup>9,10</sup>, Elina Sillanpää, associate professor<sup>10,11</sup>, Peter Nordström, professor<sup>1,b</sup>, Viktor H. Ahlqvist, postdoctoral researcher<sup>1,12,13,a,b</sup>

<sup>1</sup>Department of Public Health and Caring Sciences, Clinical Geriatrics, Uppsala University, Uppsala, Sweden.

<sup>2</sup>Centre for Epidemiology and Community Medicine, Region Stockholm, Stockholm, Sweden.

<sup>3</sup>Department of Global Public Health, Karolinska Institutet, Stockholm, Sweden.

<sup>4</sup>Center for Wellbeing, Welfare and Happiness, Stockholm School of Economics, Stockholm, Sweden.

<sup>5</sup>Department of Health, Medicine and Caring Sciences, Linköping University, Linköping, Sweden.

<sup>6</sup>Department of Medicine, Clinical Epidemiology Division, Karolinska Institutet, Stockholm, Sweden.

<sup>7</sup>Department of Medical Sciences, Uppsala University, Uppsala, Sweden.

<sup>8</sup>School of Sports Science, UiT The Arctic University of Norway, Tromsø, Norway.

<sup>9</sup>Department of Physical Education and Sports, Faculty of Sport Sciences, Sport and Health University Research Institute (iMUDS), University of Granada; CIBEROBN, ISCIII, Granada, Andalucía, Spain.

<sup>10</sup>Faculty of Sport and Health Sciences, University of Jyväskylä, Jyväskylä, Finland.

<sup>11</sup>Wellbeing Services County of Central Finland, Jyväskylä, Finland.

<sup>12</sup>Department of Biomedicine, Aarhus University, Aarhus, Denmark.

<sup>13</sup>Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden.

<sup>a</sup>Dr. Ballin and Dr. Ahlqvist are joint corresponding authors.

<sup>b</sup>Prof. Nordström and Dr. Ahlqvist are joint senior authors.

## Abstract

**Objective:** To investigate whether the higher risks of certain cancers associated with high cardiorespiratory fitness can be explained by increased detection and unobserved confounders.

**Design:** Nationwide sibling-controlled cohort study of adolescents.

**Setting:** Sweden.

**Participants:** 1 124 049 men of which 477 453 were full siblings, who underwent mandatory military conscription examinations between 1972 and 1995 at a mean age of 18.3 years.

**Main outcome measures:** Hazard ratios (HR) and 95% confidence intervals (CI) of overall cancer diagnosis and cancer mortality, and 14 site-specific cancers (diagnosis or death), as recorded in the Swedish National Patient Register or Cause of Death Register until 31 December 2023, modelled using flexible parametric regressions.

**Results:** Participants were followed until a median (maximum) age of 55.9 (73.5) years, during which 98 410 were diagnosed with cancer and 16 789 had a cancer-related death (41 293 and 6908 among full siblings respectively). The most common cancers were non-melanoma skin (27 105 diagnoses & 227 deaths) and prostate cancer (24 211 diagnoses & 869 deaths). In cohort analysis, those in the highest quartile of cardiorespiratory fitness had a higher risk of prostate (adjusted HR 1.10; 95% CI: 1.05 to 1.16) and skin cancer (e.g., non-melanoma HR 1.44; 1.37 to 1.50) compared to those in the lowest quartile, which led to a higher risk of any type of cancer diagnosis (HR 1.08; 1.06 to 1.11). However, those in the highest quartile had a lower risk of cancer mortality (HR 0.71; 0.67 to 0.76). When comparing full siblings, and thereby controlling for all behavioural, environmental, and genetic factors they share, the excess risk of prostate (HR 1.01; 0.90 to 1.13) and skin cancer (e.g., non-melanoma HR 1.09; 0.99 to 1.20) attenuated to the null. In contrast, the lower risk of overall cancer mortality was still statistically significant after control for such shared confounders (HR 0.78; 0.68 to 0.89). For other site-specific cancers, the influence of such confounding tended to vary, but none showed the same excess risk as prostate and non-melanoma skin cancer.

**Conclusions:** The association between high levels of adolescent cardiorespiratory fitness and excess risk of some cancers, such as prostate and non-melanoma skin

cancer, appears to be fully explained by unobserved confounders shared between full siblings. However, the protective association with cancer mortality persists even after control for such confounding.

## Summary box

### What is already known on this topic

- Adolescent physical activity and cardiorespiratory fitness are considered important factors for the prevention of cancer based on evidence from observational studies.
- Observational studies are, however, vulnerable to unobserved confounders and bias processes, including health-seeking behaviours and genetic and environmental confounders.
- These biases could explain why prior studies have found that high adolescent cardiorespiratory fitness is associated with higher risks of some cancers, typically low-mortality cancers such as prostate and non-melanoma skin cancer.

### What this study adds

- This nationwide cohort study of 1.1 million male adolescents showed that while higher cardiorespiratory fitness was associated with excess risk of the most common cancers - prostate and non-melanoma skin - these associations attenuated to the null when accounting for behavioural, environmental, and genetic confounders shared between full siblings.
- In contrast, high adolescent cardiorespiratory fitness was associated with a lower risk of overall cancer mortality, which remained after controlling for unobserved confounders shared between full siblings.

## Introduction

The International Agency for Research in Cancer states that physical activity is a proven way to lower the risk of cancers,<sup>1</sup> a statement also reflected through the World Health Organization guidelines for physical activity.<sup>2</sup> Recently, cardiorespiratory fitness, a robust marker reflecting long-term physical activity and genetic predisposition,<sup>3,4</sup> has also been implicated in cancer risk.<sup>5-9</sup> Considering the challenges in identifying novel modifiable risk factors for cancer prevention,<sup>10</sup> along with the downward trends in physical activity and cardiorespiratory fitness levels especially among young people,<sup>11-14</sup> public health interventions targeting these traits in adolescents are gaining traction as a means of cancer prevention.

