ETHNIC INEQUALITIES IN MORTALITY IN ENGLAND AND WALES: EXAMINING LIFE EXPECTANCY DATA AND METHODS

Harry Taylor
King's College London

Laia Bécares
King's College London

Dharmi Kapadia
University of Manchester

James Nazroo
University of Manchester

Sarah Stopforth
University of York

Chris White
Office for National Statistics

August 2024

The Nuffield Foundation is an independent charitable trust with a mission to advance social well-being. It funds research that informs social policy, primarily in Education, Welfare and Justice. The Nuffield Foundation is the founder and co-funder of the Nuffield Council on Bioethics, the Ada Lovelace Institute and the Nuffield Family Justice Observatory. The Foundation has funded this project, but the views expressed are those of the authors and not necessarily the Foundation. Visit www.nuffieldfoundation.org

ACKNOWLEDGEMENTS

This Work contains statistical data from ONS which is Crown Copyright. This work was undertaken in the Office for National Statistics Secure Research Service using data from ONS and other owners. The use of this data and does not imply the endorsement of the ONS or other data owners in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistical aggregates.

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 supported by the ESRC Census of Population Programme (Award Red: ES/V003488/1). The authors alone are responsible for the interpretation of the data.

We would like to thank Catherine Dennison at the Nuffield Foundation for her support throughout the project. We would also like to thank Rosemary Drummond, Bethan Powell, David Tabor and Jason Hilton for their valuable input into this project, and assistance in understanding how the ONS estimates of life expectancy were originally produced. We are thankful to the advisory board of the wider project: Christopher Phillipson, Justine Fitzpatrick, Aideen Young, Paul McGarry, Henglien Lisa Chen, Elizabeth Webb.

STUDY TEAM

Harry Taylor is a Postdoctoral Research Associate in the Department of Global Health and Social Medicine, King's College London, and a freelance data scientist. His academic research explores ethnic inequalities using quantitative methods, with a particular focus on health. He has also conducted research into misinformation and vaccine hesitancy for the Alan Turing Institute.

Laia Bécares is Professor of Social Science and Health in the Department of Global Health and Social Medicine, King's College London. Her research interests are in studying the pathways by which the discrimination and marginalisation of people and places lead to social and health inequalities across the life course, with a specific focus on racism and heteronormativity as systems of oppression.

Dharmi Kapadia is Senior Lecturer in Sociology at The University of Manchester. Her research is concerned with investigating ethnic inequalities in health, racism in mental health services and how disadvantage accumulates over the life course to create health and social inequalities for ethnic minority older people.

James Nazroo is Professor Emeritus of Sociology at the University of Manchester. His work on ethnic inequalities in health has led the field in demonstrating the importance of both socioeconomic inequalities and experiences of racism and racial discrimination. His publications in relation to ageing have advanced our understanding of the patterns and determinants of social and health inequalities, as well as the health and wellbeing outcomes of retirement and later life employment.

Sarah Stopforth is a Lecturer in Sociology at the University of York. She is a quantitative sociologist and her research focusses on the social stratification of health and education by social class, gender, and ethnicity.

Chris White is a principal researcher at ONS specialising in health inequalities research. He specialises in the analyses of life expectancy and health state life expectancy across a range of sociodemographic characteristics.

CONTENTS

EXECUTIVE SUMMARY

Overview

In 2021, the ONS published experimental life expectancy estimates for the period 2011 to 2014 showing longer life expectancy for minoritised ethnic groups in England and Wales when compared to a broad white ethnic category mostly composed of the White British group [1]. This was in contrast to a large body of evidence of poorer health outcomes among many minoritised ethnic groups [2–4]. Despite being labelled as experimental, the estimates of life expectancy have been used in academic journal articles [5–8] and reported in the press [9] without caveat.

This report describes findings from our study, where we examine possible explanations for this disconnect between health and mortality outcomes. Noting that migrants comprise a substantial proportion of the population of many minoritised ethnic groups in the UK, we examine what has been labelled as the "migrant mortality advantage" [10]. The migrant mortality advantage, whereby migrants have more favourable mortality outcomes than the majority population, despite often having poorer health or conditions, has been documented in several countries [11–16]. A recent review of explanations for the migrant mortality advantage found most explanations to be "insufficient or problematic" [17]. Recent evidence [18] suggests that this mortality advantage is more likely due to statistical error, whereby unrecorded out-migration of people returning to their origin country renders individuals "statistically immortal" [19] (cited by [20]): as the deaths of these people are not recorded in UK statistics, they are assumed to still be alive and present in the country. This means that life expectancy estimates, calculated by counting the population and number of deaths within a given cohort, are artificially inflated. Further intensifying this effect is the observation that migrants returning to their origin country later in life may have poorer health and higher mortality rates, as they return home in the final period of their lives, a phenomenon known as "salmon bias" [13,20–24]. Evidence from France [18] has shown that if data can provide visibility on migrants' outcomes after they return to their country of origin, the migrant mortality advantage no longer exists.

The data that would be required for us to definitively assess the presence of salmon bias in the UK is not currently available to researchers; therefore, we instead tested the sensitivity of the ONS' life expectancy estimates to selected potential sources of error. Firstly, working with some of the ONS researchers who produced the original life expectancy estimates, we attempted to replicate those estimates using the data available to us, which was a smaller sample than the data used by the ONS. We then embarked upon a process of calibration, to best match our estimates to the originals. Once we established a base model, we explored the potential sources of error in the estimates, and tested how sensitive the life expectancy estimates were to variation in those parameters. We then ran a simulation to examine how the life expectancy estimates would respond to different assumptions around outmigration, and the mortality rate of those who out-migrated Lastly, we produced estimates of disability-free life expectancy, to compare how the observed differences between ethnic groups contrasted with differences between ethnic groups in the life expectancy estimates.

Key findings

- We observe that levels of missing data are particularly high among minoritised ethnic groups at older ages, i.e. in the age groups where missing data can most effect life expectancy at birth estimates. To illustrate, missingness in the 60+ age group was 6% for the White British ethnic group, but was 28% for the Bangladeshi ethnic group.
- We show that in the dataset used to produce the ONS experimental statistics, the mortality rate generally declines over time across the study period, with this decline being sharper among minoritised ethnic groups compared with the White British group. An explanation for this is a gradual increase in statistical immortality: over time, people recorded in the 2011 Census may leave the country, but their deaths are not recorded. However, they are still counted in the population counts, causing the mortality rate to decline.
- We demonstrate that life expectancy estimates are more sensitive to missing deaths in the oldest age groups for many minoritised ethnic groups, with estimates for Arab and Bangladeshi women aged 90+ changing by more than 2 years when we simulated the effects of 20 unrecorded deaths, whereas the estimates for other ethnic groups (particularly White British) remained stable.
- We show that among people from minoritised ethnic groups the age-specific mortality rates of the UK-born cohort are higher than among the foreign-born cohort, which could demonstrate higher levels of statistical immortality among the foreign-born cohort due to unrecorded outmigration.
- We calculate disability-free life expectancy (DFLE) based on general health and long-term limiting illness. DFLE contradicts the life-expectancy estimates, being lower among most minoritised ethnic groups.
- We find particularly poor life expectancy and disability-free life expectancy outcomes for the Gypsy/Traveller ethnic group (LE at birth: 78.1 years for women and 75.4 years for men, compared with 83.1 and 79.7 years respectively for the White British group; DFLE at birth: 49.3 years for women and 46.6 for men, compared with 65.0 and 64.2 years).
- We simulate the impact of using different assumptions around outmigration to those used in the ONS methodology. To deal with statistical immortality, we observe what happens if we deviate from the assumption that missing people have the same mortality risk as those who are not missing, instead varying the mortality rate of the missing cohort. We find that estimates of life expectancy at birth varied greatly for minoritised ethnic groups, observing a change of up to 5 years in life for some ethnic groups. Conversely, we saw a change of less than 0.5 years for the White British group.
- Given the high amount of missing data among many minoritised ethnic groups, particularly at older ages, and the sensitivity of the life expectancy estimates to missingness (also particularly at older ages), we conclude that the ONS estimates of life expectancy should be treated with caution, and should not be reported or used without detailed caveat.

Conclusion

There is considerable evidence showing ethnic inequalities in health in the UK [2–4]; however, mortality estimates derived using the same methodology to the life expectancy estimates presented here [25] have been used by the previous UK Government to draw this evidence into question [26]. If left unchallenged, there is a risk that the supposed contradiction identified by these experimental estimates of life expectancy will lead to policymakers de-prioritising addressing ethnic inequalities, which will in turn have a real negative impact on the lives of people from minoritised ethnic groups. It is vital that the evidence used to inform policy on such a crucial topic is reliable and is robust to limitations in the available data. The evidence in this report demonstrates that estimates of life expectancy for minoritised ethnic groups in England and Wales are not robust to missing data, and should not be used to evidence such a key area of decision making.

Recommendations

The ONS' estimates of life expectancy for minoritised ethnic groups exhibit high sensitivity to error that is not seen in the estimates for the White British population. This is due to the large amount of missing data among these groups resulting in statistical immortality, and the potential for those missing cohorts to be at higher risk of morbidity and mortality. We would advise researchers and those engaged in policy work to treat these life expectancy estimates with extreme caution, and reference these estimates only with detailed caveats, if at all. In addition, due to the limited numbers of the oldest people (aged 80 or above) from minoritised ethnic groups living in the UK, and the importance of these cohorts on calculating accurate life expectancy statistics, we advise researchers and those engaged in policy work wishing to comment on the health of people from minoritised ethnic groups in the UK to refer to the existing body of evidence on ethnic inequalities in health, or disability-free life expectancy estimates, rather than refer to experimental life expectancy estimates.

The ONS are currently working on a new admin-based approach to producing statistics, which may lead to more accurate estimates of outmigration through observing pauses in individuals' activity in interacting with a range of government services. This would allow for more accurate estimates of life expectancy. However, to produce the most accurate estimates of life expectancy among minoritised ethnic groups would require statistical visibility of all migrants who have ever lived in the UK, regardless of when they returned to their country of origin. This suggests that when considering ethnic inequalities the admin-based approach may suffer from the same limitations as the ONS experimental life expectancy estimates. To help us begin to address underlying data problems, we recommend making available linked UK pensions data to UK researchers, as has been done in other countries.

INTRODUCTION

Understanding how health and mortality risk varies across ethnic groups is vital to informing policy that aims to achieve health equity. In the Global North, people from minoritised ethnic groups generally have poorer health than white majority groups [2–4]. Among certain minoritised ethnic groups in some countries, for example Black Americans, this health disadvantage is reflected in higher levels of mortality [27]. Much research has focussed on the life expectancy/ mortality risk of migrants rather than minoritised ethnic groups. In the United Kingdom (UK), migrants represent more than half of people from minoritised ethnic groups, although there is considerable variation in the proportion of each ethnic group who are migrants [28]. Overall, migrants are reported as having longer life expectancies than the white majority in their host country [29], with research showing lower levels of mortality among migrants in France [11], England and Wales [12], Italy [13], Scotland [14], Sweden [15], and Switzerland [16].

The so-called "migrant mortality advantage" remains largely unexplained, despite much research attention [30]. Indeed, a recent meta-analysis of theoretical explanations for the phenomenon found most of those explanations to be "problematic or insufficient" [17]. While the migrant mortality advantage is considered by some to be a paradox, others believe that the observation of higher life expectancy among migrants is actually due to methodological error [21].

A common explanation of the migrant mortality advantage is the "healthy migrant effect" [12], whereby migrants to a country have better health at any given age than those in both their origin and destination countries. However, this evidence is mixed, as some studies have found that migrants actually have poorer health [31]. Another explanation for the apparent migrant mortality advantage is "salmon bias", whereby migrants who become ill return to their country of origin in the final period of their life. One of the main methodological challenges in examining salmon bias is the incompleteness of official migration statistics. Generally, the outward migration of individuals is not officially recorded, meaning that they continue to contribute towards official population counts, even after death, effectively being rendered "statistically immortal" [19] (cited by [20]). The unknown rate of unobserved emigration hinders effective measurement of mortality using linked sources. One solution to this problem would be to ethnically code all deaths between 2020 and 2022 and use the census populations as a proxy for the denominator; however, although the UK government has called for the recording of ethnicity on death certificates [32], this has not yet been implemented.

