Cancer incidence and mortality in Pakistanis, Bangladeshis and their descendants in England and Wales

Joseph Harrison[1] (ORCID: 0000-0001-6750-6004) jh383@st-andrews.ac.uk

Frank Sullivan[2] (0000-0002-6623-4964)

Katherine Keenan[1] (0000-0002-9670-1607)

Hill Kulu[1] (0000-0001-8808-0719)

1: University of St Andrews, School of Geography and Sustainable Development, Irvine Building, North Street, St Andrews, KY16 9AL United Kingdom

2: University of St Andrews, School of Medicine, St Andrews, United Kingdom

Abstract

Background: The migrant mortality advantage is a widely observed phenomenon, thought to occur due to positive selection of migrants and better health behaviours. This paper seeks to further understand health differentials between Pakistani and Bangladeshi immigrants, their descendants, and the native population in England and Wales. We choose to focus on cancer as one of the leading causes of morbidity and mortality in developed countries.

Methods: We apply survival analysis to the Office for National Statistics Longitudinal Study of England and Wales, to compare hazard ratios of cancer incidence between these groups. Moreover, we observe the ten-year period after diagnosis to identify differences between these groups in mortality following onset of cancer. We apply stepwise models to control for

socioeconomic characteristics that have previously been found to influence health and mortality.

Results: We find that the risk of cancer onset is substantially lower for individuals born in Pakistan and Bangladesh. This advantage is also seen in their British born descendants. However, following incidence of cancer there is no significant difference in mortality between these groups, and for descendants the mortality risk after onset may be elevated.

Conclusions: We conclude that lower incidence of cancer and not better survival once diagnosed is the driver of the low cancer mortality observed in Pakistanis and Bangladeshis in England and Wales. We should investigate further how protective behaviours prevent the onset of cancer but fail to improve survivability. Using this detailed administrative data to investigate both incidence and onset of cancer across immigrant generations is a novel contribution and sheds new light on the migrant mortality advantage and immigrant health, particularly in relation to cancer.

Keywords

Immigrants, Cancer, Migrant mortality, Survival analysis, England and Wales

Declarations

Ethics approval and consent to participate: Not applicable.

Consent for publication: Not applicable.

Availability of data and materials: The data that support the findings of this study are available from Office for National Statistics – Longitudinal Study but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available.

Competing interests: The authors declare that they have no competing interests.

Funding: The study is funded by the University of St Andrews

Authors' contributions: JH prepared figures and conducted analysis. KK FS and HK conceived the topic and research design. JH prepared the initial manuscript. All authors were responsible for interpretation of the data and development of the manuscript through revisions.

Acknowledgements: The authors would like to thank the St Leonard's Postgraduate College at the University of St Andrews for the World Leading Doctoral Scholarship, which funds this project. This paper was prepared within the framework of the MigrantLife project, aimed at 'Understanding Life Trajectories of Immigrants and Their Descendants in Europe and Projecting Future Trends.' This project is led by Hill Kulu and received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 834103).

The permission of the Office for National Statistics to use the Longitudinal Study is gratefully acknowledged, as is the help provided by staff of the Centre for Longitudinal Study Information & User Support (CeLSIUS). CeLSIUS is funded by the ESRC under project ES/V003488/1. The authors alone are responsible for the interpretation of the data.

We particularly thank Dr Alison Sizer, Dr Jemima Stockton and Dr Jitka Pikhartova for their ongoing support with the data analysis.

1 Introduction

Immigrants have lower mortality compared to the native population in the destination country [1, 2]. This phenomenon has been referred to as the 'migrant mortality advantage' [3]. These findings are often generalised as immigrants overall being in better health. Recent findings increasingly suggest that despite longer life expectancies the foreign-born population spend more time with morbidities and in worse health, referred to as the migrant health mortality paradox [4]. Previous studies have addressed health inequalities through various measures such as mortality [5–8], self-reported health [9], life expectancy, including disability free life expectancy, [10, 11] and multimorbidity [12]. However, what is often neglected are the pathways involved in specific aspects of health such as disease onset and subsequent survival. Here we consider cancer, a leading cause of death in the UK.

Cancers are some of the most prevalent diseases across the western world: accounting for around 30% of deaths in high income countries in 2019 [13]. Environmental factors are considered the cause of many cancers implying they are, in theory, preventable. For example tobacco and alcohol use, dietary factors, obesity, environmental pollutants and infections [14, 15]. With global mobility increasing, it is important to develop an understanding of the differences in cancer incidence and mortality between immigrants and host populations. Cancer disparities between natives and migrants may exist in terms of incidence (i.e., chance of developing cancer in the first place) and subsequent survival. In relation to the migrant mortality advantage, cancer incidence could be lower and/or survival after diagnosis better. However, the migrant health mortality paradox would suggest that

immigrant groups experience cancer for longer, either the onset being earlier in their lives or increased longevity after incidence, but not complete remission or cure.

This study investigates cancer onset and mortality among Pakistani and Bangladeshi immigrants and their descendants in the UK. We aim to answer three questions: First, how does cancer onset differ between White British, Pakistani and Bangladeshi immigrants, and their descendants? Second, following a cancer diagnosis is there a difference in survival between White British, Pakistani and Bangladeshi immigrants, and their descendants? Third, can any differences in onset or survival be explained by socio-economic factors?

This study provides further evidence on the healthy migrant effect and if it is lost between generations, we expand previous studies in the following ways. First, we investigate differences both in the incidence and survival of cancer between immigrants, descendants, and natives. Cancer research previously has taken an approach that looks either at the incidence of cancer [16–18] or mortality [5, 19–21]. Where research has looked at both the focus is often on specific cancers and limited to certain geographic areas of England [22–24]. Second, we distinguish between immigrants and their descendants, which most previous studies have not done. This is important as the experiences and socialisation of the foreign-born will differ from their children, and these differences may lead to different health behaviour. Prior studies either use ethnicity, thus combing immigrants with their descendants [16, 22, 23], or categorise using only birth country [5, 19]. Moreover, studies often combine Pakistanis and Bangladeshis together, often with Indians too, which has wider implications due to the heterogeneity of migrants from this region [25, 26]. Lastly, we use one dataset, the Office for National Statistics Longitudinal Study (ONS-LS) of England and Wales. The ONS-LS contains linked census and life events data for a

representative 1% sample of the population of England and Wales. Using this data, we present a complete picture of cancer incidence and survival differences between Pakistani and Bangladeshi immigrants, their descendants and ancestral White British across England and Wales.

2 Background

2.1 Theories of Migrant and Minority Health

Lower mortality amongst migrants has been found in Belgium [27, 28], Nordic Countries [29], England and Wales [5, 30] and France [31]. There are competing hypotheses which seek to explain the apparent advantage in immigrant health. First, health selection, with those that experience migration coming from a certain subset of the origin population who are generally in better health than the non-migrants and that this advantage means they have above average health in the destination country [32–35]. Some also argue that considering selection alone is not sufficient and that the migration context is also important [27]. Sending countries that are similar to the destination, for example neighbouring countries or those which are culturally similar, experience less of an advantage since selection is less relevant to the migration process [8, 36].

The effects of this migrant health advantage over time are mixed. Some find that longer duration in the host country reduces the advantage, making it most pronounced between ages 20-40 [8, 31, 33, 35, 37]. Research in Germany has seen the advantage reverse at older ages, with increased migrant mortality relative to natives [38]. Other studies have found little reduction in the migrant health advantage with length of residence [27, 36]. Why the benefits reduce over time can be considered part of the Immigrant Health Transition, where the accumulated risk factors in the destination country eventually neutralise the availability of better

161 healthcare and lower risk of infectious diseases [39]. This explains why child
162 migrants do not experience a health advantage to the same extent, since their
163 extended life course in the destination country and their adaptation of 'native-like'
164 health behaviours aligns them more with the majority [40]. The reducing mortality
165 advantage over the life course has been linked with an accumulation of hardship due
166 to migrant status [31] including socioeconomic deprivation [38, 41] and experiences
167 of discrimination, particularly impactful for mental health [41, 42].

168 A second explanation of migrant health advantage is the maintenance of positive
169 health behaviours amongst immigrants, for example less smoking, lower alcohol
170 intake, better diets, and maintenance of other, healthier habits than the native
171 population [34]. This is used to partially explain differences in mortality patterns
172 between natives and immigrants. However, there are debates on the strength of this
173 as observed differences between stayers in the origin and migrants suggest that
174 there are differences in mortality patterns despite the shared culture and behaviours
175 [37].

176 A third possibility is that the advantage is a data artefact, more specifically that there
177 is over coverage in population statistics of immigrant groups because onward or
178 return migration is poorly recorded [43] this making them appear immortal in
179 analysis. Whilst this source of bias does the advantage has been found to remain
180 even if over coverage is accounted for [8, 44]. Fourth, there is potential selective
181 return migration to the origin country in later years or the salmon bias hypothesis
182 [34]. Positing that immigrants who experience health decline, return to their origin
183 country meaning that their mortality is unobserved in the destination. This does exist
184 in England and Wales amongst certain immigrant groups but cannot fully explain the
185 migrant mortality advantage [30]. The existence of mortality advantages at younger

ages alleviates the idea that selective out migration is a reason for the migrant mortality advantage [37]. Although, it has been observed internally with elevated death of return migrants to northern Sweden [45].

Lastly, are potential genetic mechanisms that alter the susceptibility of some immigrant groups to certain medical conditions. There is evidence of familial inherited susceptibility to cancer however this is not seen as a major influence on population level cancer statistics [46]. Additionally, experiences related to historical famine in the origin country can alter the epigenetic make-up of the migrating population [47]. There are also origin effects which alter vulnerability to infection related cancers, such as liver, which are often found to be higher amongst immigrant groups in high income countries [28].