Unfortunately, the current evidence supporting the role of physical activity and cardiorespiratory fitness in cancer is associated with low certainty.<sup>9,15</sup> One major limitation is that existing large-scale studies have been unable to capture important differences in behavioural, environmental and genetic confounders. This limitation is compounded by the fact that existing studies often used a composite endpoint of diagnoses and deaths, which could introduce detection bias given that individuals with high levels of physical activity more often have different health-seeking behaviours,<sup>16</sup> such as higher participation in cancer screening programs.<sup>17,18</sup> This increases the likelihood of early diagnosis, including for cancers typically associated with low mortality (e.g., non-melanoma skin cancer). When this is bundled together with cancer mortality, a protective effect on cancers associated with higher mortality may be underestimated or lost in the mix. Some studies have even found that higher physical activity and cardiorespiratory fitness are associated with a higher risk of overall cancer,<sup>6,19-22</sup> and especially skin and prostate cancer,<sup>6,7,18,23</sup> probably reflecting such bias.

These differential health behaviours remain difficult to capture and address in large-scale observational studies. One possible way to account for unobserved factors is to compare family members;<sup>24</sup> who are assumed to share a large part of these behaviours and other important genetic and environmental confounders.<sup>25,26</sup>

In this study, we aim to address the existing inconsistencies and establish a more accurate appreciation for the cancer prevention potential of cardiorespiratory fitness. We conducted a nationwide sibling-controlled cohort study over six decades, encompassing over 1.1 million young men, of which half a million were full siblings, to

investigate how the associations between adolescent cardiorespiratory and cancers in late adulthood are affected by bias processes.

## **Methods**

### **Patient and public involvement**

No patients or members of the public were involved in the study design, data collection, data analysis, interpretation of the results, decision to publish the paper or preparation of the manuscript.

### **Study design**

A cohort study with full-sibling analysis was conducted by cross-linking data from Swedish health and population registers, using the Personal Identification Number, which is unique for all Swedish residents.<sup>27</sup> The need for informed consent was waived by the Swedish Ethical Review Authority as the study only leveraged data from existing registers.<sup>28</sup> The study is reported in accordance with the STROBE guidelines.<sup>29</sup>

### **Databases**

From the Swedish Military Service Conscription Register<sup>30</sup>, we obtained an eligible study population based on all men who participated in military conscription examinations between 1972 and 1995, during which cardiorespiratory fitness was assessed. During this period, conscription around the age of 18 years was mandatory for all males in Sweden, with few exemptions (approximately 90% population coverage).<sup>30</sup> To this, family relationships were identified using the Multi-Generation Register<sup>31</sup>, subsequent mortality using the National Cause of Death Register<sup>32</sup>, inpatient and specialist outpatient care using the National Patient Register<sup>33</sup>, and emigration and socioeconomic data using the registries of Statistics Sweden<sup>31</sup>. Overall, these data cover the entire Swedish population and are mandated by law.

### **Derivation of study population**

A total of 1 249 131 men were conscripted between 1972 and 1995. From these, we excluded 33 645 (2.7%) with missing data on cardiorespiratory fitness, 72 042 (5.8%) with missing data on the covariates, and 19 395 (1.5%) with extreme values (detailed below). This resulted in 1 124 049 conscripts being included in the co-

hort analysis (90.0% retained). Among these, 477 453 were full siblings from 219 304 families and were included in the sibling analysis.

## Exposure

Cardiorespiratory fitness was assessed at conscription using validated equipment and a standardised procedure as described in detail elsewhere.<sup>30</sup> In short, conscripts performed a maximal ergometer bicycle test following normal electrocardiography. The conscript began cycling for 5 minutes at 60 to 70 revolutions per minute at a low level of predetermined (by body weight) resistance. The external resistance was then gradually increased by 25 watts every minute until exhaustion. The test results were recorded as watt maximum ( $W_{max}$ ), which correlates strongly with maximal oxygen uptake.<sup>34</sup> Similar to in previous studies, we excluded conscripts with extremely low values ( $<100 W_{max}$ ).<sup>26,35</sup>

## Outcomes

The primary outcomes were overall cancer diagnosis and cancer mortality, as recorded in the National Patient Register and the Cause of Death register until 31 December 2023. To this, we also perform an analysis of 14 site-specific cancers (either diagnosis or death), which were selected based on their previously identified associations with physical activity or cardiorespiratory fitness:<sup>2,5,6,19</sup> head and neck, oesophagus, lung, stomach, pancreas, colon, rectum, kidney, prostate, bladder, myeloma, skin (divided into melanoma and non-melanoma), and liver, bile ducts, and gallbladder. A diagnosis was ascertained as the date of first hospitalisation or visit in specialised outpatient care, recorded using diagnostic codes in a primary or secondary position, and death was classified as cancer-related if it was recorded as the primary or secondary cause on the death certificate (supplementary table 1).

## Covariates

From the Swedish Military Service Conscription Register, we collected data on age at conscription, year of conscription, and measured body mass index (BMI,  $\text{kg}/\text{m}^2$ ). As in previous studies, conscripts with extreme BMI values ( $\leq 15$  and  $\geq 60$   $\text{kg}/\text{m}^2$ ) were excluded.<sup>26,36</sup> From Statistics Sweden,<sup>37</sup> we obtained data on the socioeconomic status of both mothers and fathers of the conscripts, including information on the highest attained lifetime education and the annual disposable income standardised by birth years into highest-achieved quintiles between ages 40 and 50 (used to capture working life income, henceforth referred to as income categories).

When values for both the mother and father were available, the highest value was retained.

### Statistical analysis

Participants were followed from the day of conscription until the date of a cancer outcome, emigration, death from non-cancer causes, or end of follow-up (31 December 2023), whichever came first. Flexible parametric survival models were used to estimate hazard ratios (HRs) for each outcome by levels of cardiorespiratory fitness, using age as the underlying time scale, and baseline knots placed at the 5<sup>th</sup>, 27.5<sup>th</sup>, 50<sup>th</sup>, 72.5<sup>th</sup>, and 95<sup>th</sup> percentile of the uncensored log survival times.<sup>38,39</sup> In contrast to Cox regression, these models directly estimate the baseline hazard, which enables the computation of absolute effects measures.<sup>38</sup> We modelled cardiorespiratory fitness both as quartiles and using restricted cubic splines with knots placed at the 5<sup>th</sup>, 35<sup>th</sup>, 65<sup>th</sup>, and 95<sup>th</sup> percentile.<sup>39</sup> Since HRs are difficult to clinically contextualize,<sup>40</sup> we also calculated the standardised cumulative incidence at 65 years of age (1-Survival) and illustrated it graphically over the follow-up period. All models were adjusted for age at conscription (continuous), year of conscription (1972, 1973-1977, 1978-1982, 1983-1987, 1988-1992, and 1993-1995), BMI (continuous and quadratic term), parental education (compulsory school  $\leq 9$  years, secondary education, post-secondary education  $< 3$  years, post-secondary education  $\geq 3$  years), and parental income (5 categories).