While the impact of salmon bias upon the migrant mortality advantage has been tested in England and Wales [22,33], Italy [13], and Sweden [23], it has been found that the salmon bias hypothesis is not able to fully explain lower mortality among migrants, although these studies have not been able to overcome the issue of statistical immortality. However, recent research from France has identified higher mortality among returnees. The study used pensions data to track individuals regardless of whether they migrated, demonstrating that migrants have higher mortality (Mortality Hazard Ratio ranging from 1.6 from "Other European" countries to 3.6 in Asia) than those remaining in France, and when this mortality is fully taken into account (by including migrants who returned prior to the observation window), the migrant mortality advantage

disappears [18]. Research from the United States (US) [24] also found higher mortality among returnees, with recently-migrating Hispanic women and men having a higher mortality rate (76% and 45% respectively) than those remaining in the US, although recently-migrating non-Hispanic white women and men also saw a mortality increase of 19% and 37% respectively. However, this excess mortality among returnees could not fully explain the migrant mortality advantage; taking into account mortality outcomes among people who had left the US affected the age-specific mortality rates of both Hispanics and non-Hispanic Whites by less than 1%.

In the UK, migrants represent more than half of people from minoritised ethnic groups [28]. As such, estimates of life expectancy by ethnic group are heavily influenced by any bias or data shortcomings that affect the calculation of migrants' life expectancy. In July 2021, the ONS released a set of experimental statistics titled "Ethnic differences in life expectancy and mortality from selected causes in England and Wales: 2011 to 2014" [1]. These estimates showed higher life expectancy for nearly all minoritised ethnic groups compared with the White British group, which contradicts much of the evidence on ethnic inequalities in health [2]. This contradiction is further discussed in an article for the NHS Race and Health Observatory by Nazroo and colleagues [34]. Certain life expectancy estimates in this release were particularly surprising. For example, Bangladeshi women had an estimated life expectancy at birth of 87.3 years, and for Bangladeshi men this was 81.1 years, which compares favourably with the life expectancy of women and men living in Japan (87.7 and 81.8 respectively), a country with the third highest life expectancy in the world as of 2021 [35]. Furthermore, the life expectancy for Bangladeshi women aged 90+ was 11 years (compared to 5.3 years for the White British group), suggesting this cohort could expect to live to over 100 years old. These results are particularly surprising, given the extensive literature on the poor health outcomes for people of Bangladeshi background in the UK over many decades [36–38].

In the technical report accompanying the ONS life expectancy estimates, they describe several data limitations which would likely affect the accuracy of life expectancy estimates, including differences across ethnic groups in the rate of linkage error between the Census data and death registrations, as well as limited data on emigration [1]. Indeed, these statistics were clearly labelled as experimental in the data release, with the ONS defining experimental statistics to be those which are "in the testing phase and not yet fully developed". In the ONS' guide to experimental statistics, they state that "users should be aware that experimental statistics will potentially have a wider degree of uncertainty" [39]. Despite this caveat, the estimates of life expectancy have been referenced in academic literature [5–8] and the press [9] without mention of this uncertainty; other sources state that the estimates are experimental without any critical discussion of their methodology [40–42]. The estimates were also referenced in the Institute for Fiscal Studies' Deaton Review of Inequalities [43], although a footnote detailing concerns around the accuracy of these estimates was included. Furthermore, the ONS have continued to release mortality rate statistics by ethnic group using the same methodology [25,44]; these new releases continue to document greater life expectancies for minoritised ethnic groups which are potentially incorrect. Hence, this project presents a timely contribution to the debate around the accuracy of life expectancy estimates for minoritised ethnic groups.

AIMS AND OBJECTIVES

The ONS life expectancy estimates have several potential limitations, described in the sections below, and in this project we collaborated with the ONS to improve current estimates of ethnic differences in mortality rates, life expectancy and healthy life expectancy. The original objectives of the project were as follows:

1) To provide accurate and up to date estimates of mortality rates and life expectancy (LE) for ethnic groups in England and Wales, adjusting for age group and gender.
2) To extend the ONS analyses to provide novel healthy life expectancy (HLE) and disability-free life expectancy (DFLE) estimates.
3) To provide these estimates for younger and older groups and for men and women.
4) To disseminate the findings and underlying methods to a wide range of policy audiences.

METHODOLOGY

In the following sections, we discuss the methodology used by the ONS team in producing their estimates of life expectancy, and describe the steps we took to achieve our aims and objectives. We begin by describing how the ONS estimates were derived, and the data we had available in order to replicate these estimates. We describe our process of calibrating our data model to match the estimates produced by the ONS, and show to the extent to which we were able to replicate the original estimates. This stage was the essential first step, before we were able to make adjustments to the estimates taking into account potential sources of error we identified, and to observe how sensitive or robust the estimates were to these adjustments.

DATA AND METHODS FOR THE ONS LIFE EXPECTANCY ESTIMATES

The principal dataset used by the ONS in their estimates of life expectancy was the Public Health Data Asset (PHDA). This comprises data from all people living in England and Wales who participated in the 2011 Census, which is linked to death registrations from the years 2011-2021. The structure of this dataset can be seen in the data map in figure 1. First, the 2011 Census was mapped to the NHS Patient Registers from 2011-2013 to determine individuals' NHS numbers. This process comprised two stages: the first was a deterministic match on 24 key fields including forename, surname, date of birth, sex and geography, which obtained a 93.5% match success [45]. For those individuals who could not be matched, probabilistic matching was conducted using 13 different combinations of personal identifiers, which matched a further 1.1% of Census records [45].

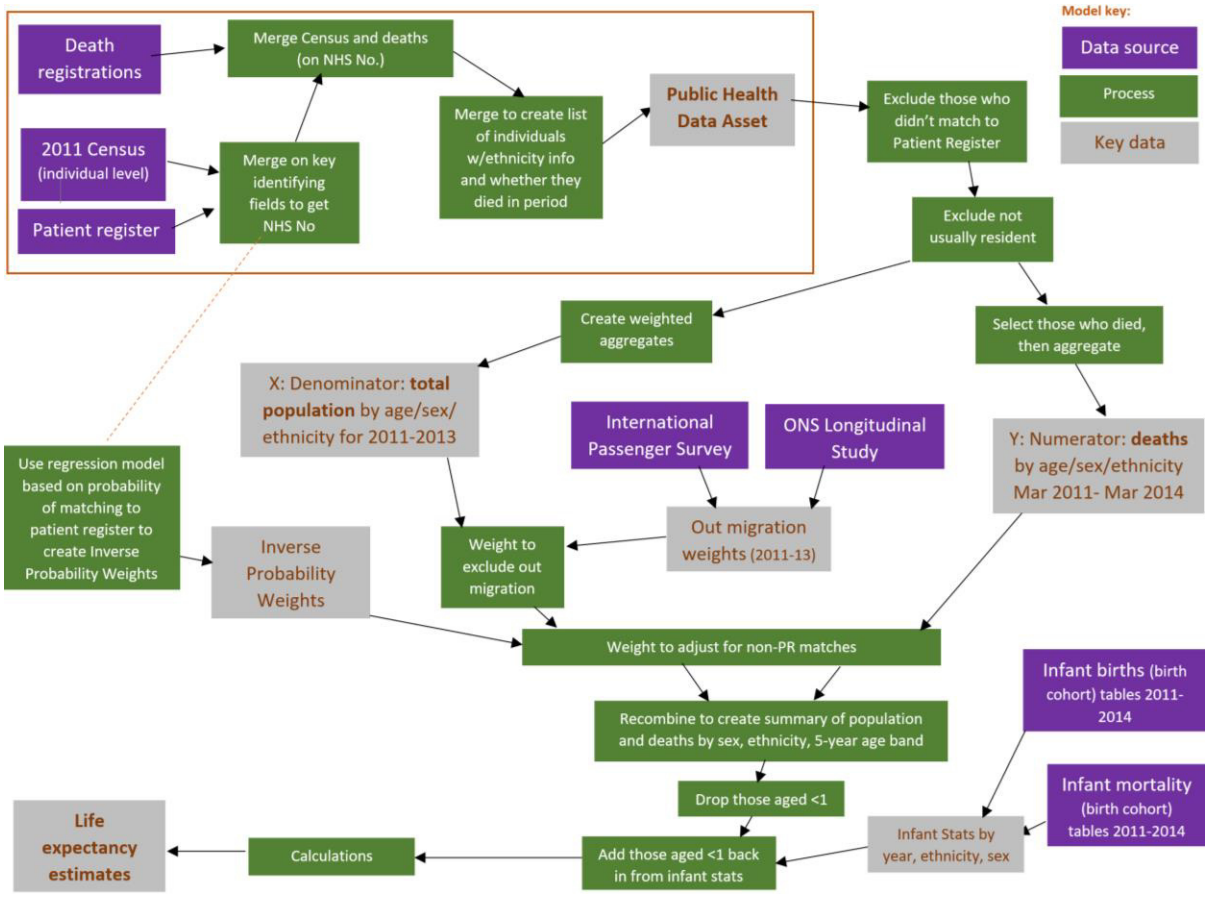


Figure 1: Map of methodology for producing ONS experimental life expectancy estimates by ethnic group

The NHS numbers were then used to merge 2011 Census data with death registrations. In their analysis, the ONS considered deaths between Census day on the 27th March 2011 and the 26th March 2014 only; anyone who died after this date was recorded as still being alive. Some deaths could not be mapped to individuals who participated in the 2011 Census; this is most likely because those individuals moved to England or Wales after the 2011 Census or because they did not participate in the Census; however, it could also be due to linkage failure between the Census and Patient Register or Patient Register and death registrations. An estimated 6.1% of individuals who did not take part in the Census were not captured in this study [45]; those who participated in the 2011 Census but who were not usually resident in England or Wales were excluded from the study population. It should be noted that although ethnicity data was not available for these unlinked deaths, country of birth information showed that linkage failure was substantially higher among people who were born outside the UK, with 37.4% of deaths of people born in Central and Western Africa being unlinked, compared with 10.8% of people born in the UK [1].

Age, sex, ethnic group, area deprivation, and region were derived from the 2011 Census returns. Ethnic group data was self-reported from a list of 18 pre-defined ethnic group categories. In their analysis, the ONS aggregated some ethnic groups resulting in 10 ethnic groups (Asian other, Bangladeshi, Black African, Black Caribbean, Black other, Indian, Mixed, Other, Pakistani,

White)[1]. The working dataset was then aggregated into total population and total deaths by ethnic group, age, and sex (X and Y respectively in figure 1).

During the study period 2011-2014, individuals may have left the country, meaning that their deaths would not have been recorded in England and Wales death statistics. However, as there is no register that records emigration (or "outmigration"), this means they would remain as part of the population figures. To account for this, the ONS adjusted the population totals using estimated outmigration totals. These estimates of outmigration were produced for age group, sex and ethnic group, and were calculated using data from the International Passenger Survey (IPS), and the Office for National Statistics Longitudinal Survey (ONS-LS) [46]. The IPS is a survey of movements in and out of the UK and can be used to estimate international migration; however, the IPS does not record ethnicity, and therefore correlations between citizenship and ethnicity were used. This is a potential source of error, as the majority of ethnic minority people in the UK are UK citizens. The ONS-LS was used to supplement the IPS data; however, the ONS-LS emigration data relies on outmigrants de-registering with the NHS, which they are not required to do.