2.2 Pakistani and Bangladeshi Cancer in the UK

Research on mortality of immigrants in England and Wales has shown a mortality advantage for those born in South Asia and low cancer mortality is thought to contribute to this [5]. Moreover, there are findings that suggest better survival, although this has narrowed in more recent years [20]. Low cancer incidence and mortality amongst Pakistanis and Bangladeshis is not consistent across cancer sites [16]. Mortality from lung, colorectal, breast and prostate cancer are all lower for the Pakistani born population [19] and other types, such as liver and stomach cancer, are more prevalent in those of Pakistani ethnicity [16, 21].

Tobacco and alcohol consumption are leading causes of cancers, increasing the risk of cancer of the lungs, liver, and throat amongst others [48, 49]. Amongst Pakistani and Bangladeshi women smoking is very low, Pakistani men smoke less than White British men and Bangladeshi men more, although for Bangladeshis deprivation can explain this gap [50]. Alcohol consumption is also substantially lower than that of the

native population [51] and although mortality from alcohol misuse in these immigrant groups has increased over time it remains lower than White British groups [52].

Environmental factors and air pollution are linked to the development of neoplasms and subsequent mortality [53], the implications of air pollution on health are found to be more pronounced for ethnic minorities including Pakistani and Bangladeshis [54]. The cause of this is thought to be residential clustering of minorities in urban areas as a result of socioeconomic disparities [55]. Pakistanis and Bangladeshis are far less mobile through the life course [56] thus dangerous levels of exposure air pollution can accumulate over time leading to worse health and more neoplasm development.

Diets high in processed foods have been associated with an increased risk of cancer [57], or predictive of obesity, a risk factor of cancer [58]. Pakistani and Bangladeshi diets for the immigrant generation remain rooted in traditional dishes from the country of origin [59]. Some of these dishes are high in fat, salt, and oil. High intake of these are linked to the increased prevalence of cardiovascular disease, diabetes and obesity amongst the South Asian population of the United Kingdom [60]. Physical activity amongst South Asians is substantially lower compared to White British groups too, with many barriers to participation [61] including residential and socioeconomic deprivation [62, 63].

The incidence of cancers can be influenced by the engagement of immigrants with healthcare, including intervention and screening programs and presentation to healthcare professionals making diagnosis more likely. For South Asians in the UK, relative to the rest of the population attendance of bowel screening is around 50% and breast screening around 80% [64]. These rates are even lower amongst Muslim

South Asians specifically a group more likely to include Pakistani and Bangladeshis. Explanations proposed for this include, lower knowledge of the existence of services which persists even when considering socioeconomic differences [65, 66]. This also combines with sociocultural beliefs which can firstly, affect the level of fatalism associated with cancer and therefore lower understanding of the benefits of screening attendance [67] and secondly, lead to reliance on faith and spiritual practices for treatment rather than western medicine [68]. Linguistic barriers are also a concern for many Asian women for presenting with symptoms or attendance at breast and cervical screenings [69] and affect South Asian participation in colorectal screening [70]. The inequalities in attendance at screening and presentation continue to be a reason that explains the slower increase in breast and prostate survival compared to other groups [20].

UK based studies often homogenise South Asian groups owing to data availability. This is problematic as health risks affecting this group vary along social and cultural lines [25]. This has consequences for cancer onset and mortality and is a limitation of many previous cancer studies which homogenise groups as South Asian [71]. In the UK context the selection mechanisms and integration pathways of Indians is vastly different to that of Pakistanis and Bangladeshis [26, 72]. These divergences materialise in varied socioeconomic outcomes and behaviours within the subgroups of South Asian [73] and is a reason for studying these groups separately.

2.3 Health of descendants

Descendants lack the same selection mechanisms as their parents; thus, the healthy migrant paradox is not as relevant in explaining health differentials. Data artefacts and return migration are also less of a factor. There can be emigration to the parental origin country, but in the UK context return is rarest for those from the Indian

subcontinent due to the large economic inequality [74], therefore over coverage of Pakistani and Bangladeshi descendants in administrative data is unlikely. Positive health selection also is less of a factor, inheritance of good genetic disposition is possible [39], and evidence exists that parental longevity is correlated with their offspring's longevity [75]. However, the fact that cancer is becoming increasingly related to lifestyle factors suggests this is not the largest driver of potential divergences [15].

The adoption of more unhealthy behaviours is offered as the explanation of why the mortality advantage is lost, and possibly reversed, for descendants of Pakistani and Bangladeshi immigrants [76]. For immigrants, increased exposure time in the destination is associated with deterioration of the benefits of positive selection [77], and the adoption of unhealthy behaviours such as, reliance fat rich foods [59] and increased smoking [78]. Hence for descendants, who experience entire life courses in the destination country, the logic would suggest that this group would assimilate to more negative health behaviours.

Evidence supports this, second generation Pakistanis and Bangladeshis are more likely to engage in tobacco and alcohol consumption compared to their immigrant parents [51] however, still less than the native population. For diets there is conflicting views about how adapted descendants diets are with some suggesting little change between generations [51] and others suggesting that British born descendants adopt the negative aspects of British diets leading to worse health [59]. Descendants do have the advantage of lower barriers to healthcare access, owing to better language knowledge and familiarity with the healthcare system, thus making them more likely to engage with cancer screening and intervention programs.

Mortality studies of second-generation immigrants are limited due to the younger age structure of the group with less observed health decline, including relatively fewer instances of cancer. Globally there are mixed results for the mortality of descendants in adulthood. In Switzerland those with a foreign background are found to have lower mortality rates [79], although not the case for descendants of Italian migrants in Switzerland [80]. US born Latin Americans also appear to have a health advantage over US born Whites [34]. In Sweden the mortality advantage found amongst first generation immigrants is reversed for their descendants [81], this is seemingly related to increased deaths at younger ages from external factors such as accidents, suicides related to mental health disorders, and substance misuse. Findings in the UK suggest little variation in self-reported health between the foreign born and the UK born descendants [82], however native-born ethnic minority groups do lose much of the mortality advantage that their immigrant ancestors have [6, 76].

2.4 Cancer in Descendants of Pakistanis and Bangladeshis in the UK

For the whole population negative health behaviours and poorer health can be influenced by material deprivation [83]. Socioeconomic health inequities are found in cancer with poorer survival rates [84, 85] and overall higher mortality rates (for most sites) amongst low social class [86]. Differences in cancer survival between South Asians and other groups do not vary across deprivation levels [20]. However, socio-economic disadvantage is more prevalent amongst Pakistani and Bangladeshis with labour market discrimination [87] resulting in occupational segregation and pay gaps [88]. The effects of these deprivation differences on ethnic lines have clear implications for overall levels of physical activity and health [61]. Controlling for these persistent disadvantages in studies of mortality explains many of the differences found that suggest a mortality disadvantage for descendants of immigrants [76]. An

additional area of disadvantage faced by Pakistani and Bangladeshi communities is residential segregation [62]. The existence of segregation and living in areas of high co-ethnic density can inhibit socialisation with the majority culture and the subsequent adoption of negative health habits such as alcohol use [89].

Cancer evidence for descendants of Pakistani and Bangladeshi immigrants in the UK context remains scant due to the young age structure and subsequent low numbers of diagnoses. What has been found is that childhood cancer for UK born Pakistanis is higher than ancestral White British but for children of Bangladeshi migrants this is not the case [90]. In adulthood the infection related cancers, such as stomach and liver, that are higher amongst Pakistani and Bangladeshi immigrants do not affect UK born descendants to the same degree [21].

3 Hypotheses

Based on the previous research on this topic we hypothesise that, Pakistani, and Bangladeshi immigrants will have lower rates of cancer onset and of subsequent mortality relative to ancestral white British group. Thus, supporting the migrant mortality advantage and the effect of positive selection. Rates of onset for the descendants' group are predicted to lie between that of first-generation immigrants and the native group, owing to waning maintenance of positive health behaviours. The mortality of descendants after diagnosis, is predicted to be comparable to that of the native population, given previous evidence that has suggested the mortality advantage is not found and sometimes reversed amongst descendants. Regarding socioeconomic factors we expect the use of these to result in further divergence of the Pakistani and Bangladeshi groups from the natives. Meaning that the rates of onset and mortality become even lower than wat is observed when analysis is unadjusted.

4 Data and Methods

We use the Office for National Statistics-Longitudinal Study (ONS-LS) [91] on a study period that runs from the census of March 1971 until the end of 2016. The ONS-LS is a longitudinal 1% sample of the population of England and Wales. It links census and life event dates such as emigration, re-entry, death, and cancer diagnosis collected from National Health Service registrations and de-registrations. An individual becomes part of the ONS-LS if they are born on one of four unspecified birth dates.

4.1 Sample construction

Eligibility for inclusion is based on all members of the ONS-LS, born in 1930 or later who participate in at least one census as an adult (aged over 20). Inclusion in the analytical sample is based on a combination of country of birth and ethnicity. The majority group who we use for comparison are ancestral White British. Membership of this group requires first, that United Kingdom as country of birth be recorded in at least half the censuses they appear at, the United Kingdom being deemed to include Channel Islands, Isle of Man, Scotland, and Northern Ireland. Second, all available parental countries of birth must also be United Kingdom. Lastly, ethnic group must be White or White British as appropriate depending on the census (earlier censuses used broader ethnic category of White, compared to later censuses where White British appears specifically), again with at least half of the censuses recording this.

Immigrants are measured as such if their country of birth is Pakistan or Bangladesh in half or more of their census appearances. Due to some immigrants from Pakistan and Bangladesh being children of expatriates born under colonialism in the early 20th century, these individuals have different exposure to risk factors and therefore differences in mortality and morbidity profiles [92]. Our restriction of the sample to

those born after 1920 removes most of these possible misclassifications [5] but additionally we require the reported ethnicity to match country of birth, meaning White British individuals born in Pakistan or Bangladesh would not make the sample.