For the full-sibling analysis, the aforementioned model was extended to a marginalised between-within model with robust standard errors, which enables control for all unobserved confounders (including genetic and shared environmental factors).<sup>24,41,42</sup> The between-within model isolates the individual-level variation (within effect) from the family-level variation (between effect) by including a term for the exposure/covariate and a term for its family average.<sup>24,41,42</sup>

To further aid public health interpretation, population-attributable fractions (PAF) were calculated,<sup>43</sup> considering a “major” intervention that would shift everyone to the top quartile of cardiorespiratory fitness, or a “moderate” intervention that would shift everyone who belonged to the bottom quartile to the second quartile. All analyses were performed using Stata MP version 16.1.



## **Sensitivity analyses**

A series of sensitivity analyses were conducted to test the robustness of the results. First, to examine whether differences in estimates in sibling analysis as compared to cohort analysis were more likely to be due to selection bias rather than control for unobserved familial confounders, the standard analysis was replicated in the sibling cohort (i.e., without controlling for shared factors).<sup>42</sup> Second, because BMI could be considered both a confounder or mediator in the associations between cardiorespiratory fitness and cancer,<sup>44</sup> the main analysis was repeated but excluding control for BMI. Third, to examine the influence of death from non-cancer causes as a competing event, the aforementioned flexible parametric models were used to compute cause-specific cumulative incidence functions of all cancer outcomes at 65 years of age, accounting for death from non-cancer causes.<sup>45</sup>

## **Results**

### **Baseline characteristics**

Baseline characteristics were similar in the total cohort and in the sibling cohort (table 1). The mean (standard deviation) age at conscription was 18.3 (0.7) years and most (81.3%) were normal weight. About a quarter had parents with a high (post-secondary) level of education and about a third had parents with a high level of annual income.

### **Cancers during follow-up**

During a median (range) follow-up of 37.7 (0.1 to 51.4) years, 98 410 (8.8%) were diagnosed with cancer and 16 789 (1.5%) died of cancer in the total cohort, and 41 293 (8.7%) and 6908 (1.5%) in the sibling cohort. The median age at first cancer diagnosis was 55.1 (19.1 to 73.0) years in the total cohort and 54.8 (19.1 to 72.6) years in the sibling cohort. The median age at cancer death was 55.7 (18.4 to 72.9) years in the total cohort and 55.3 (18.4 to 72.9) years in the sibling cohort. The most common site-specific cancers were non-melanoma skin (27 105 diagnoses and 227 deaths) and prostate cancer (24 211 diagnoses and 869 deaths) (table 1, supplementary table 2).

### **Cardiorespiratory fitness and overall cancer**

In cohort analysis, high cardiorespiratory fitness was associated with a higher risk of cancer diagnosis (fig 1, supplementary fig 1). Compared to the bottom quartile

of cardiorespiratory fitness, the HR in the top quartile was 1.08 (1.06 to 1.11), with a 1.5 (1.0 to 1.9) percentage point difference in the standardised cumulative incidence at 65 years of age (supplementary tables 3-4). Calculations of PAFs produced negative estimates, implying a theoretical increase in the proportion of cancer diagnoses should either a moderate or major intervention be carried out (supplementary table 5). In full-sibling analysis, the association attenuated to the null (HR 1.00; 0.95 to 1.06; incidence difference 0.08 (-0.9 to 1.1) percentage points).

In cohort analysis, high cardiorespiratory fitness was associated with a lower risk of cancer mortality (fig 1, supplementary fig 1). Compared to the bottom quartile, the HR in the top quartile was 0.71 (0.67 to 0.76, with an incidence difference of -1.1 (-1.2 to -0.9) percentage points (supplementary tables 3, 4). The PAF of cancer mortality associated with a moderate intervention was 16.5% (14.4 to 18.6) and for a major intervention 22.6% (19.2 to 26.1) (supplementary table 5). In full-sibling analysis, the association only modestly attenuated (HR 0.78; 0.68 to 0.89; incidence difference -0.7 (-1.1 to -0.4) percentage points). As such, the PAF associated with a moderate intervention dropped to 11.3% (6.1 to 16.6) and for a major intervention 17.2% (8.2 to 25.5).

### **Cardiorespiratory fitness and site-specific cancer**

In cohort analysis, high cardiorespiratory fitness was associated with a lower risk (HRs ranging from 0.81 to 0.49) of the rectum, head and neck, bladder, stomach, pancreas, colon, kidney, liver, bile ducts, and gallbladder, oesophagus, and lung cancer (fig 2, supplementary figs 1-3, supplementary tables 3-4). In contrast, high cardiorespiratory fitness was associated with a higher risk of prostate (HR 1.10; 1.05 to 1.16; incidence difference 0.6 (percentage points), melanoma (HR 1.50; 1.45 to 1.60; incidence difference 1.0 percentage points), and non-melanoma skin cancer (HR 1.44; 1.37 to 1.50; incidence difference 2.4 percentage points).

In full-sibling analysis, the degree of change in the associations, as compared to in the total cohort, varied between cancer types. The excess risk of non-melanoma skin cancer was largely attenuated (HR 1.09; 0.99 to 1.20; incidence difference 0.6 percentage points) as was the risk for prostate cancer (HR 1.01; 0.90 to 1.13; incidence difference 0.08 percentage points). Other tendencies toward clear attenuation were observed also for liver, bile ducts, and gallbladder cancer (from HR 0.60 to

0.97) and for kidney cancer (from HR 0.65 to 0.86), translating into reduced PAFs (supplementary table 5).

### **Sensitivity analyses**

Replication of the standard analysis in the sibling cohort yielded overall similar estimates as observed in the cohort analysis (supplementary table 6). The exception was rectum cancer, for which the estimates were more similar to those observed in sibling analysis. Excluding adjustment for BMI resulted in similar or slightly weaker associations for most outcomes, with a similar pattern in the cohort and sibling analysis (supplementary table 7). Accounting for the competing risk of non-cancer death (so-called crude risks) was similar to standard analysis across most outcomes (supplementary table 8). The exceptions were overall cancer diagnosis and site-specific cancers of the prostate and skin, for which the crude risks were slightly smaller.