The population and deaths totals were weighted to account for the incomplete match of the Census 2011 data to the NHS Patient Register data. If an individual's Patient Register data was unavailable, the linkage to death registrations was not possible, making it harder to track deaths. To address this, the ONS excluded those whose Patient Register data could not be mapped, then used Inverse Probability Weighting to scale up population and deaths totals based on the characteristics of those individuals that were successfully linked. The missingness from the Patient Register did not conform to the characteristics of the general population, with people from minoritised ethnic groups more likely to be missing than the White group. To illustrate, Patient Register data could not be mapped for approximately 17.3% of people from the Chinese ethnic group, compared with 3.7% of the White group [45]. In addition, when the weights were applied (to both the population totals, and the deaths totals), populations were scaled up more than deaths generally, but to a greater extent for minoritised ethnic groups, resulting in a denominator bias which worked to over-estimate life expectancy; this weighting differential had less of an effect on the White group, who saw the smallest change in life expectancy post-weighting. It should also be noted that this methodology assumes that the mortality rate for those who were missing from the Patient Register was the same as those with the same characteristics whose Census data could be mapped to the Patient Register. However, those missing from the Patient Register may have migrated and, under the salmon bias theory, they may consequently have a higher mortality risk if they were returning to their country of origin. The implications of these factors combined mean that the inverse probability weights are likely to over-estimate life expectancy for people from minoritised ethnic groups.

[1] The aggregations of ethnic groups were as follows - **Asian Other**: Chinese, Any other Asian background, **Mixed:** White and Black Caribbean, White and Black African, White and Asian, Any other Mixed or multiple ethnic background. **White:** English, Welsh, Scottish, Northern Irish or British (referred to in this report as White British), Irish, Gypsy or Irish Traveller, Any other White background. **Other:** Arab, Any other ethnic group.

Like in most national censuses, there is typically a proportion of the population who do not participate in the Census of England and Wales. In the 2011 Census, the response rate was approximately 94%, although this varied by ethnic group; with a 95% response rate for the White British population, ranging to a 64% response rate among the Other Black population [47]. This means that for many minoritised ethnic groups, there is a higher proportion of the resident population whose mortality outcomes are unknown, as their non-participation in the Census means that their death (if applicable) would not be recorded in the dataset used in these life expectancy estimates.

Life expectancy estimates are calculated using a life table which relies on counting the size of the population at a given point, then counting the number of people who died within a given time frame. In this report, we assume that factors that cause error in the count of populations or deaths will also cause error in life expectancy calculations. In our estimates, we count the population and deaths in groups of people separated by age band, sex and ethnic group. As part of this, we require information on births and deaths in the first year of life; however, anyone born after the 2011 Census was not captured in the study population. As such, birth cohort tables comprising infant births and mortality for the period 2011-2014 were included in the data. Anyone under 1 year of age was assumed to have been born in England or Wales; an error which will be more pronounced among minoritised ethnic groups whose populations typically have higher proportions of migrants.

SUMMARY OF LIMITATIONS IN ONS METHODOLOGY

Having described the ONS methodology of producing life expectancy estimates by ethnicity, we can summarise the main methodological limitations as follows:

1) The failure of death registration linkage is substantially higher among people born outside the UK.
2) The failure of Patient Register linkage is higher among minoritised ethnic groups, and substantially higher for some minoritised ethnic groups.
3) Participation in 2011 Census tends to be lower for people from minoritised ethnic groups, representing a larger population who are not covered by this survey, and whose outcomes are unknown, causing greater uncertainty in the level of mortality risk experienced in these groups.
4) Inverse probability weights (designed to account for missingness) do not account for potentially higher mortality rate among returning migrants.
5) There is no reliable outmigration data by ethnic group, meaning a high risk of “statistical immortality” among the population base of minoritised ethnic groups.
6) There is no data on infant (<1 year) migration into England and Wales.
7) Ethnic groups were aggregated, meaning estimates were produced for only 10 out of 18 ethnic groups recorded in the 2011 Census.

STAGE 1: CALIBRATION

The aim of our research was to provide improved life expectancy estimates by ethnic group, and extend the existing estimates with additional analyses. To do this, we first needed to replicate the ONS analysis. The ONS used an individual-level dataset called the Public Health Data Asset comprising data for the whole population enumerated on Census day 2011 and linked to the Patient Register. This dataset is only accessible by ONS researchers, and therefore we were not able to use the full dataset in our analysis. Instead, we successfully applied to access a smaller sample of the full dataset, the Public Health Research Database (PHRD) [48]. Although the PHRD contains individual-level death registration data, it contains only a 5% sample of those still alive at 2020; this represents the main difference to the full PHDA data.

The ONS calculated life tables for men and women from 10 ethnic groups in 20 age groups (in 5 year intervals ranging from <1 to 90+ years old), resulting in a total of 400 life expectancy estimates. Our aim was to be able to replicate each of the 400 estimates as closely as possible. The ONS shared their code for the model they used to create their life expectancy estimates. We adapted this for use with the PHRD, and underwent a process of calibration of the model's parameters. The adjustments made were centred around the definition of the sample population and deaths, and the creation of the Inverse Probability Weights (IPWs).

1. **Sample:** To account for the 5% sample of the alive population used in the PHRD, we created sample weights to adjust the population totals. For those people who died after 2014 (and for whom there was full individual-level data), they were treated the same as those who died. As a robustness check, we took a random 5% sample of those who died (including those who died after 2014) so that weights were not required for the calculation of life expectancy. The obtained results were similar, except for among the cohorts with the smallest sample sizes; as such, we decided the most principled approach was to retain the full sample of deaths and use sample weights.
2. **Inverse Probability Weights (IPWs):** The IPWs created by ONS were based on a regression model that modelled how the chance of not matching to the patient register varied according to age, sex, ethnic group, and region. We took two approaches to our use of IPWs. In one approach, we used the same model coefficients as calculated by the ONS and applied these to our population. In the other approach, we ran the regression model on our own data, and created our own weights from scratch. Following the creation of the IPWs, the ONS trimmed the weights to remove any extreme values. This was done by limiting any IPWs above the 99th percentile and then applying a rescaling factor to make sure they summed to the right population. Although we do not have access to the exact rescaling factor, we attempted to emulate this approach, and also tried limiting weights to 2 standard deviations above and below the mean.
3. **Other factors:** there were also several steps in the methodology that were not entirely clear. We experimented with calibrating these parameters and the ONS were later able to confirm their approach. These factors included:
    a. Whether to include the not usually resident population in the Census. The ONS confirmed that this cohort should be excluded from the analysis.
    b. Whether outmigration weights were applied cumulatively to the population base, or individually for 2011, 2012 and 2013. Also, whether the outmigration weight

should be applied to the population base for 2011. Both of these approaches were implemented by the ONS in their methodology; however, we had interpreted the outmigration weights as corrections that are applied at the end of a year to represent the outmigration that occurred during that year. As such, the 2012 population base would have the 2011 outmigration weight applied, and the 2013 population base has the product of the 2011 and 2012 outmigration weights applied). However, using the ONS' original approach produced the best-calibrating model.

c. Whether the IPWs should be applied only to deaths (with the deaths of non-patient register matching individuals excluded), or whether individuals who could not be matched to the patient register should be excluded from both the population and mortality counts. The ONS instructed that IPWs were applied to both deaths and populations, and anyone not matching to the patient register was excluded entirely.
d. What time period to use for deaths, and whether to use the age at death or the age that the individual would have been at the start of the study window. The best-calibrating option was to use period March 27, 2011 to March 26, 2014, and used the age that the individual would have been at the start of the study window.
e. Finally, as the infant birth and mortality statistics were only available by ethnic group, and not ethnic group and sex, we had to estimate this split. However, later the ONS were able to make the data publicly-available, meaning we could use the data in our model.

We ran several models varying the calibration parameters, and selected the model with the lowest average error across the 400 estimates produced. The optimal model used our own regression model for the IPWs, with weights limited to 2 standard deviations of the mean, and used the full sample of deaths with 5% sample of the alive population. This model calibrated to the ONS model with an average estimate error of +/- 0.5 years. Table 1 shows the calibration results; the table has been colour-coded for ease of interpretation: yellow indicates no change, increases and decreases in life expectancy are coloured red and green respectively, with the hue intensifying for larger increases. The estimate error was more pronounced for minoritised ethnic groups than it was for the White group, whose estimates calibrated to within +/- 0.1 years. In particular, the life expectancy at birth estimates for Black African women were almost 3 years lower in our model than in the estimates produced by the ONS. This error increased to 5 years at age 90+.

In general, the life expectancy estimates we calculated were lower for minoritised ethnic groups than those produced by the ONS. However, men and women from the Asian Other group, and men from the Black Other group had life expectancy at birth estimates approximately 0.5 years higher than the ONS estimates. The life expectancy estimates can be seen in table 2, demonstrating that life expectancy at birth is still higher for nearly all minoritised ethnic groups than the White group. This is also true for life expectancy at age 90+.

Table 1: Comparison between best-calibrating model and ONS life expectancy estimates (difference in years)

	Women										Men									
Age group	Asian other	Bangladeshi	Black African	Black Caribbean	Black other	Indian	Mixed	Other	Pakistani	White	Asian other	Bangladeshi	Black African	Black Caribbean	Black other	Indian	Mixed	Other	Pakistani	White
<1	0.4	-0.4	-3.0	-0.1	-1.9	0.0	-0.1	0.2	-0.2	0.0	0.4	-0.1	-0.2	0.1	0.4	0.0	0.4	-0.1	0.0	0.0
01-04	0.4	-0.4	-3.1	-0.1	-2.0	0.0	0.0	0.2	-0.2	0.0	0.4	-0.1	-0.2	0.0	0.4	0.0	0.4	-0.1	0.0	0.0
05-09	0.5	-0.4	-3.0	-0.1	-2.0	0.0	-0.1	0.1	-0.2	0.0	0.4	-0.1	-0.3	0.1	0.4	0.1	0.3	-0.1	0.0	0.0
10-14	0.4	-0.5	-3.0	-0.1	-2.0	0.0	-0.1	0.2	-0.2	0.0	0.4	-0.1	-0.2	0.1	0.5	0.1	0.4	-0.1	0.0	0.0
15-19	0.4	-0.4	-3.0	-0.1	-1.9	0.0	0.0	0.2	-0.2	0.0	0.5	-0.1	-0.2	0.0	0.4	0.0	0.4	-0.1	0.0	0.0
20-24	0.5	-0.4	-3.0	0.0	-2.0	0.0	-0.1	0.1	-0.2	0.0	0.4	-0.1	-0.2	0.0	0.4	0.1	0.4	-0.1	0.0	0.0
25-29	0.4	-0.4	-3.0	0.0	-1.9	0.0	-0.1	0.2	-0.2	0.0	0.4	-0.1	-0.3	0.1	0.4	0.1	0.4	-0.1	0.0	0.1
30-34	0.4	-0.4	-3.0	-0.1	-2.0	0.0	-0.1	0.2	-0.2	0.0	0.4	-0.1	-0.3	0.1	0.5	0.0	0.4	0.0	0.0	0.0
35-39	0.4	-0.4	-3.0	-0.1	-1.9	0.0	-0.1	0.2	-0.2	0.0	0.4	-0.1	-0.2	0.2	0.5	0.0	0.4	0.0	-0.1	0.1
40-44	0.4	-0.4	-3.1	0.0	-2.0	0.0	0.0	0.2	-0.2	0.0	0.4	-0.1	-0.2	0.1	0.5	0.1	0.4	0.0	0.0	0.0
45-49	0.4	-0.4	-3.1	-0.1	-2.0	0.0	0.0	0.2	-0.2	0.1	0.4	-0.1	-0.3	0.1	0.4	0.1	0.4	-0.1	0.0	0.1
50-54	0.4	-0.4	-3.1	0.0	-2.0	0.0	0.0	0.2	-0.2	0.0	0.4	0.0	-0.2	0.1	0.5	0.0	0.4	-0.1	0.0	0.0
55-59	0.4	-0.4	-3.1	0.0	-2.0	0.1	0.0	0.2	-0.2	0.0	0.4	0.0	-0.2	0.2	0.5	0.0	0.3	-0.1	0.0	0.0
60-64	0.4	-0.5	-3.2	0.0	-1.9	0.0	-0.1	0.3	-0.2	0.0	0.5	-0.1	-0.3	0.2	0.6	0.0	0.4	0.0	0.0	0.0
65-69	0.4	-0.5	-3.2	-0.1	-2.0	0.0	-0.1	0.2	-0.3	0.0	0.4	-0.2	-0.2	0.1	0.7	0.0	0.4	-0.1	0.0	0.0
70-74	0.4	-0.6	-3.3	0.0	-2.1	0.0	-0.1	0.3	-0.3	0.0	0.4	-0.3	-0.1	0.1	0.5	-0.1	0.3	-0.1	0.0	0.0
75-79	0.4	-0.6	-3.4	-0.1	-2.5	0.1	0.0	0.1	-0.2	0.0	0.7	-0.3	0.0	0.0	0.7	0.0	0.2	-0.3	0.0	0.0
80-84	0.4	-0.7	-3.6	-0.1	-2.8	0.0	0.1	0.2	-0.2	0.1	0.9	-0.4	0.0	0.0	0.8	-0.1	0.0	-0.3	-0.1	0.0
85-89	0.8	-1.0	-4.2	0.0	-3.0	0.1	-0.3	0.0	-0.2	0.0	1.1	-0.8	0.0	-0.2	1.0	-0.2	0.2	-0.7	0.0	0.0
90+	0.8	-1.5	-5.0	0.1	-3.5	0.3	-0.4	0.3	-0.1	0.0	1.1	-2.0	0.7	-0.3	2.1	-0.1	0.0	-1.2	0.2	0.1