The descendants group combines those of Pakistani and Bangladeshi background. The rationale behind combining is to ensure sufficient sample size as with the younger age structure of this group fewer events are observed in the study period. This group must also consistently have United Kingdom as place of birth. We use ethnic group membership to determine this group. Ideally parental country of birth would allow for identification of second and third generation, but that information is only available for the 1971 census. Individuals whose ethnic group is Pakistani or Bangladeshi whilst being UK born are named descendants. Speculatively this group is most likely to be second generation due to the historical migration patterns of Pakistani and Bangladeshis [26], and our study's requirement of being over age 20.

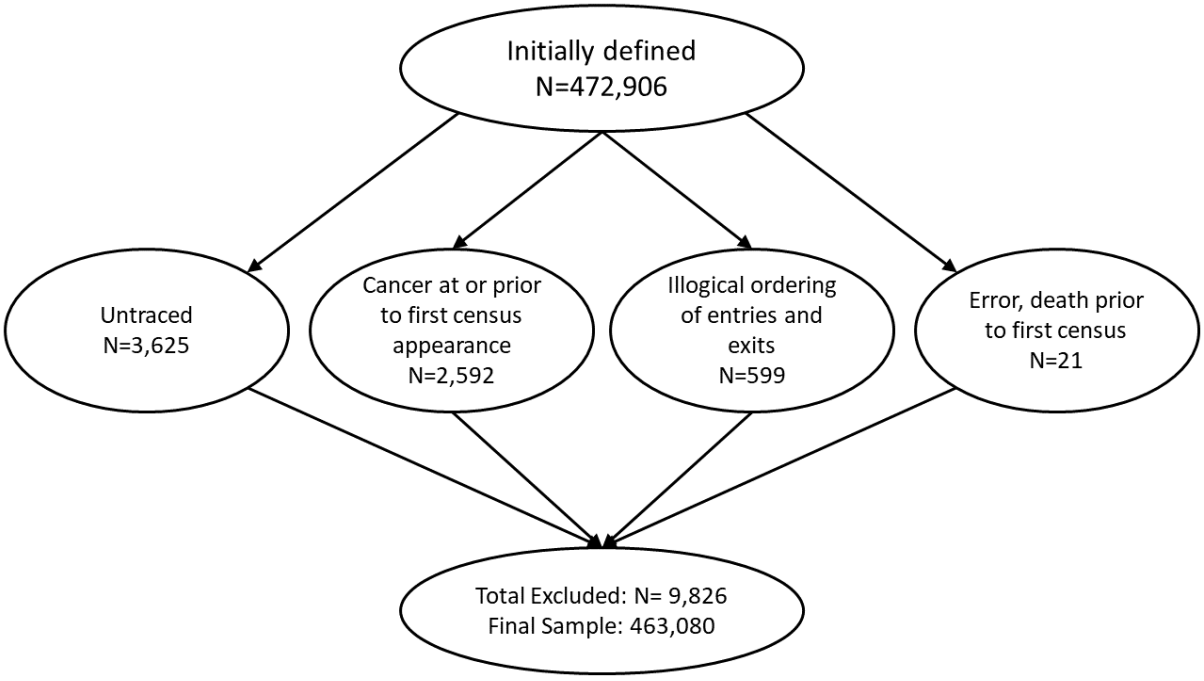
For robustness all analyses were repeated with various sample specifications, which take both stricter and less strict approaches to inclusion and consider changes in data collection and coding between censuses (see supplementary data).

4.2 Exclusions

Initially, we identified 472,906 eligible members who met the above requirements for classification of their migrant background. These individuals would have been born after 1920 and before April 1991. Meaning they should have made a census appearance at age of 20 or above. Exclusions were made on four criteria. First, being untraced, meaning there was no linkage with the national health service records meaning that life events, including cancer diagnosis, are not linked to their census record. Second, those who had cancer diagnosis prior to their first adult census were excluded as these people would continue to experience an elevated

risks of further cancer diagnosis following cancer in earlier years [93]. Third, a small number of cases were removed due to erroneous death dates which preceded their first census appearance. Lastly, we remove individuals with an illogical ordering of entries and exits. We allow for individuals to re-enter the sample following return migration and to minimise exclusions in instances where there are two re-entries and no intermediate exit date, or alternatively two exit dates and no re-entry we use the midpoint of these dates as the missing date. Figure 1 details the exclusions to reach a final sample size of 463,080.

Fig 1 Exclusion criteria and numbers excluded



Note: Initially defined are individuals born 1920-1991, who are present at one census from 1971-2011 and match the migrant origins under study.
Exclusion reasons are not mutually exclusive, hence N's do not sum to total excluded
Source: ONS-LS

4.3 Outcome Measure

Our event of interest is the incidence of first cancer. This is collected in the ONS-LS via linkage of sample members to the information provided to the English cancer

registries and the Welsh Cancer Intelligence and Surveillance Unit. These registries collect all cancer diagnosis that occur in England and Wales and are traced and matched to the ONS-LS birth dates [94]. Whilst the registry does record incidence of squamous and basal cell carcinomas, we do not include these as they are rarely a primary cause of death.

4.4 Covariates

Our main variable of interest is the migrant background, we include additional time-varying covariates. These covariates are based on answers given at the decennial census and are assumed to be fixed until the next census. Exposure time and events for each covariate can be found in Table 1. In Model 1 we control only for the ten-year period that the census covers and sex (sex stratified models were checked for robustness). Further models introduce time-varying covariates that attempt to measure levels of socioeconomic success and therefore inequalities which have been observed as associated with cancer incidence and survival. Migrant mortality studies that have used the ONS-LS have commonly used these covariates [5].

Model 2 includes a binary measure of education (having degree level education or above versus not; this dichotomy was selected to create comparable categories across censuses which have different education reporting due to changes in education policy. Social class is considered as an indicator of socioeconomic status which is associated with health inequalities including higher cancer incidence and worse survival [84, 95, 96]. Social class is measured as follows: technical and managerial, skilled, armed forces and unskilled. Model 2 also includes a measure of the spatiality with location at the time of census recorded as London, Rest of England, and Wales. We include this control to account for the devolved healthcare policies of England and Wales.

Model 3, the full model, additionally includes, marital status: never married, married, divorced, and widowed. Mortality advantages due to positive selection into marriage are observable in previous research using the ONS-LS, however cause of death specific research is less clear for Cancer [97]. We also include a variable measuring tenure; homeowner (with or without mortgage), rented and other. Other is typically a 'group home' or institutionalisation. Tenure along with social class has been seen as a reliable indicator of cancer survival and incidence in the ONS-LS [84], and is highly related to socioeconomic success.

We retain missing categories where necessary across covariates. Missing arises when sample members miss the most recent census through non-completion or being non-resident at the time. We impute where logical based on answers given at other censuses. Namely, degree level education is projected forwards and 'single never married' is projected backwards to previous census periods.

4.5 Method

Individuals are longitudinally followed through censuses every ten years. We use survival analysis to measure the exposure time before a cancer incidence, whilst resident in England and Wales. Our baseline time is measured as months since 20th birthday; however, entry to the risk set is standardised by counting exposure only after the first census appearance when aged over 20. Information on immigration before a first census appearance is obtainable, linked through the date of registration with the NHS. However, using this date would create bias since those who do register are possibly negatively health selected as they may be seeking medical treatment. Moreover, since socioeconomic variables are only collected at census date, including immigrants at their arrival date would result in more missing amongst covariates.

Individuals can exit the sample at death and emigration. Individuals with no information relating to death or emigration and who are not present at the 2011 census are deemed ‘lost to follow up’ (LTFU). These individuals are apportioned four years of exposure time following their last census appearance which is deemed the optimal amount of time based on the exit dates available in the sample [8].

To study both incidence of, and subsequent mortality from cancer we run two separate analyses using survival analysis, i.e., Cox proportional hazards models. First, we study individuals from their first census appearance until the event of first cancer registration. If they are never diagnosed with cancer then individuals are censored at death, emigration, end of the study period of December 2016 or being deemed LTFU. We allow for entry and exit to the sample based on the emigration dates and re-entry dates that are linked to NHS health records. We use the mid-point of dates where there is missing information, for example estimating the exit date when we have two re-entry dates and no exit date between them, or two exit dates but no re-entry date. Cases where the ordering was illogical such as having re-entry dates before a recorded emigration date, were removed.

To study survival, we restrict the sample to only those who experience a diagnosis of any cancer during the study period N=72,358. These individuals are followed for a maximum of 10 years from the diagnosis date, with the event of interest being death by any cause. Again, censoring happens at emigration, being LTFU and the end of the study period. 4,424 individuals are recorded as dying in the same month as their cancer diagnosis, these observations are allocated half a month of exposure time between diagnosis and death, on the assumption that there can be a maximum of one month variance between diagnosis and subsequent death, which if normally distributed would tend towards half a month. We conduct sensitivity analysis

assigning 0.03 months (approx. one day) of survival and results were not impacted. Since the baseline is now time since diagnosis, instead of age, we include a control for five-year age bands across all models. Moreover, due to different prognoses of different cancers we introduce a control for the site where the cancer is diagnosed.