## **Discussion**

### **Principal findings**

In this nationwide sibling-controlled cohort study spanning six decades, we find evidence suggesting that epidemiological associations between adolescent cardiorespiratory fitness and later-life cancer risk may be biased. Specifically, individuals with high adolescent cardiorespiratory fitness were at a higher risk of being diagnosed with non-melanoma skin cancer and prostate cancer. However, when adjusting for unobserved behavioural, environmental, and genetic confounders shared between full siblings, these excess risks disappeared. In contrast, high cardiorespiratory fitness was associated with lower cancer-related mortality, which remained after adjustment for such confounding.

### **Comparisons with other studies**

Physical activity and cardiorespiratory fitness have been linked to lower cancer risk, although the certainty in the evidence is limited, including a largely unknown contribution of unobserved confounding and bias processes.<sup>5-9,19,25</sup> In this study, high adolescent cardiorespiratory fitness was linked to excess risk of non-melanoma skin and prostate cancer, as in previous observational studies and meta-analyses in men.<sup>6,19</sup> However, we provide evidence to suggest that this risk appears to be fully explained by bias as it could not be replicated after accounting for unobserved confounders shared between full siblings. The sibling analysis cannot pinpoint the exact

confounding factor(s), but it may be a combination of differential health-seeking behaviours (including participation in cancer screening), as well as other shared environmental factors and genetics. Previous studies showing associations between physical activity and cardiorespiratory fitness and excess risk of skin and prostate cancer have often suggested that their findings may be the result of bias,<sup>6,18,46</sup> speculating that it may be explained by fit people also spending more time outdoors in the sun and that they participate in screening programs to a greater extent.<sup>47,48</sup> Formal validation of these confounding and bias processes has, however, been lacking until now. Yet, even in this large-scale study, the excess risk of melanoma skin cancer could not be fully explained by confounders shared between full siblings. This may suggest the presence of biases beyond those shared by siblings, and we encourage future research to further scrutinize this association.

The association between high adolescent cardiorespiratory fitness and lower risk of overall cancer mortality remained statistically significant after accounting for such confounding factors (HR 0.78; 95% CI: 0.68 to 0.89). One previous study based on the same population reported a lower risk of both cancer diagnosis and cancer mortality in sibling analysis.<sup>25</sup> Yet, interpretation of those estimates and comparisons to ours is difficult considering that the study relied on a younger cohort with substantially fewer cancer diagnoses and deaths (15 093 [1.3%] & 4900 [0.4%], compared to our study with 98 410 [8.8%] diagnoses & 16 789 [1.5%] deaths).<sup>25</sup> Additionally, only time-specific hazard ratios were reported<sup>25</sup>, which are difficult to interpret as they are largely influenced by the depletion of susceptibles.<sup>40</sup> Nevertheless, based on the results presented here, it appears important to distinguish a diagnosis of cancer from cancer-related mortality since collapsing the two may dilute the benefits associated with high-mortality cancers by introducing biased associations with a diagnosis of low-mortality cancers detected in screening. By extension, this could suggest that previous large-scale studies and meta-analyses of cardiorespiratory fitness or physical activity, which relied on composite endpoints, might have underestimated the protective associations.<sup>5,6,19-22</sup> If this is true, and can be verified by future causal analysis, the potential of public health initiatives targeting cardiorespiratory fitness or physical activity to reduce cancer mortality might be underappreciated.<sup>49,50</sup>

This study also revealed a heterogeneous pattern regarding adolescent cardiorespiratory fitness and site-specific cancer after comparing full siblings. For example,

the associations showing lower risks of kidney and liver, bile ducts, and gallbladder cancer seemed to attenuate more, whereas those with colon and head and neck cancer seemed more robust. These findings may suggest a more nuanced role of adolescent cardiorespiratory fitness in site-specific cancer, whereby there may be a greater preventive potential for certain types and less for others. Assuming part of the attenuation is the result of control for familial confounders, these findings align with research showing a clear variation in the contribution of genetic and shared environmental factors to the risk of different cancers.<sup>51,52</sup>

### **Strengths and limitations**

This study has several notable strengths, including its nationwide coverage, family-based analysis, standardised measures of cardiorespiratory fitness, extensive follow-up until late adulthood with virtually zero attrition, and investigation of both diagnoses of cancer in a universal healthcare system and cancer mortality.

However, some limitations should be considered. First, our study was limited to men due to historical conscription practices in Sweden. While we may believe there are limited reasons for sex-specific mechanisms linking cardiorespiratory fitness and physical activity to cancer, it is important to note that cancer prevalence varies by sex, and we were unable to examine cancers specific to females. Second, our study focuses on adolescent cardiorespiratory fitness and may provide insights into public health policies aimed at young individuals. However, it may not directly inform interventions targeted at other populations, despite evidence that cardiorespiratory fitness moderately tracks from adolescence into adulthood.<sup>53</sup> Third, due to the limited number of cases of site-specific cancers, we were unable to conduct separate analyses of diagnosis and mortality. Furthermore, the small counts of some site-specific cancers can explain the confidence intervals overlapping the null in the sibling analysis, which is not necessarily evidence of the absence of effects. Fourth, while comparisons of full siblings are expected to control for several important confounders that are difficult to measure at scale, this design hinges on assumptions<sup>42,54</sup> and there likely remains residual confounding (e.g., the remaining 50% of genetics not shared between full siblings and behaviours not shared by full siblings with different fitness levels). Future studies utilizing monozygotic twins or other causal designs would be beneficial for triangulating these findings.

## Implications

Population-wide initiatives targeting cardiorespiratory fitness early in life might hold promise in contributing to reducing future cancer mortality burden. Such initiatives may be particularly warranted given that population levels of cardiorespiratory fitness are trending downwards across all ages, and especially in young people.<sup>11-13</sup> While more causal data is needed before definitive policy statements can be made, the influence of behavioural, environmental, and genetic confounders shared between full siblings appear to only partly diminish the public health potential of such interventions.