Source: ONS Public Health Research Database

Table 2: Life expectancy estimates from the best-calibrating model

	Women										Men									
Age group	Asian other	Bangladeshi	Black African	Black Caribbean	Black other	Indian	Mixed	Other	Pakistani	White	Asian other	Bangladeshi	Black African	Black Caribbean	Black other	Indian	Mixed	Other	Pakistani	White
U1	87.3	86.9	85.9	84.5	84.9	85.4	83.0	87.1	84.6	83.1	84.9	81.0	83.6	80.8	82.4	82.3	79.7	83.9	82.3	79.7
01-04	86.6	86.3	85.3	84.0	84.4	84.7	82.4	86.4	84.2	82.4	84.4	80.5	83.2	80.4	82.0	81.7	79.0	83.3	81.9	79.0
05-09	82.7	82.3	81.4	80.0	80.5	80.8	78.4	82.4	80.3	78.4	80.4	76.5	79.2	76.5	78.0	77.8	75.0	79.4	78.0	75.1
10-14	77.7	77.3	76.4	75.1	75.5	75.8	73.4	77.5	75.3	73.4	75.4	71.6	74.3	71.6	73.1	72.8	70.1	74.4	73.1	70.1
15-19	72.7	72.4	71.5	70.1	70.6	70.8	68.5	72.5	70.4	68.5	70.5	66.6	69.4	66.6	68.1	67.8	65.1	69.5	68.1	65.1
20-24	67.8	67.4	66.5	65.2	65.7	65.9	63.5	67.5	65.5	63.5	65.5	61.7	64.5	61.7	63.2	62.9	60.2	64.5	63.2	60.2
25-29	62.8	62.5	61.6	60.3	60.8	60.9	58.6	62.6	60.6	58.6	60.6	56.8	59.6	56.9	58.3	58.0	55.3	59.6	58.3	55.4
30-34	57.8	57.6	56.7	55.4	55.9	56.0	53.7	57.6	55.6	53.7	55.6	51.9	54.7	52.1	53.5	53.0	50.5	54.7	53.4	50.5
35-39	52.9	52.6	51.8	50.5	51.1	51.0	48.8	52.6	50.7	48.8	50.7	47.0	49.9	47.3	48.8	48.1	45.7	49.8	48.5	45.7
40-44	48.0	47.7	46.9	45.7	46.2	46.1	44.0	47.7	45.8	43.9	45.8	42.1	45.0	42.5	43.9	43.3	40.9	44.9	43.7	40.9
45-49	43.1	42.9	42.1	40.9	41.4	41.2	39.3	42.8	41.0	39.2	41.0	37.2	40.2	37.8	39.3	38.5	36.3	40.0	38.9	36.2
50-54	38.3	38.1	37.4	36.3	36.9	36.4	34.6	38.0	36.2	34.4	36.2	32.5	35.5	33.2	34.9	33.8	31.7	35.2	34.1	31.5
55-59	33.5	33.5	32.8	31.7	32.5	31.7	30.0	33.3	31.6	29.8	31.5	27.9	30.9	28.8	30.4	29.3	27.2	30.5	29.5	27.0
60-64	28.8	28.9	28.2	27.2	28.1	27.0	25.6	28.7	27.0	25.4	27.0	23.6	26.3	24.4	26.1	24.9	23.1	26.0	25.1	22.7
65-69	24.3	24.8	23.8	22.8	23.7	22.5	21.5	24.2	22.6	21.1	22.6	19.7	22.3	20.2	22.0	20.7	19.1	21.6	21.0	18.7
70-74	20.0	21.0	19.6	18.8	19.5	18.2	17.5	19.9	18.6	17.0	18.5	15.9	18.4	16.3	18.4	16.7	15.4	17.5	17.0	14.9
75-79	15.9	17.5	15.8	14.9	15.5	14.4	13.7	15.8	15.1	13.3	14.9	13.0	15.4	12.8	15.0	13.2	12.0	13.7	13.5	11.5
80-84	12.2	14.8	12.1	11.5	11.7	11.0	10.5	12.3	11.8	10.0	12.0	10.6	12.7	9.9	12.1	10.0	9.0	10.5	10.6	8.6
85-89	9.7	12.5	9.1	8.9	9.2	8.5	7.7	9.4	9.2	7.2	9.8	8.0	10.5	7.3	10.1	7.4	7.0	7.6	8.4	6.3
90+	8.0	9.5	6.7	6.9	7.1	6.9	5.6	7.8	8.0	5.3	8.4	5.9	10.2	5.7	10.4	6.0	5.7	5.2	7.4	4.7

Source: ONS Public Health Research Database

STAGE 2: IDENTIFYING POTENTIAL SOURCES OF ERROR

After calibrating our life expectancy estimates by ethnicity to those calculated by the ONS, the next stage in our analysis was to closely examine the potential sources of error in their production to consider how the estimates could be improved. Figure 2 presents a data map to demonstrate the interconnections between datasets and processing steps implemented in the production of the life expectancy statistics based on information provided with the ONS data release [1] and from conversations with the ONS team who produced the estimates. The map also documents key potential sources of error in the model (figure 2).

For each of the potential error sources identified in the map, we considered the extent to which the error could likely be mitigated, and introduced corrections into our model to observe how sensitive the estimates were to these mitigations. Where we identified procedural potential sources of error which we were unable to adjust for in our data, we also discuss these here. In figure 2, we classify potential sources of error according to whether the amount of error or missingness is known, and whether steps have been implemented to correct it. A key for this classification is below.

We discuss the potential sources of error under the following categories: linkage failure and under-enumeration, unmatched death records, sampling frame issues, outmigration error and salmon bias.

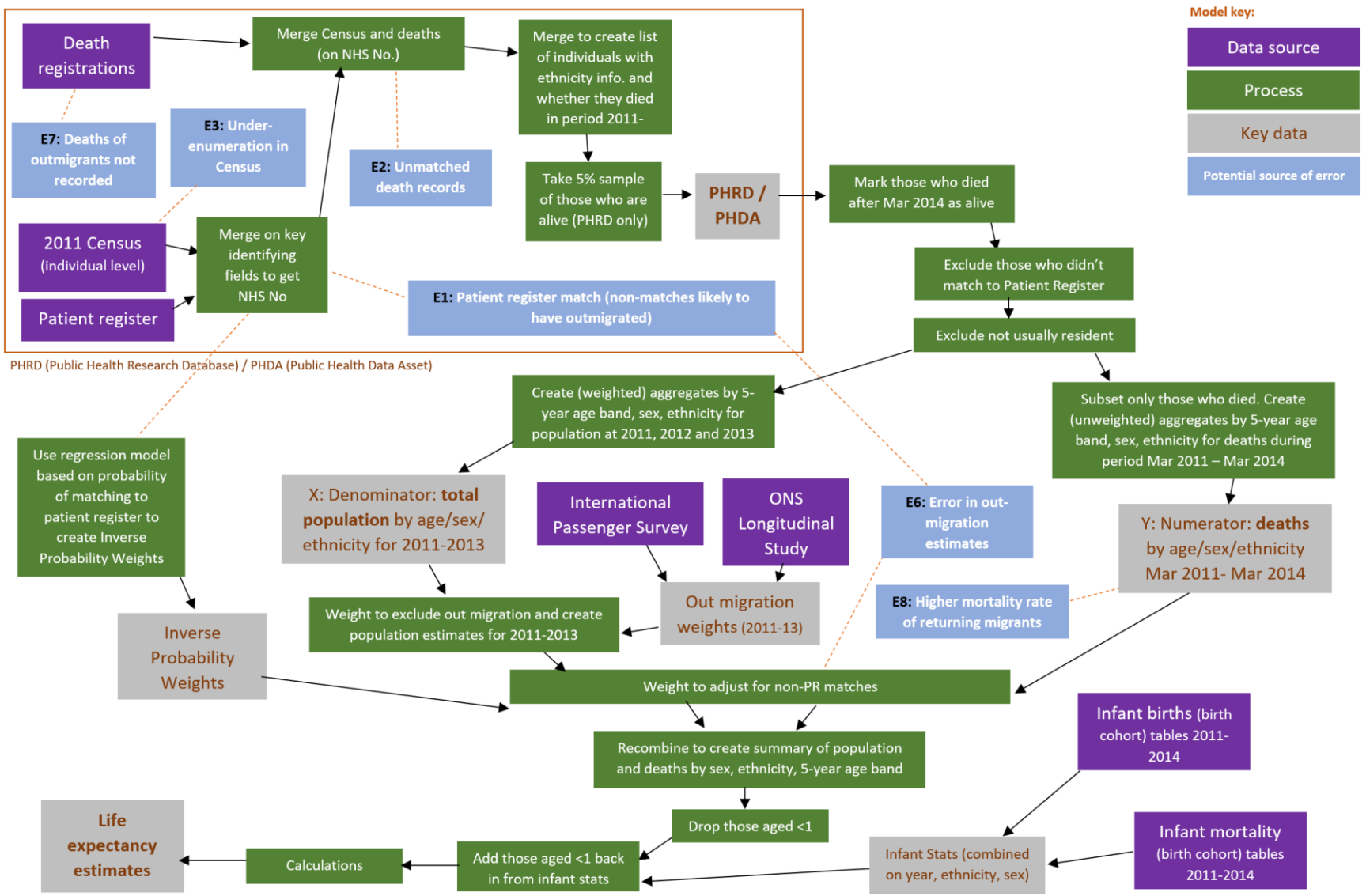


Figure 2: Map of methodology for producing ONS experimental life expectancy estimates by ethnic group with potential sources of error

ERROR 1 - PATIENT REGISTER MATCH

The main purpose of matching the base 2011 Census data to the Patient Register (PR) was to obtain their NHS number, which was subsequently used to match to the death registration data. Not all people who participated in the 2011 Census could be matched to a record in the PR; while some deaths among this cohort were matched using other fields, the reliability of this match was poor. Therefore, the group of people whose Census data could not be matched to the PR were excluded, representing approximately 2.5 million (4.8%) people when scaling weights are applied.