Table 1 Number of events and total exposure time in 1000 person years for each covariate

	Panel A: Cancer Incidence		**Panel B: Death after diagnosis**	
Covariate	Exposure time (1000 Person Years)	Events	Exposure time (1000 Person Years)	Events
Total	11092	72358	331.4	42256
Immigrant Background				
Natives	10885.5	71926	329.7	42055
Pakistani-born	125.7	295	1.2	135
Bangladeshi-born	53.9	113	0.4	55
Descendants	26.9	24	0.1	11
Sex				
Men	5481.7	33945	132.2	22474
Women	5610.3	38413	199.2	19782
Age Band				
20-25	367.6	411	0.5	11
25-30	988.3	1849	4.8	77
30-35	1272.6	2343	12.4	208
35-40	1301.9	2477	17.3	376
40-45	1307.5	3023	18.6	834
45-50	1278.5	4304	21.5	1431
50-55	1204.4	5980	26.9	2530
55-60	1014.6	7552	32.9	3695
60-65	822.6	9373	39.5	5087
65-70	625.0	10363	43.7	6227
70-75	429.2	9656	41.8	6541
75-80	270.2	7819	35.4	6441
80-85	142.8	4751	23.5	5070
85+	66.8	2457	12.5	3728
Census Period				
1971-1981	2057.4	3364	8.9	1549
1981-1991	2715.7	8470	30.6	4845
1991-2001	3259.8	18462	75.7	10770
2001-2011	3304.1	27411	134.5	16222

2011-2016	1712.4	14651	81.7	8870
Education				
Degree	1374.7	8500	46.8	3585
No Degree	9717.3	63858	284.6	38671
Social Class				
Professional, technical, and managerial	2783.1	16443	83.8	7308
Skilled	5623.4	34323	159.4	18249
Unskilled	528.7	3657	15.9	2213
Armed Forces	29.9	76	0.4	52
Missing	2126.9	17859	71.9	14434
Location				
London	1089.5	5534	22.3	3312
Rest of England	9011.3	58986	276.0	33424
Wales	649.7	4438	20.4	2592
Unknown/Missing	341.5	3400	12.6	2928
Marital Status				
Single Never Married	2425.9	8214	34.7	3385
Married	7175.9	46925	217.4	25913
Divorced	408.4	7792	34.8	7060
Widowed	798.1	6455	33.3	3345
Missing	283.7	2972	11.2	2553
Tenure				
Owner Occupied	7355.7	48797	238.9	25926
Renter	3196.4	18830	74.0	12056
Other	170.9	930	4.2	1010
Missing	368.9	3801	14.4	3264

Source: Authors' calculations using ONS-LS

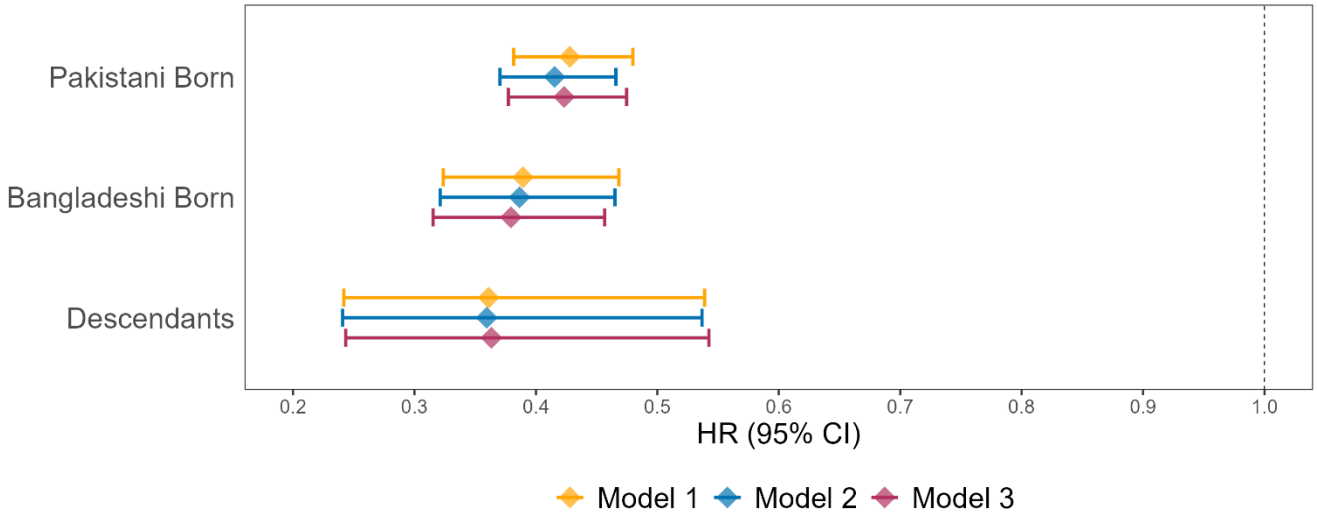
5 Results

5.1 Incidence

Figure 2 shows the hazard ratios for the incidence of cancer for each migrant background and each additive model. The reference line of one indicates the Ancestral native majority group. Incidence of cancer amongst Pakistani born, Bangladeshi born, and their descendants is substantially lower than amongst White British. Relative to White British majority, the risk of cancer onset for Pakistani immigrants is approximately 42%, for Bangladeshi born 38% and for descendants 36%.

The introduction of covariates does little to change the magnitude of the association, with much lower rates remaining for Pakistani Born, Bangladeshi Born and the descendants across all three models, see Table 2. Across all models, sex is non-significant, but there is increased risk of diagnoses in later time periods, with twice the risk of cancer in 2011-2016 compared to 1971-1981 census period. This is to be expected given the average age of sample members in later time periods and better cancer detection through screening programs. We observe a significant association of lower cancer incidence amongst those living in London compared to the rest of England and Wales. In the full model Wales and the Rest of England has approximately 12% higher relative risk of cancer onset than those living in London. There is variation in the effect of Social Class; compared to Managerial positions, Skilled workers have a slightly increased risk of cancer incidence. Whereas Unskilled and Armed forces have a lower risk, although the sample size which is coded as Armed forces is small. Those with degree level education or higher have a reduced risk of cancer diagnosis with a hazard ratio of 0.95. The final socioeconomic variable of tenure finds a substantially increased cancer incidence for those in rented accommodation compared to people living in owner occupied homes with hazard ratio of 1.15. The other category sees lower risk of 0.87. Rates by marital status show some significant differences. Those who are divorced have elevated risk of cancer incidence whilst those widowed have lower rates. No difference is observed between those never married and married.

Fig 2 Hazard Ratios of Cancer Incidence by migrant background



Note: Reference Category: Ancestral Natives = 1. 95% CIs shown.

Source: Authors' calculations using ONS-LS

Table 2 Cox proportional hazards model: Hazard Ratios of first cancer incidence in adulthood

	Model 1		Model 2		Model 3	
	Haz Ratio	95% CI	Haz Ratio	95% CI	Haz Ratio	95% CI
Immigrant Background						
Ancestral Native	1	N/A	1	N/A	1	N/A
Pakistan Born	0.43	0.38-0.48	0.42	0.37-0.47	0.42	0.38-0.47
Bangladeshi Born	0.39	0.32-0.47	0.39	0.32-0.47	0.38	0.32-0.46
Descendants	0.36	0.24-0.54	0.36	0.24-0.54	0.36	0.24-0.54
Sex						
Male	1	N/A	1	N/A	1	N/A
Female	1.01	0.99-1.02	1.00	0.98-1.01	1.01	0.99-1.02
Time Period						
1971-1981	1	N/A	1	N/A	1	N/A
1981-1991	1.28	1.23-1.33	1.27	1.22-1.32	1.29	1.23-1.34
1991-2001	1.66	1.60-1.72	1.65	1.59-1.72	1.70	1.64-1.77
2001-2011	1.93	1.86-2.00	1.94	1.87-2.02	2.00	1.93-2.08
After 2011	1.94	1.86-2.02	2.01	1.93-2.09	2.06	1.97-2.14
Location						
London			1	N/A	1	N/A

Rest of England		1.10	1.07-1.13	1.12	1.09-1.15
Wales		1.12	1.07-1.16	1.14	1.09-1.18
Missing		1.29	1.23-1.35	1.38	1.21-1.58
Social Class					
Managerial, Technical and Professional		1	N/A	1	N/A
Skilled		1.04	1.02-1.06	1.03	1.00-1.05
Unskilled		1.00	0.96-1.04	0.96	0.92-0.99
Armed Forces		0.77	0.62-0.97	0.77	0.62-0.97
Missing/Other		1.09	1.06-1.11	1.05	1.03-1.08
Education					
No Degree		1	N/A	1	N/A
Has Degree		0.94	0.91-0.96	0.95	0.92-0.97
Marital Status					
Never married				1	N/A
Married				1.02	0.99-1.04
Widowed				0.93	0.90-0.96
Divorced/Separated				1.08	1.04-1.11
Missing				0.89	0.81-0.99
Tenure					
Owner Occupied				1	N/A
Rented				1.17	1.15-1.19
Other				0.88	0.82-0.94
Missing				1.10	1.00-1.21

Source: Authors' calculations using ONS-LS

One cause of bias that could influence these findings of low cancer incidence amongst the Pakistani and Bangladeshi minority is the censoring of individuals at mortality prior to a cancer diagnosis. In theory these premature deaths are found in the unhealthiest individuals who would be the most likely to develop cancer later in the life course but are never observed experiencing the event. Meaning that the survivors into later ages are part of a select healthier group. Considering there is evidence of elevated risk of cardiovascular disease amongst Pakistanis and Bangladeshis [60, 98] and also more deaths at younger ages amongst descendants [76] it is possible that the minorities under study at older ages where cancer is more