Taken at face value, the hypothetically preventable share of cancer mortality at 65 years of age, should everyone from the bottom quartile be moved to the second lowest quartile, was estimated to be 16.5% when comparing unrelated individuals in the population and 11.3% when comparing full siblings. This shift assumed that those with the lowest fitness would increase their fitness level by 14% on average. This corresponds roughly to the effect seen in exercise interventions,<sup>55</sup> although it is unclear to what extent that findings from controlled and typically supervised exercise interventions can be extrapolated to large-scale public health and community interventions.

## Conclusion

This study suggests that high cardiorespiratory fitness in adolescence is associated with lower risk of overall cancer mortality in late adulthood after accounting for behavioural, environmental, and genetic confounders shared by full siblings. Furthermore, our findings provide evidence suggesting that epidemiological studies on this topic may yield biased conclusions if not conducted carefully, as observed excess risks of certain cancers may stem from methodological biases rather than indicating genuine cause for concern.

**Table 1. Baseline characteristics and number of cancer outcomes during follow-up in the total cohort and the full sibling cohort.**

	<b>Total cohort (N=1 124 049)</b>	<b>Full sibling cohort (N=477 453)</b>
<b>Baseline data</b>		
<b>Birth year, median (IQR)</b>	1966 (1960 to 1971)	1965 (1961 to 1970)
<b>Age at conscription, mean (SD)</b>	18.3 (0.7)	18.3 (0.7)
<b>Body mass index, kg/m<sup>2</sup></b>		
Median (range)	21.4 (15.0 to 59.2)	21.3 (15.0 to 59.2)
<b>Body mass index categories, n (%)</b>		
Underweight (<18.5 kg/m <sup>2</sup> )	88 966 (7.9)	37 884 (7.9)
Normal weight (18.5-24.9 kg/m <sup>2</sup> )	913 586 (81.3)	390 811 (81.9)
Overweight (25.0-29.9 kg/m <sup>2</sup> )	102 267 (9.1)	41 248 (8.6)
Obesity (>30.0 kg/m <sup>2</sup> )	19 230 (1.7)	7510 (1.6)
<b>Cardiorespiratory fitness, Watt max</b>		
Median (range)	271 (100 to 999)	271 (100 to 999)
<b>Watt max by quartiles, median (range)</b>		
Quartile 1	217 (100 to 236)	217 (100 to 236)
Quartile 2	253 (237 to 270)	253 (237 to 270)
Quartile 3	290 (271 to 312)	290 (271 to 312)
Quartile 4	339 (313 to 999)	339 (313 to 999)
<b>Parental level of education, n (%)</b>		
Compulsory school <9 years	319 242 (28.4)	139 246 (29.1)
Secondary education	505 510 (45.0)	210 246 (44.0)
Post-secondary education <3 years	123 757 (11.0)	50 210 (10.5)
Post-secondary education >3 years	175 540 (15.6)	77 751 (16.3)
<b>Parental highest income, n (%)</b>		
Category 1 (low income)	55 055 (4.9)	18 918 (4.0)
Category 2	108 836 (9.7)	41 820 (8.8)
Category 3	237 953 (21.2)	101 150 (21.2)
Category 4	338 444 (30.1)	145 735 (30.5)
Category 5 (high income)	383 761 (34.1)	169 830 (35.6)
<b>Number of cancer events (median age at event) during follow-up</b>		
Overall cancer diagnosis	98 410 (55.1)	41 293 (54.8)
Overall cancer mortality	16 789 (55.7)	6908 (55.3)
Head and neck	4026 (54.1)	1692 (53.7)
Oesophagus	1178 (57.1)	464 (57.0)

Lung	3131 (57.4)	1263 (57.1)
Stomach	1430 (55.4)	599 (55.3)
Pancreas	2255 (57.2)	943 (57.0)
Liver, bile ducts, and gallbladder	2246 (57.1)	925 (56.8)
Colon	5320 (55.1)	2177 (54.8)
Rectum	3917 (55.7)	1625 (55.6)
Kidney	2741 (54.6)	1151 (54.5)
Prostate	24 225 (59.5)	10 042 (59.1)
Bladder	3490 (56.9)	1432 (56.3)
Myeloma	1460 (55.8)	580 (55.9)
Melanoma skin	10 026 (52.5)	4216 (52.5)
Non-melanoma skin	27 302 (54.5)	11 509 (54.2)

---

IQR = interquartile range. SD = standard deviation.



## Figure legends

**Figure 1. Standardised cumulative incidences and hazard ratios for overall cancer diagnosis and cancer mortality by quartiles of cardiorespiratory fitness in cohort and sibling analysis.** Estimates obtained using flexible parametric survival models, extended to a marginalised between-within model in the sibling cohort, with baseline knots placed at the 5<sup>th</sup>, 27.5<sup>th</sup>, 50<sup>th</sup>, 72.5<sup>th</sup>, and 95<sup>th</sup> percentile of the uncensored log survival times, and using age as the underlying time scale. The bottom quartile was the referent. All models were adjusted for age at conscription, year of conscription, body mass index, parental education, and parental income. Inferential measures for the incidences were omitted for clarity as they are also reflected by the hazard ratios and are reported in supplementary tables 3 and 4.

**Figure 2. Hazard ratios for diagnosis or death from site-specific cancers by quartiles of cardiorespiratory fitness in cohort and sibling analysis.** Estimates were obtained using flexible parametric survival models, extended to a marginalised between-within model in the sibling cohort, with baseline knots placed at the 5<sup>th</sup>, 27.5<sup>th</sup>, 50<sup>th</sup>, 72.5<sup>th</sup>, and 95<sup>th</sup> percentile of the uncensored log survival times, and using age as the underlying time scale. The lowest quartile was the referent. All models were adjusted for age at conscription, year of conscription, body mass index, parental education, and parental income.

## References

1. International Agency for Research on Cancer. World Health Organization. World Cancer Day 2021: Spotlight on IARC research related to physical activity and nutrition. Available from: <https://www.iarc.who.int/featured-news/world-cancer-day-2021-physical-activity/>.
2. World Health Organization. Guidelines on physical activity and sedentary behaviour. Geneva: World Health Organization, 2020. Available from: <https://www.who.int/publications/i/item/9789240015128>.
3. Raghuveer G, Hartz J, Lubans DR, et al. Cardiorespiratory Fitness in Youth: An Important Marker of Health: A Scientific Statement From the American Heart Association. *Circulation* 2020; **142**(7): e101-e18.