The characteristics of these people differed from those of the general population; with higher levels of missingness among minoritised ethnic groups (particularly at older ages), people born outside the UK, men, and people living in London. The effect of this additional missingness for minoritised ethnic populations is difficult to quantify. If it could be assumed that the mortality rates of the non-matching population were the same as those who matched to the PR, the only impact of higher missingness among minoritised ethnic cohorts would be wider confidence intervals, due to the reduction in sample size. However, if we were to assume that many of those who did not match to the Patient Register could not be linked as they had out-migrated, we would reasonably expect their mortality rate to be higher than the remaining population (who may in fact have better health, due to positive health selection among migrants) as seen in the research by Guillot et al. [18].

Inverse Probability Weights (IPWs) were used to correct for linkage failure between the Census 2011 data and the Patient Register. The IPWs are effective in ensuring that the population age structures (in terms of what proportion of the population is a given age) in each sex and ethnic group are maintained. This is important, as substantial changes in the age-population structure would risk biasing the LE estimates. The ONS noted that the IPWs increased the population of people from minoritised groups more than the IPWs increased the number of deaths from people from minoritised groups, which means that the application of IPWs would over-estimate life expectancy. However, we found the impact of the IPWs on life expectancy estimates to be minimal. In our calibration process, we tried several different approaches to specifying the IPWs; however, the IPWs had a small impact on the estimates. Removing IPWs altogether resulted in very small changes to life expectancy, with none of the 400 life expectancy estimates changing more than 0.1 years.

The reason for the IPWs only having a very slight impact on life expectancy is that the regression model used to define the IPWs has only 5 parameters: age (including single year and age band terms), sex, ethnic group, area deprivation, and region. As the life expectancy estimates are applied to the numerator and denominator (i.e. deaths and population respectively) for each age-sex-ethnicity cell, the only factors that can vary are area deprivation and region. Therefore, for a life expectancy estimate to increase in a cell, the deprivation and regional characteristics of those who died would have to be substantially different to the characteristics of those who were still alive. Furthermore, region and area deprivation had coefficients with a relatively small range (odds ratio between 0.55 and 1 for deprivation; odds ratio between 0.63 and1.01 for region).

While the IPWs prevent changes in age structure biasing the life expectancy estimates, they are unable to account for potential differences in the mortality rate among those who did not match to the PR, and generally reinforce the mortality rate of the observable population. The IPWs could have been configured to identify characteristics of those who didn't map to the patient register that might cause higher mortality rate due to salmon bias (e.g. a combination of being born outside the UK, poor health, and older age). However, this argument is cyclical: we are not able to demonstrate a higher mortality rate among this at-risk cohort precisely because they are missing from the dataset and become statistically immortal.

ERROR 2 – DEATH RECORDS UNMATCHED

A proportion (12.4%) of the death registration data could not be matched to anyone who was enumerated in the 2011 Census [1]. These deaths were mainly people who migrated in to England or Wales after the 2011 Census [45], but also included people who were present in England and Wales at the 2011 Census who did not participate in the census. There were also a small number of people whose NHS number was missing from death registrations, and so could not be matched on that basis. The characteristics of non-matched people were different to the general population. Although ethnicity information was unavailable, country of birth information showed that there were much higher levels of people born in Africa, Pakistan and Bangladesh, than there were people born in the UK.

It would be fair to exclude the deaths of those people who migrated to England or Wales after 2011, as they do not fall within our sampling frame, i.e. all people present in England and Wales during the 2011 Census. However, missingness among the cohort who did not participate in the 2011 Census may be problematic if their mortality rate differs to those who did participate in the census. It may be possible to simulate this by observing the characteristics of people who did not participate in the census, and mapping them to unmatched deaths; however, it would be necessary to know whether an unmatched death was due to non-participation in the census, or arrival after the census date, and therefore we have been unable to formally test this.

ERROR 3 - UNDER-ENUMERATION IN CENSUS WAS HIGHER FOR MINORITISED ETHNIC GROUPS

Although it was a requirement for everyone to participate in the 2011 Census of England and Wales, not all people living in England or Wales were enumerated in the 2011 Census. The published Census counts are upscaled using weights based on estimates of Census non-participation to reflect their estimate of the true population count. The characteristics of those who were not enumerated differ from those of the general population, with higher numbers of people from minoritised ethnic groups [47]. As discussed in Error 2, under-enumeration may be problematic if the mortality rate of non-participants in the census differs from those who did participate.

SAMPLING ISSUES

ERROR 4 - NO VISIBILITY OF PEOPLE WHO HAVE MIGRATED BEFORE SAMPLING WINDOW

(Not on map)

In our model, we consider only people who participated in the Census in March 2011. Anyone who migrated before this window would not be included in the data. Indeed, this missingness may appear to be inconsequential, as those who were not enumerated in the Census are also not included in the mortality figures. However, Guillot and colleagues (2023) show the importance of taking into account migrants who returned earlier in the life course, as they also have higher mortality, and must be taken into account to get a full picture of the effects of salmon bias. Unfortunately, this would require visibility of the mortality outcomes of migrants who have left England and Wales; such data are not presently available.

ERROR 5 – AGGREGATED WHITE GROUP MASKS VARIATION BETWEEN SUB-GROUPS

(Not on map)

Given the heterogeneity among ethnic groups grouped together under a single category, we disaggregated the ethnic groups as much as possible. In the ONS life expectancy estimates, the White ethnic group classification combined White British, White Other, White Gypsy/Traveller and White Irish groups. These groups have different health outcomes, with the Gypsy/Traveller group having particularly poor health [49]. We disaggregated the White group, as well as the Other/Arab group and the Asian Other/Chinese group. We were unable to split out the Mixed group due to the low sample size at certain ages for this group making the resulting estimates unreliable. We then re-calculated the IPWs using our own regression model. The outmigration estimates and infant statistics were not available at the disaggregated ethnic group level, so we used aggregated ethnic group data. The results for disaggregated ethnic groups can be seen in table 3, demonstrating the substantially lower life expectancy at birth estimates for the Gypsy/Traveller and Irish groups. It should be noted that the life expectancy at the eldest ages was high for Gypsy/Traveller men. This was likely due to the very small number of Gypsy/Traveller men surviving to age 90+ resulting in insufficient data to produce reliable estimates. We also observe that life expectancy estimates for the Arab group are notably higher than the Other group, reinforcing the need to disaggregate these two groups in our analysis.

Table 3: Disaggregated life expectancy estimates

Women															
Age	Asian other	Chinese	Bangladeshi	Black African	Black Caribbea	Black other	Indian	Mixed	Arab	Other	Pakistani	White British	White other	White Gypsy / Traveller	White Irish
U1	87.4	87.2	86.9	85.9	84.5	84.9	85.4	83.0	88.3	86.7	84.6	83.1	85.7	78.1	82.3
01-04	86.7	86.5	86.3	85.3	84.0	84.4	84.7	82.4	87.6	86.0	84.2	82.3	84.9	77.3	81.6
05-09	82.7	82.6	82.3	81.4	80.0	80.5	80.8	78.4	83.7	82.0	80.3	78.4	81.0	73.3	77.6
10-14	77.8	77.6	77.4	76.4	75.1	75.5	75.8	73.4	78.7	77.1	75.3	73.4	76.0	68.3	72.7
15-19	72.8	72.6	72.4	71.5	70.1	70.6	70.8	68.5	73.7	72.1	70.4	68.4	71.0	63.4	67.8
20-24	67.8	67.6	67.4	66.5	65.2	65.7	65.9	63.5	68.8	67.1	65.5	63.5	66.1	58.4	62.9
25-29	62.9	62.6	62.5	61.6	60.3	60.7	60.9	58.6	63.8	62.2	60.6	58.5	61.1	53.5	57.9
30-34	58.0	57.7	57.6	56.7	55.4	55.9	56.0	53.7	58.8	57.2	55.7	53.6	56.2	48.6	53.0
35-39	53.0	52.7	52.6	51.8	50.5	51.1	51.0	48.8	53.8	52.3	50.7	48.7	51.2	44.0	48.1
40-44	48.1	47.7	47.8	46.9	45.7	46.2	46.1	44.0	48.9	47.3	45.8	43.9	46.3	39.3	43.2
45-49	43.2	42.9	42.9	42.1	40.9	41.4	41.2	39.3	44.1	42.4	41.0	39.1	41.4	34.6	38.4
50-54	38.4	38.1	38.1	37.4	36.3	36.9	36.4	34.6	39.3	37.6	36.2	34.4	36.6	30.0	33.6
55-59	33.6	33.3	33.5	32.8	31.7	32.5	31.7	30.0	34.5	32.9	31.6	29.8	31.9	25.5	29.1
60-64	28.9	28.7	28.9	28.2	27.2	28.1	27.0	25.6	30.0	28.3	27.0	25.4	27.3	21.9	24.7
65-69	24.4	24.1	24.8	23.8	22.8	23.7	22.5	21.5	25.6	23.8	22.6	21.1	22.8	17.7	20.5
70-74	20.2	19.8	21.0	19.6	18.8	19.5	18.2	17.5	21.2	19.5	18.6	17.0	18.5	13.9	16.5
75-79	16.1	15.7	17.5	15.8	14.9	15.5	14.4	13.7	17.3	15.4	15.1	13.3	14.5	10.8	12.9
80-84	12.4	12.0	14.8	12.1	11.5	11.7	11.0	10.5	13.4	12.0	11.8	9.9	10.7	8.3	9.7
85-89	9.9	9.4	12.5	9.1	8.9	9.2	8.5	7.7	10.4	9.2	9.2	7.2	7.9	5.6	7.2
90+	8.4	7.6	9.5	6.7	6.9	7.1	6.9	5.6	7.6	7.8	8.0	5.3	6.0	4.5	5.3

Men															
Age	Asian other	Chinese	Bangladeshi	Black African	Black Caribbea	Black other	Indian	Mixed	Arab	Other	Pakistani	White British	White other	White Gypsy / Traveller	White Irish
U1	84.6	85.4	81.0	83.6	80.8	82.4	82.3	79.7	86.4	83.1	82.3	79.7	81.9	75.4	77.8
01-04	84.0	84.8	80.5	83.2	80.4	82.0	81.7	79.0	85.8	82.5	81.9	79.0	81.2	74.7	77.1
05-09	80.1	80.9	76.5	79.2	76.5	78.0	77.8	75.0	81.8	78.6	78.0	75.0	77.2	70.7	73.1
10-14	75.1	75.9	71.6	74.3	71.6	73.1	72.8	70.1	76.9	73.6	73.1	70.1	72.3	65.7	68.1
15-19	70.1	70.9	66.6	69.3	66.6	68.1	67.8	65.1	71.9	68.6	68.1	65.1	67.3	60.7	63.2
20-24	65.2	66.0	61.7	64.5	61.7	63.2	62.9	60.2	67.0	63.7	63.2	60.2	62.4	55.8	58.2
25-29	60.3	61.0	56.8	59.6	56.9	58.3	58.0	55.3	62.2	58.8	58.3	55.3	57.5	50.8	53.4
30-34	55.3	56.1	51.9	54.7	52.1	53.5	53.0	50.5	57.2	53.8	53.4	50.5	52.6	45.9	48.6
35-39	50.4	51.1	47.0	49.9	47.3	48.8	48.1	45.7	52.3	49.0	48.5	45.7	47.7	41.4	43.8
40-44	45.5	46.3	42.1	45.0	42.5	43.9	43.3	40.9	47.3	44.1	43.7	40.9	42.8	36.8	39.0
45-49	40.7	41.4	37.2	40.2	37.8	39.3	38.5	36.3	42.4	39.2	38.9	36.2	38.0	32.2	34.3
50-54	35.9	36.6	32.5	35.5	33.2	34.9	33.8	31.6	37.6	34.5	34.1	31.5	33.2	27.7	29.7
55-59	31.2	32.0	27.9	30.9	28.8	30.4	29.3	27.2	32.9	29.8	29.5	27.0	28.6	23.3	25.2
60-64	26.7	27.5	23.6	26.3	24.4	26.1	24.9	23.0	28.3	25.2	25.1	22.8	24.2	19.3	21.0
65-69	22.3	23.2	19.7	22.3	20.2	22.0	20.7	19.1	24.1	20.8	21.0	18.7	20.0	15.8	17.1
70-74	18.2	19.1	15.9	18.4	16.3	18.4	16.7	15.4	19.8	16.9	17.0	14.9	16.0	12.6	13.6
75-79	14.4	15.9	13.0	15.4	12.8	15.0	13.2	12.0	16.1	13.1	13.5	11.6	12.3	9.5	10.5
80-84	11.3	13.2	10.6	12.7	9.9	12.1	10.0	9.0	12.9	9.8	10.6	8.6	9.1	7.6	8.0
85-89	9.1	11.1	8.0	10.5	7.3	10.1	7.4	7.0	10.3	6.8	8.4	6.3	6.5	7.9	5.8
90+	7.8	9.4	5.9	10.2	5.7	10.4	6.0	5.7	7.8	4.6	7.4	4.6	5.1	8.6	4.4

Source: ONS Public Health Research Database

OUTMIGRATION ISSUES

ERROR 6 – ERROR IN OUT-MIGRATION ESTIMATES HIGHER FOR MINORITISED ETHNIC GROUPS

The outmigration weights used by the ONS in their life expectancy models may have been the best available source of outmigration estimates available at the time; however, cross-referencing them with subsequent estimates of outmigration shows that they may substantially under-estimate outmigration among people born outside the UK.