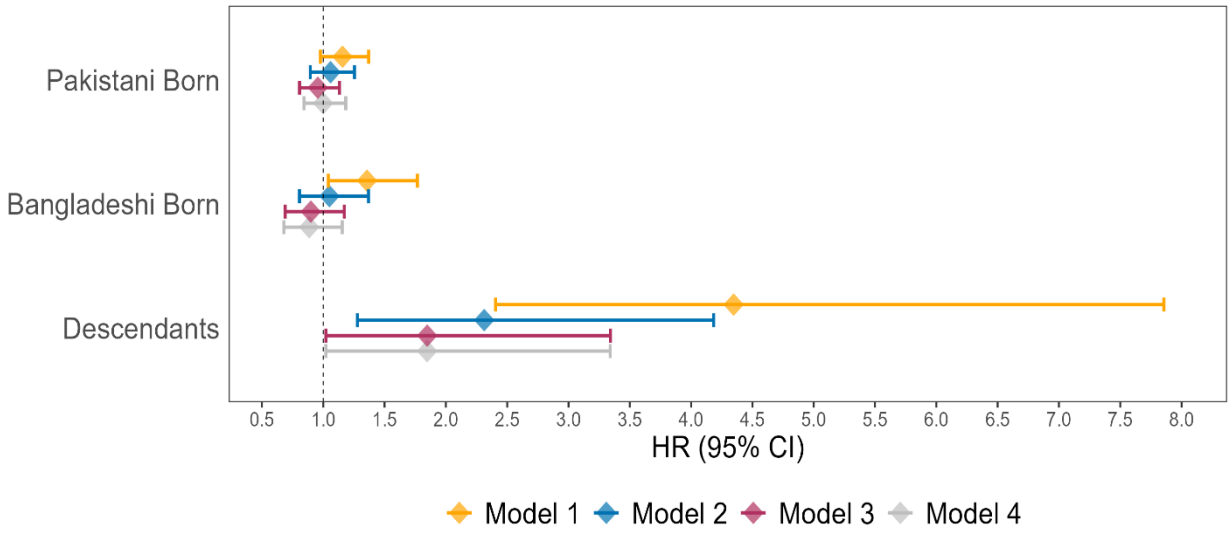
prevalent are healthier. To ensure that the low cancer incidence rates are not due to premature deaths in migrant populations we used the same model specifications as Model 3 above with the outcome of mortality prior to a cancer diagnosis. The hazard ratios of mortality prior to cancer are significantly lower for Pakistani born (95% CI 0.59-0.69) and Bangladeshi born (95% CI 0.55-0.71) with no significant difference for descendants (95% CI 0.74-1.51). Suggesting that if there is any bias due to the censoring of unhealthy individuals prior to cancer diagnosis it is influencing the native reference population more so than the minority populations under study.

5.2 Mortality after Diagnosis

Our secondary analysis focussing on all-cause mortality in the ten years following diagnosis can be seen graphically in Figure 3. Model 1 contains controls for sex, age and time-period, Model 2 adds in a control for type of cancer. Model 3 adds location, social class, and education. Lastly, Model 4 considers tenure and marital status. Across all models there is no significance difference in the hazard ratio after a cancer diagnosis, for Pakistani born individuals, compared to the ancestral native reference group. For Bangladeshi born Model 1 suggests some elevated risk of death yet controlling for type of cancer in Model 2 explains that gap. Before the socioeconomic controls are added there is some non-significant evidence that mortality after diagnosis is slightly higher for the foreign born, however this is reversed once socioeconomic controls are introduced, and never significant. For the descendants Model 1 suggests a substantially higher risk of all-cause mortality following a cancer diagnosis. Again, the strength of this association is reduced heavily through introducing the type of cancer and further still with the introduction of socioeconomic variables. The relative risk of mortality in the full model is weakly significant with a hazard ratio of 1.85. We caution though that, for descendants in

particular, statistical power is limited due to a small number of cancer onset as the previous analysis showed.

Fig 3 Hazard Ratios of death following cancer incidence by migrant background



Note: Reference Category: Ancestral White British =-1. 95% CIs shown
Source: Authors' calculations using ONS-LS

Results for all covariates can be seen in Table 3, patterns are in line with expectations, risk of death after diagnosis increases with age and over time periods the risk has deceased, a sign of the better treatment and medical developments which increased cancer survival. We see a significant effect for sex, the risk of all-cause mortality for women is over 20% lower compared to similar men. Differences by location are limited, London dwellers appear to have a higher relative risk of mortality compared to the rest of England. Gradients by social class are apparent, skilled, and unskilled both have higher relative risk of mortality after diagnosis compared to those defined in the most prestigious social class of technical, managerial and professional. Moreover, amongst those who obtain degree level education the risk of death following diagnosis is reduced by 12%. Lastly, we see an association between marital status and death following diagnosis with both married

574 and divorced individuals having lower relative risk compared to those who have

575 never married.

576 **Table 3** Cox proportional hazards model: Hazard Ratios of mortality following diagnosis

	Model 1		Model 2		Model 3		Model 4	
	Haz Ratio	95% CI	Haz Ratio	95% CI	Haz Ratio	95% CI	Haz Ratio	95% CI
Immigrant Background								
Ancestral Native	1	N/A	1	N/A	1	N/A	1	N/A
Pakistan Born	1.16	0.98-1.37	1.06	0.89-1.25	0.96	0.81-1.13	1.00	0.84-1.18
Bangladeshi Born	1.36	1.04-1.77	1.05	0.81-1.37	0.90	0.69-1.17	0.89	0.68-1.15
Descendants	4.35	2.41-7.86	2.31	1.28-4.18	1.85	1.02-3.34	1.85	1.02-3.34
Age Band								
20-25	0.16	0.09-0.29	0.24	0.13-0.43	0.23	0.13-0.41	0.20	0.11-0.36
25-30	0.16	0.12-0.20	0.22	0.18-0.28	0.22	0.18-0.28	0.20	0.16-0.25
30-35	0.21	0.18-0.24	0.28	0.24-0.32	0.28	0.24-0.32	0.26	0.23-0.30
35-40	0.29	0.26-0.32	0.35	0.32-0.39	0.35	0.31-0.39	0.34	0.31-0.38
40-45	0.53	0.49-0.57	0.60	0.56-0.65	0.60	0.55-0.65	0.59	0.55-0.64
45-50	0.73	0.69-0.78	0.77	0.72-0.82	0.77	0.72-0.82	0.77	0.72-0.82
50-55	1	N/A	1	N/A	1	N/A	1	N/A
55-60	1.25	1.19-1.31	1.17	1.11-1.23	1.16	1.10-1.22	1.16	1.10-1.22
60-65	1.51	1.44-1.59	1.39	1.32-1.46	1.36	1.30-1.43	1.36	1.30-1.43
65-70	1.81	1.73-1.90	1.65	1.57-1.73	1.59	1.52-1.67	1.59	1.52-1.67
70-75	2.15	2.05-2.25	1.97	1.88-2.07	1.88	1.79-1.97	1.87	1.78-1.96
75-80	2.74	2.61-2.88	2.62	2.49-2.75	2.41	2.30-2.54	2.39	2.27-2.51
80-85	3.74	3.55-3.93	3.62	3.44-3.81	3.14	2.97-3.31	3.09	2.93-3.26
85+	5.55	5.26-5.86	5.47	5.18-5.78	4.67	4.41-4.94	4.49	4.24-4.76
Sex								
Male	1	N/A	1	N/A	1	N/A	1	N/A
Female	0.77	0.75-0.78	0.84	0.83-0.86	0.80	0.78-0.82	0.79	0.77-0.81
Time Period								
1971-1981	1	N/A	1	N/A	1	N/A	1	N/A
1981-1991	0.69	0.65-0.73	0.75	0.70-0.79	0.74	0.69-0.78	0.74	0.70-0.79
1991-2001	0.47	0.45-0.50	0.57	0.54-0.61	0.58	0.54-0.61	0.59	0.56-0.62
2001-2011	0.30	0.29-0.32	0.40	0.38-0.42	0.41	0.39-0.44	0.42	0.40-0.45
After 2011	0.25	0.23-0.26	0.32	0.30-0.34	0.36	0.34-0.39	0.37	0.34-0.39
Cancer Type								
Colorectal			1	N/A	1	N/A	1	N/A
Bronchus/Lung			3.45	3.32-3.59	3.36	3.24-3.50	3.27	3.15-3.40
Prostate			0.56	0.53-0.59	0.57	0.54-0.60	0.58	0.55-0.61

Kidney	1.24	1.15-1.34	1.25	1.16-1.35	1.24	1.15-1.33
Bladder	0.78	0.74-0.83	0.78	0.74-0.83	0.78	0.74-0.83
Stomach	2.53	2.39-2.69	2.47	2.33-2.62	2.43	2.29-2.58
Non-Hodgkin's lymphoma	1.01	0.95-1.08	1.02	0.95-1.09	1.02	0.95-1.09
Melanoma/Skin	0.51	0.47-0.56	0.53	0.48-0.58	0.54	0.49-0.59
Pancreatic	4.82	4.52-5.13	4.82	4.52-5.14	4.82	4.52-5.14
Leukaemia	1.28	1.19-1.38	1.28	1.19-1.38	1.29	1.20-1.38
Oesophageal	2.84	2.66-3.03	2.80	2.62-2.99	2.74	2.57-2.93
Oral	1.05	0.96-1.13	1.02	0.94-1.11	1.00	0.92-1.09
Brain	4.24	3.95-4.56	4.32	4.01-4.64	4.33	4.02-4.65
Myeloma	1.64	1.50-1.79	1.64	1.50-1.79	1.64	1.50-1.79
Liver	4.15	3.77-4.56	4.11	3.73-4.52	4.12	3.75-4.54
Thyroid	0.59	0.49-0.71	0.60	0.49-0.72	0.61	0.51-0.73
Breast	0.65	0.62-0.68	0.66	0.63-0.69	0.66	0.63-0.69
Uterine	0.62	0.57-0.68	0.62	0.57-0.68	0.62	0.57-0.68
Ovary	1.57	1.49-1.69	1.58	1.47-1.69	1.59	1.48-1.70
Cervical	0.95	0.87-1.05	0.93	0.84-1.02	0.91	0.83-1.00
Other malignant neoplasm	0.57	0.55-0.59	0.57	0.55-0.59	0.57	0.55-0.59
Location						
London			1	N/A	1	N/A
Rest of England			0.93	0.90-0.97	0.96	0.92-0.99
Wales			0.94	0.90-0.99	0.98	0.93-1.03
Missing			1.19	1.13-1.26	1.17	1.00-1.36
Social Class						
Managerial, technical, and professional			1	N/A	1	N/A
Skilled			1.14	1.11-1.18	1.12	1.09-1.15
Unskilled			1.27	1.21-1.34	1.20	1.14-1.26
Armed Forces			1.08	0.82-1.42	1.05	0.80-1.38
Missing/Other			1.39	1.34-1.44	1.31	1.27-1.36
Education						
No Degree			1	N/A	1	N/A
Has Degree			0.86	0.83-0.89	0.88	0.84-0.91
Marital Status						
Never married					1	N/A
Married					0.85	0.82-0.88
Widowed					0.96	0.92-1.00
Divorced/Separated					0.90	0.86-0.95
Missing					0.76	0.68-0.84
Tenure						