4. Schutte NM, Nederend I, Hudziak JJ, Bartels M, de Geus EJ. Twin-sibling study and meta-analysis on the heritability of maximal oxygen consumption. *Physiol Genomics* 2016; **48**(3): 210-9.
5. Pozuelo-Carrascosa DP, Alvarez-Bueno C, Cavero-Redondo I, Morais S, Lee IM, Martínez-Vizcaíno V. Cardiorespiratory fitness and site-specific risk of cancer in men: A systematic review and meta-analysis. *Eur J Cancer* 2019; **113**: 58-68.
6. Onerup A, Mehlig K, Geijerstam AA, et al. Associations between cardiorespiratory fitness in youth and the incidence of site-specific cancer in men: a cohort study with register linkage. *Br J Sports Med* 2023; **57**(19): 1248-56.
7. Ekblom-Bak E, Bojsen-Moller E, Wallin P, et al. Association Between Cardiorespiratory Fitness and Cancer Incidence and Cancer-Specific Mortality of Colon, Lung, and Prostate Cancer Among Swedish Men. *JAMA Netw Open* 2023; **6**(6): e2321102.
8. Han M, Qie R, Shi X, et al. Cardiorespiratory fitness and mortality from all causes, cardiovascular disease and cancer: dose–response meta-analysis of cohort studies. *British Journal of Sports Medicine* 2022; **56**(13): 733-9.
9. Lang JJ, Prince SA, Merucci K, et al. Cardiorespiratory fitness is a strong and consistent predictor of morbidity and mortality among adults: an overview of meta-analyses representing over 20.9 million observations from 199 unique cohort studies. *British Journal of Sports Medicine* 2024.
10. Davey Smith G, Hofman A, Brennan P. Chance, ignorance, and the paradoxes of cancer: Richard Peto on developing preventative strategies under uncertainty. *Eur J Epidemiol* 2023; **38**(12): 1227-37.
11. Tomkinson GR, Lang JJ, Tremblay MS. Temporal trends in the cardiorespiratory fitness of children and adolescents representing 19 high-income and upper middle-income countries between 1981 and 2014. *Br J Sports Med* 2019; **53**(8): 478-86.
12. Lamoureux NR, Fitzgerald JS, Norton KI, Sabato T, Tremblay MS, Tomkinson GR. Temporal Trends in the Cardiorespiratory Fitness of 2,525,827 Adults Between 1967 and 2016: A Systematic Review. *Sports Med* 2019; **49**(1): 41-55.
13. Ekblom-Bak E, Ekblom Ö, Andersson G, et al. Decline in cardiorespiratory fitness in the Swedish working force between 1995 and 2017. *Scand J Med Sci Sports* 2019; **29**(2): 232-9.

14. Conger SA, Toth LP, Cretsinger C, et al. Time Trends in Physical Activity Using Wearable Devices: A Systematic Review and Meta-analysis of Studies from 1995 to 2017. *Medicine & Science in Sports & Exercise* 2021.
15. WHO guidelines on physical activity and sedentary behaviour: Web Annex. Evidence profiles. Geneva: World Health Organization; 2020.
16. Byun W, Sui X, Hebert JR, et al. Cardiorespiratory fitness and risk of prostate cancer: findings from the Aerobics Center Longitudinal Study. *Cancer epidemiology* 2011; **35**(1): 59-65.
17. Lakhani NA, Saraiya M, Thompson TD, King SC, Guy Jr GP. Total body skin examination for skin cancer screening among US adults from 2000 to 2010. *Preventive medicine* 2014; **61**: 75-80.
18. Crump C, Stattin P, Brooks JD, et al. Early-Life Cardiorespiratory Fitness and Long-term Risk of Prostate Cancer. *Cancer Epidemiol Biomarkers Prev* 2020; **29**(11): 2187-94.
19. Moore SC, Lee IM, Weiderpass E, et al. Association of Leisure-Time Physical Activity With Risk of 26 Types of Cancer in 1.44 Million Adults. *JAMA Intern Med* 2016; **176**(6): 816-25.
20. Stamatakis E, Ahmadi MN, Friedenreich CM, et al. Vigorous intermittent lifestyle physical activity and cancer incidence among nonexercising adults: The UK Biobank Accelerometry Study. *JAMA oncology* 2023; **9**(9): 1255-9.
21. Su J, Jiang Y, Fan X, et al. Association between physical activity and cancer risk among Chinese adults: a 10-year prospective study. *International Journal of Behavioral Nutrition and Physical Activity* 2022; **19**(1): 150.
22. Garcia L, Pearce M, Abbas A, et al. Non-occupational physical activity and risk of cardiovascular disease, cancer and mortality outcomes: a dose-response meta-analysis of large prospective studies. *Br J Sports Med* 2023; **57**(15): 979-89.
23. Lakoski SG, Willis BL, Barlow CE, et al. Midlife Cardiorespiratory Fitness, Incident Cancer, and Survival After Cancer in Men: The Cooper Center Longitudinal Study. *JAMA Oncol* 2015; **1**(2): 231-7.
24. Sjölander A, Frisell T, Öberg S. Sibling comparison studies. *Annual Review of Statistics and Its Application* 2022; **9**: 71-94.
25. Hogstrom G, Ohlsson H, Crump C, Sundquist J, Sundquist K. Aerobic fitness in late adolescence and the risk of cancer and cancer-associated mortality in

adulthood: A prospective nationwide study of 1.2 million Swedish men. *Cancer epidemiology* 2019; **59**: 58-63.

26. Ballin M, Neovius M, Ortega FB, et al. Genetic and Environmental Factors and Cardiovascular Disease Risk in Adolescents. *JAMA Netw Open* 2023; **6**(11): e2343947.

27. Ludvigsson JF, Otterblad-Olausson P, Pettersson BU, Ekblom A. The Swedish personal identity number: possibilities and pitfalls in healthcare and medical research. *Eur J Epidemiol* 2009; **24**(11): 659-67.

28. Ludvigsson JF, Haberg SE, Knudsen GP, et al. Ethical aspects of registry-based research in the Nordic countries. *Clin Epidemiol* 2015; **7**: 491-508.

29. von Elm E, Altman DG, Egger M, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ* 2007; **335**(7624): 806-8.