The ONS has used several approaches to estimate outmigration. The International Passenger Survey (IPS) is a continuous survey conducted at major ports of entry and exit to the UK. The IPS asks passengers what the intended purpose of their travel is; from this, the ONS can estimate international outmigration. The IPS covers approximately 800,000 (0.34%) of the 240 million passengers travelling through UK ports annually; although only 3,000 – 4,000 of these passengers are defined as long-term international migrants [50]. The ONS use the IPS to produce estimates of long-term international migrants (LTIM estimates) [51]. These are produced mainly using IPS data, with some additional data, which ONS describe as adjustments to "asylum seekers, visitor and migrant switchers and flows to and from Northern Ireland" [51]. In recent years, the ONS have been focused on working towards an admin-based approach to calculating international migration using Registration and Population Interaction Database (RAPID) data [52]. Each of these methods has produced varying estimates of the level of outmigration from the UK among UK-born and foreign-born residents.

Estimates of outmigration by ethnic group are not produced by the ONS. The weights used in their life expectancy estimates were created by using the ONS Longitudinal Study (ONS-LS), a longitudinal survey of approximately 1% of the population of England and Wales which records observed emigration events of study members and unobserved likely emigration events following de-registration from the Patient Register. As this survey contains ethnic group data, it could be used to estimate emigration by ethnic group. These figures were combined with figures from the IPS in order to produce more accurate estimates. The estimates were produced according to ethnic group by sex by age group (<25, 25-44, 45-64, 65+).

The major limitation with using ONS-LS data is that outmigration can only be estimated by reviewing de-registrations from the NHS, and those who do not provide an exit date become lost to follow-up ([53] cited by [54]). While the NHS advises all patients to cancel their NHS registration if they plan to emigrate [54], this is only advisory, and many de-registrations are likely to have been missed, or not taken place, should the patient have decided to stay; however, cancelled ciphers allows some of the attrition to be accounted for [55]. The International Passenger Survey samples only approximately 3-4,000 international migrants a year, meaning that the sample size is likely to be small for many birth countries [56]. Furthermore, the IPS asks respondents about their intentions to migrate, which can be subject to change.

We tested how the outmigration assumptions used in the ONS life expectancy estimates compared with other estimates of outmigration. We used results from two ONS publications, one based on LTIM data [51], and one based on an experimental, in-development, methodology

published by the ONS in 2023 [57]. The latter approach used various datasets to produce the estimates: figures for non-EU nationals were based on Home Office Borders and Immigration data, figures for EU-nationals were based on Registration and Population Interaction Database (RAPID) data, figures for British nationals were based on the IPS. We aimed to summarise our estimates over the period of our study (March 2011 - March 2014); however, in the ONS' 2023 migration estimate methodology, estimates were only available from June 2012. To compare these alternative estimates of outmigration with our model, we attempted to summarise our outmigration assumptions by country of birth. However, the outmigration estimates were not defined according to country of birth, being specified by age, sex, and ethnic group. As such, we calculated the proportion of people born inside/outside the UK within each age-sex-ethnic group cohort, and used this to estimate the country of birth of outmigrants within our model. To achieve this, we used Census data stratified by ethnic group, sex, age group, and country of birth (inside/outside UK). We multiplied this by our outmigration weights to estimate outmigration split by country of birth (table 4). Although we recognise that British nationality is not synonymous with being born in the UK, the ONS' alternative estimates of outmigration suggest that our model substantially under-estimates outmigration among people born outside the UK. As we theorise that mortality would be higher for outmigration among people returning to their countries of origin, this is an important factor in our model.

Table 4: Outmigration by country of birth in the life expectancy model Vs LTIM estimates

Implied in our model:		ONS methodology:	2014	2023
Born inside UK	194,686	**British nationals**	155,375	143,000
Born outside UK	52,558	**Not British nationals**	307,500	190,000

Source: Our model: ONS Public Health Research Database. ONS 2014 methodology: *"Migration Statistics Quarterly Report: November 2014"* [51] (figure 3.1, averaged for Mar 2011 – Mar 2014). ONS 2023 methodology: *"Long-term international migration, provisional: year ending June 2023"* (figure 2, averaged for Jun 2012 – Mar 2014) [57]

As our study population comprises people participating in the 2011 Census, individuals who generally do not ever participate in Censuses will not be included. However, if we know the amount of dropout between Censuses, i.e. the number of people who took part in one Census, but were not recorded at the next Census, this may be a useful indicator of outmigration. While there will be a proportion of people who took part in the 2001 Census, but decided not to participate in the 2011 Census, the likelihood of participation in the Census increases with age; therefore, if someone has participated in one Census, then it is likely that they will participate in the next. As such, we could assume a large proportion of people who are "lost to follow up" between Censuses have in fact outmigrated. While it has been assumed elsewhere that people who are lost to follow-up in the ONS-LS may be lost due to unrecorded outmigration [54], it is also feasible that they could be lost due to incorrect information being recorded [54]; the exact reason for loss to follow-up in the ONS-LS cannot be known ([58] cited by [54]). Those who are lost to follow up, but have actually migrated, are likely to become "statistically immortal" in our

analyses. Additionally, there is evidence that the mortality rate of outmigrants from minoritised ethnic groups is higher than those who remain [18].

Datasets linking the 2011 Census to the 2021 Census were not available when our research was being conducted; however, we were able to estimate loss to follow-up using linked 2001 and 2011 Census data. Table 5 shows the proportion of participants participating in the 2001 Census, who were subsequently lost to follow-up (or who were recorded as having outmigrated) in the 2011 Census. We see that loss to follow-up is far higher in minoritised ethnic groups compared with the White British group.

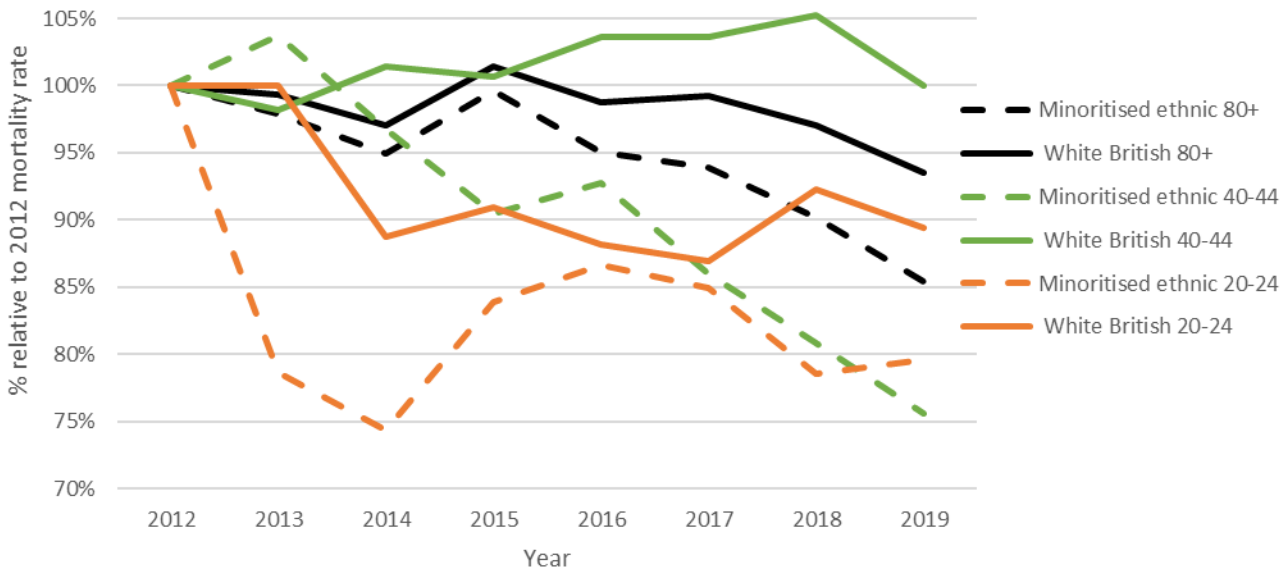
Table 5: Proportion of ONS-LS participants who either outmigrated or were lost to follow-up between 2001 and 2011 Census by age group and ethnic group

Ethnic group	0-18	19-29	30-39	40-49	50-59	60+
Asian Other	18%	33%	22%	19%	16%	19%
Bangladeshi	18%	18%	12%	16%	14%	28%
Black African	27%	38%	27%	26%	18%	26%
Black Caribbean	21%	26%	20%	18%	19%	18%
Black Other	26%	30%	21%	27%	*	*
Chinese	30%	49%	27%	18%	22%	22%
Indian	12%	18%	14%	11%	13%	16%
Mixed	20%	29%	24%	17%	14%	17%
Other	37%	47%	43%	28%	25%	27%
Pakistani	19%	23%	16%	16%	16%	20%
White British	14%	14%	11%	8%	7%	6%
White Irish	31%	43%	25%	18%	17%	11%
White Other	36%	58%	40%	27%	21%	15%

Source: ONS Longitudinal Study. * denotes cells suppressed due to statistical disclosure

ERROR 7 – DEATHS OF OUTMIGRANTS NOT RECORDED

The ONS death registration data does not record whether someone who migrated later died. There is a risk that individuals may have migrated, but neither their emigration nor their deaths are recorded, rendering them statistically immortal. This risk is compounded over time, as the section of the population defined in March 2011 who subsequently emigrated increases over time. To investigate the impact of this risk, we explored how the mortality rate of minoritised groups changed over time, comparing mortality for minoritised ethnic groups against the White group. We excluded 2011 as the mortality figures for April and May 2011 were significantly lower than the average mortality in those months from 2012-2019 (see Appendix A). We present data for selected age groups (20-24, 40-44, and 80+); it is not appropriate to provide an aggregate mortality rate, as the changing population age structure for each group over time would influence results. Although mortality varied substantially during this period, it can be observed that mortality rate falls more quickly over time for minoritised ethnic groups (figure 3). One potential explanation for this pattern is statistical immortality., although we also consider that the greater rate of fall among minoritised groups may be an artefact of differences in age structure between the two cohorts.



Source: ONS Public Health Research Database

Figure 3: Change in mortality rate since 2012 (ages 20-24, 40-44 and 80+, all genders)

To investigate reasons behind the falling mortality rate, we comprehensively tabulated the missing data in our model (table 6). We compared the weighted sample size of our model with published Census counts to estimate the proportion of each ethnic group's population which remains unaccounted for in our model. The table demonstrates how the section of each cohort whose outcome is not known by our model varies by ethnic group.