Owner Occupied		1	N/A
Rented		1.19	1.16-1.22
Other		1.73	1.62-1.84
Missing		1.30	1.17-1.45

Source: Authors' calculations using ONS-LS

The mortality observed in the secondary analysis is all-cause mortality. This does not necessarily give an accurate picture of mortality from cancer. We use the International Classifications of Diseases (ICD) Code which is available in the ONS-LS through linkage to death registrations. The ONS-LS exists over three revisions of ICD codes, 1971 to 1981 is ICD-8, 1981–1999 is ICD-9 and from 2000 onwards has been ICD-10. We harmonise these ICD codes across the sample to create broad categories of deaths [5] enabling us to dichotomise primary cause of death into either from cancer or another cause. In total approximately 25% of the deaths observed in the ten years following cancer diagnosis were not primarily caused by cancer. Thus, we use a semi-parametric competing risk model where the risk of death from cancer competes with the risk of death from any other cause [99]. Covariates are specified in the same way as Model 4. We continue to find non-significant differences between ancestral natives, and both Pakistani Born (95% CI 0.69-1.06) and Bangladeshi Born (95% CI 0.84-1.34). Moreover, the hazard ratio for descendants whilst remaining positive becomes non-significant (HR=1.48, 95% CI 0.66-3.32).

5.3 Sensitivity Analysis

We ran several sensitivity analyses to investigate different sample specifications, descriptions, and subsequent sample sizes of which can be seen in Online Resource 1. None of these specifications altered the results. Further since the use of ‘missing’ as a category generates scepticism in health research [100], we repeated the

analysis using only complete cases, the results hold with only small changes to the magnitude, see Online Resource 2. Further we consider different ways to capture socioeconomic status by using economic position as a covariate instead of, and as well as, social class. Once more, the differences are minimal, see Online Resource 3.

Lastly, we considered sex stratified models (Online Resources 4 & 5). The socioeconomic determinants of health are likely to differ by sex [79], meaning that susceptibility to cancer may also differ. Sex stratified models still find stable results, albeit with larger confidence intervals. Thus, due to the low number of events and data restrictions we maximise sample size and statistical power by using a non-stratified sample, with sex as a covariate.

6 Discussion

This study supports previous findings of low cancer mortality amongst Pakistani and Bangladeshi born individuals [5, 19]. We can add to this more certainty that it is driven by the lower incidence and not by a better survival after diagnosis. We find evidence that suggests amongst descendants the advantage of lower cancer incidence persists. However, in the ten years following diagnosis there is little evidence to suggest that mortality rates differ across any groups. There is some weak evidence that mortality rates after a cancer diagnosis may be elevated for the descendants of Pakistani and Bangladeshi immigrants, however the small number of events in this group must be considered when interpreting the findings for that group. Overall, this finding supports the concept of the healthy migrant effect [3] and that ill-health which has given rise to the theory of the migrant health mortality paradox [4] is unlikely to be driven by cancer.

Our first research aim was to uncover differences in onset of cancer between the groups. We use event history analysis and Cox proportional hazards models and do find these clear differences. We build on previous research that has identified lower incidence amongst South Asian as a broad group or that has identified lower incidence for specific cancer sites [16, 22–24]. The analysis finds that there is a considerable advantage for immigrants from Pakistan and Bangladesh, and their descendants when it comes to the of cancer incidence in England and Wales, when it comes to the onset of cancer. We speculate that there are both environmental effects related to the overall lower burden of cancer found in the origin countries [101]. Alongside a maintenance of healthy behaviours [34, 102], which has been found in relation to, alcohol consumption and tobacco smoking [50, 51]. The low levels of incidence amongst descendants suggests a combination of inheritance of the positive selection from their immigrant parents and a continuation of these healthy behaviours [14]. We could speculate that the lack of change in cancer incidence between generations could a reflection of the low socialisation and assimilation with the majority population which has entrenched behavioural norms [89], however we do acknowledge that some studies do find worse health behaviours in the descendants of Pakistani and Bangladeshi immigrants [51].

Previous research finds excess rates of cardiovascular disease [5, 60, 98], therefore, we investigated if survival bias is a reason behind these findings. Yet, the attrition of individuals due to deaths before cancer incidence still suggests the existence of the healthy migrant effect for the foreign born and that mortality is not different for the descendants group compared to ancestral natives either. Salmon bias and data artefacts could also be explanations of this apparent benefit, but we believe our results are robust to this based on previous analysis of this data set [8, 30], we

649 employed similar analytical strategy and specified risk time to those lost to follow-up
650 in line with previous findings.

651 Research question two sought to find if there were differences in all-cause mortality
652 in the ten-years after cancer diagnosis. Here we find that the healthy migrant effect is
653 only present as an advantage in cancer incidence, the risk of death was not different
654 between observed groups. Health protective behaviours or genetic benefits do not
655 appear to provide a relative advantage in survival after onset. Perhaps a factor of the
656 universal health care system of England and Wales acting as an equaliser across
657 socioeconomic boundaries of society [103]. The universal health care coverage
658 includes screening and preventative care, however previous research has identified
659 that they are less utilised by Pakistani and Bangladeshis [64, 104]. Whilst this might
660 be due to a, potentially, justified belief that cancer is less prevalent in their
661 communities [67] it may lead to late detection and therefore worse survival rates.

662 We find that all-cause mortality for descendants following a cancer diagnosis is
663 higher than that of ancestral natives. However, when we use a competing risk model
664 to identify cancer specific mortality, this difference is attenuated. This provides
665 evidence that the previous observed higher mortality in descendants [76] is not a
666 factor of worse cancer outcomes and is related to causes outwith cancer. This
667 compliments results from other contexts which has found that the reversal of the
668 healthy migrant effect is due to rises in mortality caused by external factors not
669 illness [81].

670 Our final research question was concerned with how controlling for additional
671 socioeconomic covariates would influence the results. Overall, the covariates in our
672 models generally follow the expected patterns, with a positive association of

socioeconomic success and lower risk of cancer incidence and subsequent mortality [84, 96]. Clear gradients exist across age in line with expectations, cancer diagnoses and deaths become more common in older ages [105]. Over time there are more diagnoses of cancer, attributed to better screening methods, but less risk of death due to medical interventions improving survivability for all cancers [106].

The use of socioeconomic variables does little to change the magnitude of the results for cancer incidence for any of the observed groups. Given the relatively worse socioeconomic outcomes of Pakistani and Bangladeshis [87, 107] this may be deemed surprising, however further supports that it is a positive selection effect and health behaviours that drive low onset of neoplasms. When predicting the all-cause mortality risk after cancer onset the inclusion of socioeconomic controls have little impact on risk for the Pakistani and Bangladeshi born groups. For descendants, these covariates do explain much of the elevated risk after diagnosis like prior findings [76]. However, this analysis has a relatively small number of cancer onsets, we await a time where enough descendants have reached peak cancer and mortality ages, to truly see if accumulated disadvantage across the life course has negatively affected their longevity.

Our study is not without limitations, whilst we use a rich source of representative administrative data, there is no information around the health behaviour of the individuals in the study. Census questions do not pertain to behaviours which could be considered risk factors in cancer incidence. Therefore, we can only speculate on the persistence of health behaviours as a reason for low cancer incidence amongst the Pakistani and Bangladeshi group. Further, despite the large sample size we have limited scope to talk about the types of cancer that effects these groups, we lack statistical power to make statements about differences in the types of cancer that

both immigrants and descendants develop and how it differs from ancestral natives and each other.

To our knowledge we are the first to research both the incidence and survival of cancer in Pakistani and Bangladeshi immigrants and their descendants in a way that considers the potential generational differences. Differences between Pakistani and Bangladeshi immigrants and their descendants are limited for cancer onset and inconclusive for subsequent mortality. These findings are robust to a wide variety of sample specifications. Our approach and analysis add validity to previous studies which have combined Pakistani and Bangladeshi immigrant groups, with similar hazard ratios for both groups. Unfortunately, due to the lower event counts we are unable to investigate the descendants specifically as separate origin groups and with some evidence of divergence between Pakistanis and Bangladeshis in other life domains [73] how the health of second generation adults in these specific groups differs remains to be seen.

7 Acknowledgements

Omitted due to potential disclosure.

8 References

1. Shor E, Roelfs D. A Global Meta-analysis of the Immigrant Mortality Advantage. Int Migr Rev. 2021;55:999–1028.

2. Aldridge RW, Nellums LB, Bartlett S, Barr AL, Patel P, Burns R, et al. Global patterns of mortality in international migrants: a systematic review and meta-analysis. Lancet. 2018;392:2553–66.

3. Razum O. Migrant Mortality, Healthy Migrant Effect. In: Kirch W, editor. Encyclopedia of Public Health. Springer Dordrecht; 2008. p. 933–5.

4. Wallace M, Darlington-Pollock F. Poor health, low mortality? Paradox found among immigrants in England and Wales. Popul Space Place. 2022;28:1–15.

5. Wallace M, Kulu H. Mortality among immigrants in England and Wales by major causes of death, 1971-2012: A longitudinal analysis of register-based data. Soc Sci Med. 2015;147:209–21.

6. Scott AP, Timæus IM. Mortality differentials 1991-2005 by self-reported ethnicity: Findings from the ONS Longitudinal Study. J Epidemiol Community Health. 2013;67:743–50.

7. Wild SH, Fischbacher C, Brock A, Griffiths C, Bhopal R. Mortality from all causes and circulatory disease by country of birth in England and Wales 2001-2003. J Public Health (Bangkok). 2007;29:191–8.