30. Ludvigsson JF, Berglind D, Sundquist K, Sundstrom J, Tynelius P, Neovius M. The Swedish military conscription register: opportunities for its use in medical research. *Eur J Epidemiol* 2022; **37**(7): 767-77.

31. Ludvigsson JF, Almqvist C, Bonamy A-KE, et al. Registers of the Swedish total population and their use in medical research. *Eur J Epidemiol* 2016; **31**(2): 125-36.

32. Brooke HL, Talbäck M, Hörnblad J, et al. The Swedish cause of death register. *Eur J Epidemiol* 2017; **32**(9): 765-73.

33. Ludvigsson JF, Andersson E, Ekblom A, et al. External review and validation of the Swedish national inpatient register. *BMC Public Health* 2011; **11**: 450.

34. Andersen LB. A maximal cycle exercise protocol to predict maximal oxygen uptake. *Scand J Med Sci Sports* 1995; **5**(3): 143-6.

35. Ballin M, Nordstrom A, Nordstrom P. Cardiovascular Disease and All-Cause Mortality in Male Twins With Discordant Cardiorespiratory Fitness: A Nationwide Cohort Study. *Am J Epidemiol* 2020; **189**(10): 1114-23.

36. Henriksson P, Henriksson H, Tynelius P, et al. Fitness and Body Mass Index During Adolescence and Disability Later in Life: A Cohort Study. *Ann Intern Med* 2019; **170**(4): 230-9.

37. Ludvigsson JF, Svedberg P, Olen O, Bruze G, Neovius M. The longitudinal integrated database for health insurance and labour market studies

(LISA) and its use in medical research. *European Journal of Epidemiology* 2019; **34**(4): 423-37.

38. Lambert PC, Royston P. Further development of flexible parametric models for survival analysis. *The Stata Journal* 2009; **9**(2): 265-90.

39. Harrell FE. Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis: Springer; 2001.

40. Hernan MA. The hazards of hazard ratios. *Epidemiology* 2010; **21**(1): 13-5.

41. Sjolander A. Estimation of marginal causal effects in the presence of confounding by cluster. *Biostatistics* 2021; **22**(3): 598-612.

42. Sjolander A, Oberg S, Frisell T. Generalizability and effect measure modification in sibling comparison studies. *Eur J Epidemiol* 2022; **37**(5): 461-76.

43. Chen YQ, Hu C, Wang Y. Attributable risk function in the proportional hazards model for censored time-to-event. *Biostatistics* 2006; **7**(4): 515-29.

44. Watts EL, Gonzales TI, Strain T, et al. Observational and genetic associations between cardiorespiratory fitness and cancer: a UK Biobank and international consortia study. *British Journal of Cancer* 2024; **130**(1): 114-24.

45. Lambert PC, Wilkes SR, Crowther MJ. Flexible parametric modelling of the cause-specific cumulative incidence function. *Stat Med* 2017; **36**(9): 1429-46.

46. Behrens G, Niedermaier T, Berneburg M, Schmid D, Leitzmann MF. Physical activity, cardiorespiratory fitness and risk of cutaneous malignant melanoma: Systematic review and meta-analysis. *PloS one* 2018; **13**(10): e0206087.

47. Reiter-Brennan C, Dzaye O, Al-Mallah MH, et al. Fitness and prostate cancer screening, incidence, and mortality: Results from the Henry Ford Exercise Testing (FIT) Project. *Cancer* 2021; **127**(11): 1864-70.

48. Wei G, Farooq J, Castelo-Soccio L, Mhaskar R. Correlates between physical activity and sunburn prevalence among a nationally representative sample of US high school students, 2015–2017. *Journal of Physical Activity and Health* 2021; **18**(9): 1113-9.

49. Brazo-Sayavera J, Silva DR, Lang JJ, et al. Physical Fitness Surveillance and Monitoring Systems Inventory for Children and Adolescents: A Scoping Review with a Global Perspective. *Sports Med* 2024.

50. Ortega FB, Leskosek B, Blagus R, et al. European fitness landscape for children and adolescents: updated reference values, fitness maps and country

rankings based on nearly 8 million test results from 34 countries gathered by the FitBack network. *Br J Sports Med* 2023; **57**(5): 299-310.

51. Lichtenstein P, Holm NV, Verkasalo PK, et al. Environmental and heritable factors in the causation of cancer—analyses of cohorts of twins from Sweden, Denmark, and Finland. *New England journal of medicine* 2000; **343**(2): 78-85.

52. Mucci LA, Hjelmborg JB, Harris JR, et al. Familial risk and heritability of cancer among twins in Nordic countries. *Jama* 2016; **315**(1): 68-76.

53. Garcia-Hermoso A, Izquierdo M, Ramirez-Velez R. Tracking of physical fitness levels from childhood and adolescence to adulthood: a systematic review and meta-analysis. *Transl Pediatr* 2022; **11**(4): 474-86.

54. Frisell T, Oberg S, Kuja-Halkola R, Sjolander A. Sibling comparison designs: bias from non-shared confounders and measurement error. *Epidemiology* 2012; **23**(5): 713-20.

55. Scribbans TD, Vecsey S, Hankinson PB, Foster WS, Gurd BJ. The Effect of Training Intensity on VO<sub>2</sub>max in Young Healthy Adults: A Meta-Regression and Meta-Analysis. *Int J Exerc Sci* 2016; **9**(2): 230-47.

## Footnotes

### Contributors

MB and VHA conceived the study. MB, MN, PH, and VHA designed the study. PN acquired the ethical permission and the data. MB performed the statistical analyses and drafted the manuscript. VHA verified the underlying data. PN and VHA performed data managing. PN and VHA are joint senior authors. All authors interpreted the data. All authors critically revised the manuscript for intellectual content. The corresponding authors MB and VHA attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted. MB and VHA is the manuscript's guarantors and accept full responsibility for the conduct of the study, had access to the data, and controlled the decision to publish. All authors reviewed and approved the manuscript for submission.

### Role of the funding source

The author(s) received no specific funding for this work. As such, no funder had any role in the study design, data collection and analysis, decision to publish, or

preparation of the manuscript. MN is funded via grants from the Swedish Research Council (2019-00738). VHA is funded via grants from the National Institute for Aging and the National Institute of Neurological Disorders and Stroke (1R01NS131433-01).