To illustrate, for the Black Other group, the 2011 Census reports a population of 280,437. In our model, our weighted total, excluding those not usually resident, gives a population of 190,912. Of these, we do not know the outcome of 18,612 people whose data could not be mapped to the patient register, and we are estimating that we would not be able to know the outcome of 3,527 people who have been effectively excluded through the application of outmigration weights. This leaves a deficit of 111,664 against the 2011 Census totals: these are people whose outcome is unknown. The summary row in table 6 entitled "% not enumerated in the 2011 Census" shows the ONS' estimates of under-enumeration from the 2011 Census [47]; we see that most of the difference between the published census counts and the counts in our data is due to under-enumeration. However, we see that the proportion for whom their mortality outcome is unknown in our model is substantially higher among minoritised ethnic groups than it is for the White group. It should also be noted that the White group in this table includes White minority groups such as White Other, who likely have substantially higher missingness than the White British group.

Table 6: Missingness in life expectancy model by ethnic group

	Asian other	Bangla-deshi	Black African	Black Caribbean	Black other	Indian	Mixed	Other	Pakistani	White
Published Census counts	1,228,861	447,201	989,628	594,825	280,437	1,412,958	1,224,400	563,696	1,124,511	48,209,395
Counts in the data*	1,075,037	415,649	876,125	545,619	190,912	1,357,663	1,034,214	430,074	1,056,201	45,778,511
Not usually resident	1,916	80	334	51	27	1,014	222	443	323	3,886
Died during period	4,863	1,808	2,685	8,290	900	11,796	5,083	2,289	5,629	1,230,433
Did not map to patient register	131,433	33,679	103,902	42,315	18,612	101,245	74,149	42,656	74,640	1,885,693
Excluded due to outmigration weights	36,348	6,895	28,544	8,735	3,527	34,589	18,773	19,353	15,184	483,058
Remaining	900,477	373,187	740,660	486,228	167,846	1,209,019	935,987	365,333	960,425	42,175,441
Outcome unknown	**321,605**	**72,126**	**245,949**	**100,256**	**111,664**	**191,129**	**283,108**	**195,631**	**158,134**	**4,799,635**
Comparison to published Census counts:										
Counts in the data*	87.5%	92.9%	88.5%	91.7%	68.1%	96.1%	84.5%	76.3%	93.9%	95.0%
Not usually resident										
Died during period	0.4%	0.4%	0.3%	1.4%	0.3%	0.8%	0.4%	0.4%	0.5%	2.6%
Did not map to patient register	10.7%	7.5%	10.5%	7.1%	6.6%	7.2%	6.1%	7.6%	6.6%	3.9%
Excluded due to outmigration weights	3.0%	1.5%	2.9%	1.5%	1.3%	2.4%	1.5%	3.4%	1.4%	1.0%
Remaining	73.3%	83.4%	74.8%	81.7%	59.9%	85.6%	76.4%	64.8%	85.4%	87.5%
% outcome unknown	**26.2%**	**16.1%**	**24.9%**	**16.9%**	**39.8%**	**13.5%**	**23.1%**	**34.7%**	**14.1%**	**10.0%**
% not enumerated in 2011 Census^	15.0%	7.0%	12.0%	8.0%	36.0%	6.0%	16.5%	27.0%	6.0%	5.0%
% adjusted for in model	13.7%	9.1%	13.4%	8.6%	7.9%	9.6%	7.6%	11.0%	8.0%	4.9%

Source: ONS Public Health Research Database (except ^, source: ONS Census response rate tables [47])
*excluding not usually resident population. All figures from PHRD are weighted, except deaths which are individual-level counts.

If we were to assume that those who were missing had the same likelihood of having died in the study period as the observable population (taking their age into account), then this missingness would not impact the life expectancy results. However, the proportion of people not responding to the Census is estimated by the ONS using a model that considers many area-level characteristics including proportion of individuals claiming Jobseeker's Allowance, relative median house price, dwelling density, proportion of people aged 16-29 years old, and rate of notable offences recorded [59]. Several of these variables are potentially correlated with socioeconomic status, which itself correlates highly with life expectancy. As such, under-enumeration in the Census cannot be considered to be data "missing at random", and has

therefore implications for life expectancy (people not enumerated in the Census would have lower life expectancy based on their socioeconomic profile). It should also be noted that one of the predictors for non-response at an area level is the proportion of the school population that is from a minoritised ethnic group [59], suggesting that the ONS researchers who devised the Census weighting methodology had found that minoritised ethnicity was strongly correlated with non-response; although the authors do not offer reasons as to why this might be.

SALMON BIAS

ERROR 8 – MORTALITY RATE OF MIGRANTS HIGHER FOR MINORITISED ETHNIC GROUPS

The mortality of migrants is not simulated in our model, as outmigrants are effectively excluded from our calculations. As per the salmon bias hypothesis, the mortality rate of outmigrants may differ from migrants remaining in their host country. Guillot and colleagues [18] demonstrated this variability in mortality rate using French pensions data, showing elevated mortality hazard ratios associated with outmigrating (as opposed to remaining in France) among migrants born in Eastern and Central Europe (Mortality Hazard Ratio 1.5), Southern Europe (MHR 2), Africa (MHR 2.5) and Asia and elsewhere (MHR 3.5).

The mortality rate of UK-born people emigrating from the UK may also differ from the remaining UK-born population, but in the converse direction. Health selection among the UK-born population may mean that those who emigrate to a new country are more likely to be healthy (see [60] for a discussion on health selection among migrants). As such, those who stay in the country are likely to have a higher mortality rate and lower LE estimate.

ERROR 9 – SENSITIVITY OF LE ESTIMATES TO MISSINGNESS AT OLDER AGE

(Not on map)

During our calibration process, we noticed that the LE estimates were much more sensitive to missingness at older age, and the lower the sample size in that cohort, the greater the sensitivity. Table 5 in section E6 has demonstrated that missingness is likely higher for people from minoritised ethnic groups in the oldest age groups, making this a potentially important factor in the sensitivity of the life expectancy estimates.

To further explore this observation, we tested the sensitivity of the life expectancy at birth estimates to the mortality rates in these age groups in two ways. Firstly, we selected an ethnic group with higher life expectancy at birth than the White British group. Bangladeshi women had a life expectancy of 86.8 years at birth, compared with 83.0 for White British women. To assess the impact of the small sample size in the Bangladeshi group for those aged 80+, we substituted in the White British mortality rate at these ages. The results can be seen in table 7, demonstrating

that the Bangladeshi life expectancy at birth estimate reduces to 83.4. This highlights how critical the accuracy of the data is at older ages in ensuring robust life expectancy estimates. To illustrate the risk involved, there were only 80, 23 and 21 deaths (rounded to a whole number) among Bangladeshi women in the 80-84, 85-89 and 90+ age groups, respectively, during our study period. Introducing the White British mortality rates increased these to 106, 63 and 38. Therefore, just 84 missing deaths could cause a substantial change in the life expectancy at birth estimate. We also estimated just substituting in the White British mortality rate for the 90+ group, resulting in an increase of 17 deaths: the life expectancy at birth estimate for Bangladeshi women reduced to 84.8, representing a change of 2 years.

To investigate this further, and to demonstrate the small amount of missing data that would be required to cause a considerable change in the life expectancy estimates, we simulated what would happen if we introduced 20 additional deaths into each age group-sex-ethnic group cohort. Table 8 shows the change in years to the life expectancy at birth estimate when these deaths were introduced. We see that missing deaths at age 90+ have the largest impact on life expectancy at birth for many ethnic groups. Bangladeshi and Arab women see an increase of more than 2 years in life expectancy at birth with the addition of these deaths. Conversely, the estimates of the White British group experienced only a nominal change. Given the large proportion of missingness seen in table 8, it is fully conceivable that a number of deaths of this magnitude may not have been recorded due to statistical immortality.

Table 7: Effect on life expectancy at birth on Bangladeshi women of substituting in White British mortality rate for those aged 80+.

Age group	Bangladeshi - original	White British	Bangladeshi – substituted mortality rate
<1	86.8	83.0	83.4
01-04	86.3	82.3	82.8
05-09	82.3	78.4	78.9
10-14	77.3	73.4	73.9
15-19	72.4	68.4	68.9
20-24	67.4	63.5	64.0
25-29	62.5	58.5	59.0
30-34	57.6	53.6	54.1
35-39	52.6	48.7	49.2
40-44	47.7	43.9	44.3
45-49	42.9	39.1	39.4
50-54	38.1	34.4	34.6
55-59	33.5	29.8	29.9
60-64	28.9	25.4	25.3
65-69	24.8	21.1	21.1
70-74	21.0	17.0	17.1
75-79	17.5	13.3	13.3
80-84	14.8	9.9	9.9
85-89	12.5	7.2	7.2
90+	9.5	5.3	5.3

Source: ONS Public Health Research Database

Table 8: Effect on life expectancy at birth of adding 20 deaths into the mortality figures for each age group-sex-ethnic group cohort