8. Wallace M, Kulu H. Low immigrant mortality in England and Wales: A data artefact? Soc Sci Med. 2014;120:100–9.

9. Evandrou M, Falkingham J, Feng Z, Vlachantoni A. Ethnic inequalities in limiting health and self-reported health in later life revisited. J Epidemiol Community Health. 2016;70:653–62.

10. Rees PH, Wohland PN, Norman PD. The estimation of mortality for ethnic groups at local scale within the United Kingdom. Soc Sci Med. 2009;69:1592–607.

11. Wohland P, Rees P, Nazroo J, Jagger C. Inequalities in healthy life expectancy between ethnic groups in England and Wales in 2001. Ethn Heal. 2015;20:341–53.

12. Charles A. Health and well-being. In: Finney N, Nazroo J, Becares L, Kapadia D, Shlomo N, editors. Effective Directors: The Right Questions to Ask (QTA). Bristol University Press; 2021. p. 63–7.

13. Institute for Health Metrics and Evaluation. GBD Results. 2020. https://vizhub.healthdata.org/gbd-results/. Accessed 26 Apr 2023.

14. Anand P, Kunnumakara AB, Sundaram C, Harikumar KB, Tharakan ST, Lai OS, et al. Cancer is a preventable disease that requires major lifestyle changes. Pharm Res. 2008;25:2097–116.

15. Brown KF, Rumgay H, Dunlop C, Ryan M, Quartly F, Cox A, et al. The fraction of cancer attributable to modifiable risk factors in England, Wales, Scotland, Northern Ireland, and the United Kingdom in 2015. Br J Cancer. 2018;118:1130–41.

16. Delon C, Brown KF, Payne NWS, Kotrotsios Y, Vernon S, Shelton J. Differences in cancer incidence by broad ethnic group in England, 2013–2017. Br J Cancer. 2022;126:1765–73.

17. Winter H, Cheng KK, Cummins C, Maric R, Silcocks P, Varghese C. Cancer incidence in the south Asian population of England (1990–92). Br J Cancer. 1999;79:645–54.

18. Harding S, Rosato M. Cancer incidence among first generation Scottish, Irish, West Indian and South Asian migrants living in England and Wales. Ethn Heal. 1999;4:83–92.

19. Wild SH, Fischbacher C, Brock A, Griffiths C, Bhopal R. Mortality from all cancers and lung, colorectal, breast and prostate cancer by country of birth in England and Wales, 2001-2003. Br J Cancer. 2006;94:1079–85.

20. Maringe C, Li R, Mangtani P, Coleman MP, Rachet B. Cancer survival differences between South Asians and non-South Asians of England in 1986-2004, accounting for age at diagnosis and deprivation. Br J Cancer. 2015;113:173–81.

21. Mangtani P, Maringe C, Rachet B, Coleman MP, Dos Santos Silva I. Cancer mortality in ethnic South Asian migrants in England and Wales (1993-2003): Patterns in the overall population and in first and subsequent generations. Br J Cancer. 2010;102:1438–43.

771 22. Jack RH, Davies EA, Møller H. Prostate cancer incidence, stage at diagnosis, treatment
772 and survival in ethnic groups in South-East England. BJU Int. 2010;105:1226–30.

773 23. Jack RH, Davies EA, Møller H. Breast cancer incidence, stage, treatment and survival in
774 ethnic groups in South East England. Br J Cancer. 2009;100:545–50.

775 24. Jack RH, Konfortion J, Coupland VH, Kocher HM, Berry DP, Allum W, et al. Primary liver
776 cancer incidence and survival in ethnic groups in England, 2001-2007. Cancer Epidemiol.
777 2013;37:34–8.

778 25. Bhopal R, Unwin N, White M, Yallop J, Walker L, Alberti KGMM, et al. Populations :
779 Cross Sectional Study. BMJ. 1999;319:215–20.

780 26. Peach C. South Asian migration and settlement in Great Britain, 1951-2001. Contemp
781 South Asia. 2006;15:133–46.

782 27. Anson J. The migrant mortality advantage: A 70 month follow-up of the Brussels
783 population. Eur J Popul. 2004;20:191–218.

784 28. Vanthomme K, Vandenheede H. Trends in Belgian cause-specific mortality by migrant
785 origin between the 1990s and the 2000s. BMC Public Health. 2019;19.

786 29. Honkaniemi H, Bacchus-Hertzman J, Fritzell J, Rostila M. Mortality by country of birth in
787 the Nordic countries - A systematic review of the literature. BMC Public Health. 2017;17:1–
788 12.

789 30. Wallace M, Kulu H. Can the salmon bias effect explain the migrant mortality advantage
790 in England and Wales? Popul Space Place. 2018;24.

791 31. Wallace M, Khlat M, Guillot M. Mortality advantage among migrants according to
792 duration of stay in France, 2004-2014. BMC Public Health. 2019;19.

793 32. Chiswick BR, Lee YL, Miller PW. Immigrant selection systems and immigrant health.
794 Contemp Econ Policy. 2008;26:555–78.

795 33. Constant AF. The healthy immigrant paradox and health convergence. CESifo DICE
796 Rep. 2017;15:20–5.

797 34. Abraido-Lanza AF, Dohrenwend BP, Ng-Mak DS, Blake Turner J, Abraido-Lanza AF.
798 The Latino Mortality Paradox: A Test of the “Salmon Bias” and Healthy Migrant Hypotheses
799 Requests for reprints and correspondence should be sent to .I. Am J Public Health.
800 1999;89:1543–8.

801 35. Constant AF, Milewski N. Self-selection in physical and mental health among older intra-
802 European migrants. J Econ Ageing. 2021;19:100322.

803 36. Juárez SP, Drefahl S, Dunlavy A, Rostila M. All-cause mortality, age at arrival, and
804 duration of residence among adult migrants in Sweden: A population-based longitudinal
805 study. SSM - Popul Heal. 2018;6:16–25.

806 37. Wallace M, Wilson B. Migrant Mortality Advantage Versus Origin and the Selection
807 Hypothesis. Popul Dev Rev. 2019;45:767–94.

808 38. Goettler A. What Happens to the Healthy Immigrant Later in Life? – The Health of
809 (Forced) Migrants Through the Life Course. In: Crepaz K, Becker U, Wacker E, editors.
810 Health in Diversity – Diversity in Health. Wiesbaden: Springer VS; 2020. p. 103–19.

811 39. Spallek J, Zeeb H, Razum O. What do we have to know from migrants’ past exposures
812 to understand their health status? A life course approach. Emerg Themes Epidemiol.
813 2011;8:6.

40. Shor E, Roelfs D, Vang ZM. The “Hispanic mortality paradox” revisited: Meta-analysis and meta-regression of life-course differentials in Latin American and Caribbean immigrants’ mortality. Soc Sci Med. 2017;186:20–33.

41. Nazroo J. The structuring of ethnic inequalities in health: Economic position, racial discrimination, and racism. Am J Public Health. 2003;93:277–84.

42. Wallace S, Nazroo J, Bécares L. Cumulative Effect of Racial Discrimination on the Mental Health of Ethnic Minorities in the United Kingdom. Am J Public Health. 2016;106:1294–300.

43. Monti A, Drefahl S, Mussino E, Härkönen J. Over-coverage in population registers leads to bias in demographic estimates. Popul Stud (NY). 2020;74:451–69.

44. Wallace M, Wilson B. Age variations and population over-coverage: Is low mortality among migrants merely a data artefact? Popul Stud (NY). 2022;76:81–98.

45. Andersson G, Drefahl S. Long-Distance Migration and Mortality in Sweden: Testing the Salmon Bias and Healthy Migrant Hypotheses. Popul Space Place. 2017;23:e2032.

46. Lindor NM, McMaster ML, Lindor CJ, Greene MH. Concise Handbook of Familial Cancer Susceptibility Syndromes - Second Edition. JNCI Monogr. 2008;2008:3–93.

47. Bygren LO, Tinghög P, Carstensen J, Edvinsson S, Kaati G, Pembrey ME, et al. Change in paternal grandmotherśearly food supply influenced cardiovascular mortality of the female grandchildren. BMC Genet. 2014;15:2–7.

48. Rumgay H, Shield K, Charvat H, Ferrari P, Sornpaisarn B, Obot I, et al. Global burden of cancer in 2020 attributable to alcohol consumption: a population-based study. Lancet Oncol. 2021;22:1071–80.

49. Jemal A, Center MM, DeSantis C, Ward EM. Global Patterns of Cancer Incidence and Mortality Rates and Trends. Cancer Epidemiol Biomarkers Prev. 2010;19:1893–907.

50. Karlsen S, Millward D, Sandford A. Investigating ethnic differences in current cigarette smoking over time using the health surveys for England. Eur J Public Health. 2012;22:254–6.

51. Wang S, Li S. Exploring generational differences of british ethnic minorities in smoking behavior, frequency of alcohol consumption, and dietary style. Int J Environ Res Public Health. 2019;16.

52. Harrison L, Sutton M, Gardiner E. Ethnic differences in substance use and alcohol-use-related mortality among first generation migrants to england and wales. Subst Use Misuse. 1997;32:849–76.

53. Abed Al Ahad M, Sullivan F, Demšar U, Melhem M, Kulu H. The effect of air-pollution and weather exposure on mortality and hospital admission and implications for further research: A systematic scoping review. PLoS One. 2020;15 10 October:1–24.

54. Abed Al Ahad M, Demšar U, Sullivan F, Kulu H. Does Long - Term Air Pollution Exposure Affect Self - Reported Health and Limiting Long Term Illness Disproportionately for Ethnic Minorities in the UK ? A Census - Based Individual Level Analysis. Appl Spat Anal Policy. 2022;:1557–82.

55. Clark K, Drinkwater S. Enclaves, neighbourhood effects and employment outcomes : Ethnic minorities in England and Wales. J Popul Econ. 2002;15:5–29.

56. Finney N. Understanding ethnic differences in the migration of young adults within Britain from a lifecourse perspective. Trans Inst Br Geogr. 2011;36:455–70.