### **Competing interests**

All authors have completed the ICMJE uniform disclosure form and declare: MN reported serving on advisory boards for Johnson & Johnson and Itrim, and serving as a consultant for the Armed forces; no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

### **Ethical approval**

This study was approved by the Regional Ethical Review Board in Umeå (no. 2010-113-31M).

### **Data sharing**

The data in this study are not available to the public and will not be shared according to regulations under Swedish law. Researchers interested in obtaining the data may seek ethical approvals and inquire through Statistics Sweden. For further advice see: <https://www.scb.se/en/services/guidance-for-researchers-and-universities/>.

### **Transparency statement**

The lead author MB and the last author VHA (the manuscript's guarantors) affirm that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

### **Dissemination to participants and related patients and public communities**

The findings of this study will be disseminated to the public through a press release, popular science articles and seminars.



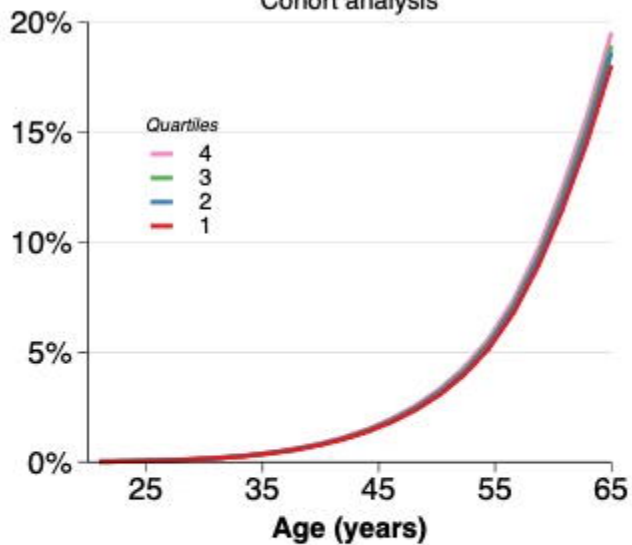
## **Copyright/licence for publication**

The Corresponding Authors have the right to grant on behalf of all authors and do grant on behalf of all authors, a worldwide licence to the Publishers and its licensees in perpetuity, in all forms, formats and media (whether known now or created in the future), to i) publish, reproduce, distribute, display and store the Contribution, ii) translate the Contribution into other languages, create adaptations, reprints, include within collections and create summaries, extracts and/or, abstracts of the Contribution, iii) create any other derivative work(s) based on the Contribution, iv) to exploit all subsidiary rights in the Contribution, v) the inclusion of electronic links from the Contribution to third party material where-ever it may be located; and, vi) licence any third party to do any or all of the above.



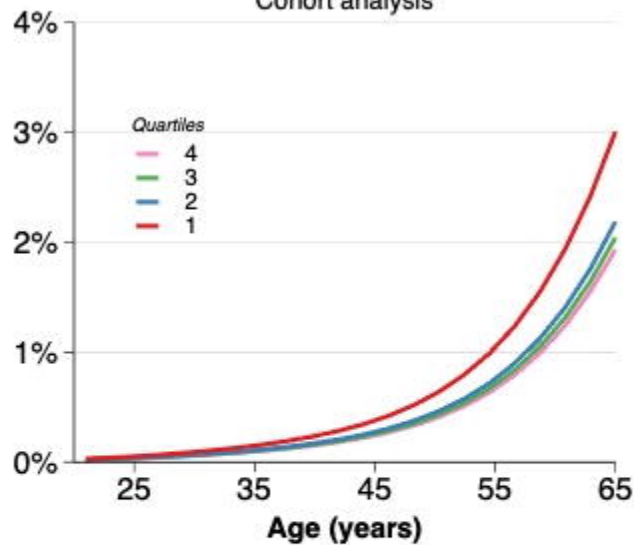
### Cancer diagnosis

Cohort analysis

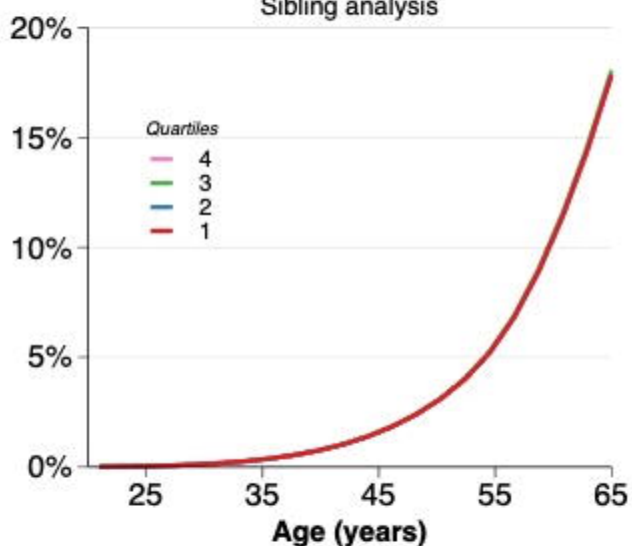


### Cancer mortality

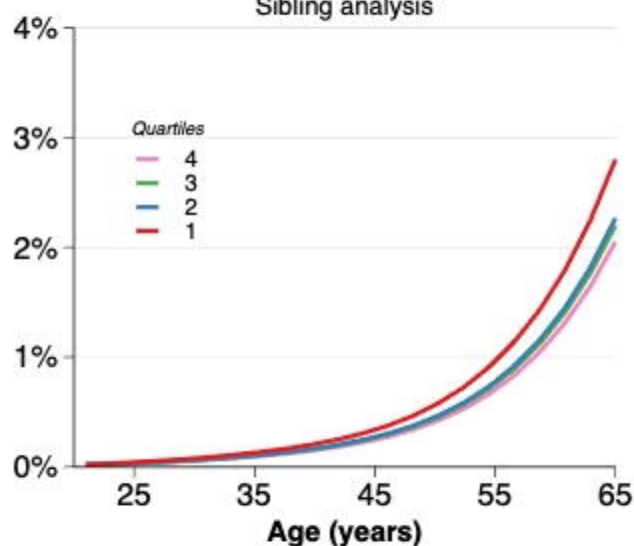
Cohort analysis



Sibling analysis



Sibling analysis



Hazard Ratio (95% CI)