57. Aune D, Giovannucci E, Boffetta P, Fadnes LT, Keum NN, Norat T, et al. Fruit and vegetable intake and the risk of cardiovascular disease, total cancer and all-cause mortality-A systematic review and dose-response meta-analysis of prospective studies. Int J Epidemiol. 2017;46:1029–56.

58. Renehan AG, Tyson M, Egger M, Heller RF, Zwahlen M. Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies. Lancet. 2008;371:569–78.

59. Jamal A. Food consumption among ethnic minorities: The case of British-Pakistanis in Bradford, UK. Br Food J. 1998;100:221–7.

60. Gholap N, Davies M, Patel K, Sattar N, Khunti K. Type 2 diabetes and cardiovascular disease in South Asians. Prim Care Diabetes. 2011;5:45–56.

61. Fischbacher C, Hunt S, Alexander L. How physically active are South Asians in the United Kingdom? A literature review. J Public Health (Bangkok). 2004;26:250–8.

62. Shankley W, Finney N. Ethnic minorities and housing in Britain. In: Byrne B, Alexander C, Khan O, Nazroo J, Shankley W, editors. Ethnicity, Race and Inequality in the UK: State of the Nation. 2020. p. 149–66.

63. Li Y, Heath A. Persisting disadvantages: a study of labour market dynamics of ethnic unemployment and earnings in the UK (2009–2015). J Ethn Migr Stud. 2020;46:857–78.

64. Szczepura A, Price C, Gumber A. Breast and bowel cancer screening uptake patterns over 15 years for UK south Asian ethnic minority populations, corrected for differences in socio-demographic characteristics. BMC Public Health. 2008;8:1–15.

65. Robb K, Wardle J, Stubbings S, Ramirez A, Austoker J, MacLeod U, et al. Ethnic disparities in knowledge of Cancer Screening Programmes in the UK. J Med Screen. 2010;17:125–31.

66. Robb K, Solarin I, Power E, Atkin W, Wardle J. Attitudes to colorectal cancer screening among ethnic minority groups in the UK. BMC Public Health. 2008;8:1–11.

67. Vrinten C, Wardle J, Marlow LAV. Cancer fear and fatalism among ethnic minority women in the United Kingdom. Br J Cancer. 2016;114:597–604.

68. Patel M, Phillips-Caesar E, Boutin-Foster C. Barriers to lifestyle behavioral change in migrant South Asian populations. J Immigr Minor Heal. 2012;14:774–85.

69. Szczepura A. Access to health care for ethnic minority populations. Postgrad Med J. 2005;81:141–7.

70. Palmer CK, Thomas MC, McGregor LM, Von Wagner C, Raine R. Understanding low colorectal cancer screening uptake in South Asian faith communities in England - A qualitative study Health behavior, health promotion and society. BMC Public Health. 2015;15:1–7.

71. Arnold M, Razum O, Coebergh JW. Cancer risk diversity in non-western migrants to Europe: An overview of the literature. Eur J Cancer. 2010;46:2647–59.

72. Luthra RR, Platt L. The Changing Face of Pakistani Migration to the United Kingdom. AAPI Nexus Policy, Pract Community. 2017;15:15–56.

73. Georgiadis A, Manning A. Change and continuity among minority communities in Britain. J Popul Econ. 2011;24:541–68.

74. Dustmann C, Weiss Y. Return migration: Theory and empirical evidence from the UK. Br J Ind Relations. 2007;45:236–56.

902 75. Gudmundsson H, Gudbjartsson DF, Kong A. Inheritance of human longevity in Iceland.
903 Eur J Hum Genet. 2000;8:743–9.

904 76. Wallace M. Adult mortality among the descendants of immigrants in England and Wales:
905 does a migrant mortality advantage persist beyond the first generation? J Ethn Migr Stud.
906 2016;42:1558–77.

907 77. Harding S. Mortality of Migrants from the Indian Subcontinent to England and Wales:
908 Effect of Duration of Residence. Epidemiology. 2003;14:287–92.

909 78. Nazroo J. The Health of Britain's Ethnic Minorities: Findings from a National Survey.
910 Policy Studies Institute; 1997.

911 79. Zufferey J. Investigating the migrant mortality advantage at the intersections of social
912 stratification in Switzerland: The role of vulnerability. Demogr Res. 2016;34:899–926.

913 80. Tarnutzer S, Bopp M, the SNC study group. Healthy migrants but unhealthy offspring? A
914 retrospective cohort study among Italians in Switzerland. BMC Public Health. 2012;12.

915 81. Wallace M. Mortality Advantage Reversed: The Causes of Death Driving All-Cause
916 Mortality Differentials Between Immigrants, the Descendants of Immigrants and Ancestral
917 Natives in Sweden, 1997–2016. Eur J Popul. 2022;38:1213–41.

918 82. Smith NR, Kelly YJ, Nazroo JY. Intergenerational continuities of ethnic inequalities in
919 general health in England. J Epidemiol Community Health. 2009;63:253–8.

920 83. Marmot MG. Health equity in England: The Marmot review 10 years on. BMJ. 2020;368
921 February:1–4.

922 84. Sloggett A, Young H, Grundy E. The association of cancer survival with four
923 socioeconomic indicators: A longitudinal study of the older population of England and Wales
924 1981-2000. BMC Cancer. 2007;7:1–8.

925 85. Woods LM, Rachet B, Coleman MP. Origins of socio-economic inequalities in cancer
926 survival: A review. Ann Oncol. 2006;17:5–19.

927 86. Power C, Hyppönen E, Smith GD. Socioeconomic position in childhood and early adult
928 life and risk of mortality: A prospective study of the mothers of the 1958 British birth cohort.
929 Am J Public Health. 2005;95:1396–402.

930 87. Larsen EN, Di Stasio V. Pakistani in the UK and Norway: different contexts, similar
931 disadvantage. Results from a comparative field experiment on hiring discrimination. J Ethn
932 Migr Stud. 2019. https://doi.org/10.1080/1369183X.2019.1622777.

933 88. Brynin M, Güveli A. Understanding the ethnic pay gap in Britain. Work Employ Soc.
934 2012;26:574–87.

935 89. Bécares L, Nazroo J, Stafford M. The ethnic density effect on alcohol use among ethnic
936 minority people in the UK. J Epidemiol Community Health. 2011;65:20–5.

937 90. Sayeed S, Barnes I, Ali R. Childhood cancer incidence by ethnic group in England, 2001-
938 2007: A descriptive epidemiological study. BMC Cancer. 2017;17:1–11.

939 91. Office for National Statistics. ONS SRS Metadata Catalogue. 2019.

940 92. Marmot MG, Adelstein AM, Bulusu L. Lessons from the study of Immigrant Mortality.
941 Lancet. 1984;323:1455–7.

942 93. Turcotte LM, Whitton JA, Friedman DL, Hammond S, Armstrong GT, Leisenring W, et al.
943 Risk of Subsequent Neoplasms During the Fifth and Sixth Decades of Life in the Childhood
944 Cancer Survivor Study Cohort. J Clin Oncol. 2015;33:3568–75.

945 94. Henson KE, Elliss-brookes L, Coupland VH, Payne E, Vernon S, Rous B, et al. Data
946 Resource Profile: National Cancer Registration Dataset in England. Int J Epidemiol.
947 2020;49:16-16h.

948 95. Rachet B, Ellis L, Maringe C, Chu T, Nur U, Quaresma M, et al. Socioeconomic
949 inequalities in cancer survival in England after the NHS cancer plan. Br J Cancer.
950 2010;103:446–53.

951 96. Quaglia A, Lillini R, Mamo C, Ivaldi E, Vercelli M, Capocaccia R, et al. Socio-economic
952 inequalities: A review of methodological issues and the relationships with cancer survival.
953 Crit Rev Oncol Hematol. 2013;85:266–77.

954 97. Franke S, Kulu H. Cause-specific mortality by partnership status: Simultaneous analysis
955 using longitudinal data from England and Wales. J Epidemiol Community Health.
956 2018;72:838–44.

957 98. Harding S, Rosato M, Teyhan A. Trends for coronary heart disease and stroke mortality
958 among migrants in England and Wales, 1979-2003: Slow declines notable for some groups.
959 Heart. 2008;94:463–70.

960 99. Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing
961 Risk. J Am Stat Assoc. 1999;94:496–509.

962 100. Greenland S, Finkle WD. A Critical Look at Methods for Handling Missing Covariates in
963 Epidemiologic Regression Analyses. Am J Epidemiol. 1995;142:55–64.

964 101. Parkin DM. International variation. Oncogene. 2004;23:6329–40.

965 102. Azagba S, Shan L, Latham K. Assessing trends and healthy migrant paradox in
966 cigarette smoking among U.S. immigrant adults. Prev Med (Baltim). 2019;129.

967 103. Asaria M, Ali S, Doran T, Ferguson B, Fleetcroft R, Goddard M, et al. How a universal
968 health system reduces inequalities: Lessons from England. J Epidemiol Community Health.
969 2016;70:637–43.

970 104. Crawford J, Ahmad F, Beaton D, Bierman AS. Cancer screening behaviours among
971 South Asian immigrants in the UK, US and Canada: A scoping study. Heal Soc Care
972 Community. 2016;24:123–53.

973 105. White MC, Holman DM, Boehm JE, Peipins LA, Grossman M, Jane Henley S. Age and
974 Cancer Risk. Am J Prev Med. 2014;46:S7–15.

975 106. Arnold M, Rutherford MJ, Bardot A, Ferlay J, Andersson TML, Myklebust TÅ, et al.
976 Progress in cancer survival, mortality, and incidence in seven high-income countries 1995–
977 2014 (ICBP SURVMARK-2): a population-based study. Lancet Oncol. 2019;20:1493–505.

978 107. Li Y, Heath A. Minority ethnic men in British labour market (1972-2005). Int J Sociol Soc
979 Policy. 2008;28:231–44.

980