Age	Sex	Arab	Asian.othe	Bangladesh	Black.Afric	Black.Carib	Black.othe	Chinese	Indian	Mixed	Other	Pakistani	White.Briti	White.Gyp	White.Irish	White.othe
<1	Females	-0.22	-0.14	-0.12	-0.04	-0.17	-0.25	-0.32	-0.05	-0.03	-0.10	-0.04	0.00	-0.17	-0.06	-0.04
01-04	Females	-0.35	-0.11	-0.13	-0.06	-0.24	-0.23	-0.39	-0.07	-0.03	-0.33	-0.05	0.00	-1.11	-0.75	-0.04
05-09	Females	-0.33	-0.11	-0.11	-0.06	-0.20	-0.24	-0.45	-0.07	-0.03	-0.32	-0.05	0.00	-1.04	-0.61	-0.05
10-14	Females	-0.39	-0.11	-0.11	-0.06	-0.17	-0.24	-0.40	-0.07	-0.04	-0.32	-0.05	0.00	-0.78	-0.51	-0.06
15-19	Females	-0.43	-0.11	-0.13	-0.06	-0.13	-0.25	-0.20	-0.06	-0.04	-0.30	-0.06	0.00	-0.79	-0.40	-0.05
20-24	Females	-0.40	-0.09	-0.12	-0.06	-0.11	-0.34	-0.08	-0.05	-0.04	-0.24	-0.05	0.00	-0.95	-0.23	-0.03
25-29	Females	-0.28	-0.06	-0.09	-0.05	-0.09	-0.32	-0.10	-0.03	-0.04	-0.17	-0.04	0.00	-0.92	-0.15	-0.01
30-34	Females	-0.23	-0.05	-0.08	-0.03	-0.10	-0.27	-0.11	-0.03	-0.04	-0.14	-0.03	0.00	-0.87	-0.12	-0.01
35-39	Females	-0.26	-0.04	-0.09	-0.03	-0.09	-0.24	-0.11	-0.03	-0.05	-0.14	-0.04	0.00	-0.83	-0.11	-0.01
40-44	Females	-0.31	-0.04	-0.12	-0.03	-0.05	-0.17	-0.13	-0.03	-0.05	-0.14	-0.04	0.00	-0.62	-0.08	-0.01
45-49	Females	-0.31	-0.05	-0.18	-0.03	-0.04	-0.14	-0.12	-0.03	-0.05	-0.14	-0.06	0.00	-0.64	-0.06	-0.02
50-54	Females	-0.38	-0.06	-0.22	-0.05	-0.04	-0.20	-0.12	-0.03	-0.06	-0.17	-0.05	0.00	-0.62	-0.05	-0.02
55-59	Females	-0.57	-0.06	-0.17	-0.08	-0.06	-0.44	-0.11	-0.03	-0.08	-0.18	-0.05	0.00	-0.88	-0.04	-0.02
60-64	Females	-0.66	-0.06	-0.19	-0.12	-0.07	-0.51	-0.13	-0.03	-0.10	-0.16	-0.07	0.00	-0.62	-0.03	-0.02
65-69	Females	-0.71	-0.09	-0.27	-0.12	-0.06	-0.40	-0.19	-0.03	-0.10	-0.19	-0.08	0.00	-0.76	-0.02	-0.02
70-74	Females	-0.72	-0.10	-0.21	-0.12	-0.04	-0.36	-0.20	-0.03	-0.11	-0.20	-0.08	0.00	-0.76	-0.02	-0.02
75-79	Females	-0.58	-0.10	-0.26	-0.20	-0.03	-0.39	-0.19	-0.03	-0.09	-0.27	-0.07	0.00	-0.51	-0.02	-0.02
80-84	Females	-0.92	-0.14	-0.43	-0.27	-0.04	-0.53	-0.22	-0.03	-0.07	-0.19	-0.08	0.00	-0.48	-0.01	-0.01
85-89	Females	-1.14	-0.15	-0.84	-0.52	-0.05	-0.84	-0.25	-0.04	-0.07	-0.26	-0.15	0.00	-0.53	-0.01	-0.01
90+	Females	-2.13	-0.41	-2.22	-1.07	-0.12	-1.32	-0.49	-0.07	-0.10	-0.47	-0.33	0.00	-0.37	-0.01	-0.02
<1	Males	-0.20	-0.13	-0.10	-0.04	-0.16	-0.23	-0.30	-0.05	-0.03	-0.08	-0.04	0.00	-0.16	-0.06	-0.03
01-04	Males	-0.31	-0.11	-0.11	-0.05	-0.23	-0.22	-0.35	-0.06	-0.03	-0.32	-0.05	0.00	-0.96	-0.65	-0.03
05-09	Males	-0.29	-0.10	-0.10	-0.06	-0.17	-0.20	-0.46	-0.06	-0.03	-0.29	-0.04	0.00	-0.96	-0.59	-0.04
10-14	Males	-0.33	-0.10	-0.11	-0.06	-0.15	-0.22	-0.41	-0.06	-0.04	-0.31	-0.05	0.00	-0.75	-0.50	-0.05
15-19	Males	-0.36	-0.09	-0.11	-0.06	-0.12	-0.22	-0.21	-0.06	-0.04	-0.25	-0.05	0.00	-0.86	-0.36	-0.05
20-24	Males	-0.32	-0.09	-0.11	-0.08	-0.13	-0.32	-0.11	-0.04	-0.04	-0.25	-0.04	0.00	-0.93	-0.25	-0.03
25-29	Males	-0.26	-0.08	-0.10	-0.07	-0.13	-0.38	-0.14	-0.03	-0.05	-0.17	-0.04	0.00	-0.89	-0.16	-0.01
30-34	Males	-0.22	-0.06	-0.08	-0.05	-0.12	-0.32	-0.14	-0.02	-0.05	-0.13	-0.03	0.00	-0.87	-0.12	-0.01
35-39	Males	-0.21	-0.04	-0.07	-0.04	-0.11	-0.31	-0.16	-0.02	-0.06	-0.11	-0.03	0.00	-0.75	-0.10	-0.01
40-44	Males	-0.18	-0.05	-0.08	-0.04	-0.07	-0.18	-0.17	-0.03	-0.05	-0.12	-0.03	0.00	-0.66	-0.07	-0.01
45-49	Males	-0.21	-0.06	-0.12	-0.04	-0.04	-0.14	-0.15	-0.03	-0.05	-0.13	-0.05	0.00	-0.49	-0.05	-0.02
50-54	Males	-0.25	-0.06	-0.12	-0.05	-0.05	-0.21	-0.14	-0.03	-0.06	-0.13	-0.05	0.00	-0.50	-0.05	-0.02
55-59	Males	-0.26	-0.07	-0.15	-0.07	-0.06	-0.42	-0.14	-0.02	-0.09	-0.14	-0.04	0.00	-0.52	-0.04	-0.02
60-64	Males	-0.34	-0.07	-0.35	-0.12	-0.09	-0.68	-0.16	-0.03	-0.10	-0.16	-0.06	0.00	-0.50	-0.03	-0.02
65-69	Males	-0.35	-0.09	-0.32	-0.18	-0.08	-0.56	-0.20	-0.03	-0.09	-0.17	-0.08	0.00	-0.58	-0.02	-0.02
70-74	Males	-0.49	-0.10	-0.12	-0.15	-0.04	-0.49	-0.20	-0.03	-0.08	-0.18	-0.05	0.00	-0.49	-0.02	-0.02
75-79	Males	-0.49	-0.10	-0.11	-0.15	-0.03	-0.43	-0.18	-0.02	-0.07	-0.22	-0.04	0.00	-0.57	-0.01	-0.02
80-84	Males	-0.55	-0.13	-0.12	-0.28	-0.03	-0.53	-0.23	-0.03	-0.07	-0.16	-0.06	0.00	-0.62	-0.01	-0.02
85-89	Males	-0.72	-0.19	-0.28	-0.55	-0.04	-0.83	-0.35	-0.04	-0.06	-0.18	-0.09	0.00	-0.66	-0.01	-0.01
90+	Males	-1.68	-0.44	-0.75	-1.24	-0.09	-1.62	-0.98	-0.07	-0.12	-0.27	-0.25	0.00	-0.53	-0.02	-0.02

Source: ONS Public Health Research Database

SUMMARY

In this section, we discussed the potential sources of error in the ONS' methodology for producing life expectancy estimates by ethnic group. We have summarised these potential sources of error in table 9, detailing for each potential source of error the amount of imprecision or error, the likely impact of that error upon LE estimates, and then the potential to correct that error in some form. We can see that there is considerable potential for error in the estimates, and that much of this error is difficult to mitigate.

Given this difficulty to mitigate, which is driven by the availability of suitable data, we conclude that it may not be possible to create revised estimates of life expectancy. Instead, for the remainder of the report, we will focus on simulating the sensitivity to error in outmigration and assumptions around the mortality of outmigrants, and extending the existing analyses. We present the details of this work in the next section, *Stage 3: Extension of results*.

Table 9: Summary of potential sources of error

Error	Amount of error	Impact of error	Potential to mitigate
Linkage failure and under-enumeration			
E1 - Patient register (PR) match - MITIGATED UNKNOWN ERROR	**High** – those from minoritised ethnic groups had a much higher likelihood of not matching to the PR.	**Medium** – non PR matches could have higher mortality rate than general population.	**Low**- see text.
E2 – Death records unmatched - UNMITIGATED KNOWN ERROR	**Medium** – high numbers of unmatched deaths from non-UK countries.	**Medium** – dependent on whether mortality rate for Census participants is the same as for those who did participate, which is not known.	**Medium** – could be simulated.
E3 - Under-enumeration in Census was higher for minoritised ethnic groups - UNMITIGATED KNOWN ERROR	**High** for people from minoritised groups.	**Medium** – if missingness due to under-enumeration is not at random.	**Low** – could be simulated, but would reinforce existing mortality rate.
Sampling issues			
E4 - No visibility of people who have migrated before sampling window - UNMITIGATED UNKNOWN ERROR *(Not on map)*	**High** for people from minoritised groups.	**High** – Guillot et al. demonstrate that it is necessary to consider this cohort in order to demonstrate the full impact of salmon bias.	**Low** – the data required to demonstrate this is presently unavailable.
E5 – Aggregated White group masks variation between sub-groups - MITIGATED KNOWN ERROR *(Not on map)*	**High** – masks substantial variation in life expectancy within the White ethnic group.	**High** – Sub-groups, particularly Gypsy/Traveller group, have very different outcomes.	**High** – subgroups are easily split out.
Outmigration issues			
E6 – Error in out-migration estimates higher for minoritised ethnic groups - MITIGATED KNOWN ERROR	**Medium** – quite large errors in outmigration when compared to other ONS estimates of outmigration by country of birth.	**Medium** – the result is an moderate overestimation of the population for many minoritised ethnic groups.	**High** – we can alter outmigration to more accurately reflect other estimates.
E7 – Deaths of outmigrants not recorded - MITIGATED KNOWN ERROR	**Medium** – there are likely considerable missing deaths.	**High** – not only are there missing deaths, but these deaths are likely to be not missing at random.	**Medium** – can improve with alternative outmigration estimates, and data from other studies on mortality rate of returnees.
Salmon bias			
E8 – Mortality rate of migrants higher for minoritised ethnic groups - MITIGATED KNOWN ERROR	**High** – Guillot et al. suggest considerably higher mortality among returnees.	**High** – Guillot et al. demonstrate that correcting for this can remove migrant mortality advantage.	**Medium** – we can simulate higher mortality rate, but it would be estimate based on Guillot et al, which is data from France, not the UK.
E9 – Sensitivity of LE estimates to missingness at older age - MITIGATED KNOWN ERROR *(Not on map)*	Depends on missingness in each cohort.	**High** – Loss of 20 deaths in the age 90+ cohort causes large changes to life expectancy in some ethnic groups, but no change for White British group. High missingness at older ages for some minoritised ethnic groups.	**Medium** – it is dependent on the other errors. Again, we can simulate it, but not correct using actual data.

Key for missing data blocks:

MITIGATED UNKNOWN ERROR – the error cannot be quantified, but attempts have been made to account for it.

MITIGATED KNOWN ERROR – the error can be quantified, and attempts have been made to account for it.

UNMITIGATED KNOWN ERROR – the error can be quantified, but no attempt has been made to correct for it.

UNMITIGATED UNKNOWN ERROR – the error cannot be quantified, and no attempt has been made to correct for it.

STAGE 3: EXTENSION OF RESULTS

Following our exploration of the potential sources of error, we conducted a series of experiments to produce additional estimates of life expectancy. In the first experiment, we simulate the sensitivity of life expectancy estimates to alternative assumptions. Although we had aimed to produce revised estimates of life expectancy, our exploration of the potential sources of error demonstrated to us that the data we would require to produce accurate estimates of life expectancy by ethnic group in England and Wales is not available. As such, the purpose of these experiments was not to produce revised estimates, but to demonstrate the sensitivity across the likely range of error in the estimates. We then produced estimates of life expectancy with minoritised ethnic groups split into UK-born and foreign-born cohorts. Finally, we produced estimates of disability-free life expectancy.

TESTING SENSITIVITY OF LIFE EXPECTANCY ESTIMATES TO OUTMIGRATION ASSUMPTIONS

Given i) the increased risk of missingness among those from minoritised ethnic groups, particularly at the eldest ages (80+), ii) the very small populations among minoritised ethnic groups at the eldest ages which are prone to high random error, iii) the evidence on higher risk of mortality among returnees [18] and the salmon bias hypothesis, and iv) the incomplete data (and large variation) in outmigration estimates, we devised a simulation to test the sensitivity of life expectancy estimates by ethnic group to these factors.

Our aim was to use an alternative source of data to inform outmigration, and to simulate the effects of a higher mortality rate among outmigrants. Our approach involved using loss to follow-up data from the ONS-LS by age group, sex, and ethnic group (scaled linearly to represent our study period) as our estimate of outmigration, and using the mortality rates of outmigrants reported in the work of Guillot et al. [18]. We then tested the sensitivity of the life expectancy estimates to these new assumptions on the quantity of outmigration, and the mortality rate of outmigrants.

The loss to follow-up figures from the ONS Longitudinal Study demonstrated that dropout/loss to follow-up, and hence statistical immortality, is likely to be higher among minoritised ethnic groups. Life expectancy varied greatly across outmigrant mortality scenarios. Figure 4 shows LE estimates at birth by ethnic group for each of those scenarios. The chart shows estimates from our best-calibrating model using the ONS' methodology. We then plot our model with our assumptions around outmigration, using a series of mortality scaling scenarios. In using these different methods and assumptions, we observe large variation in life expectancy estimates at birth for most minoritised ethnic groups. In contrast, the life expectancy estimates for the White British group are relatively immune to these methodological and assumption changes.

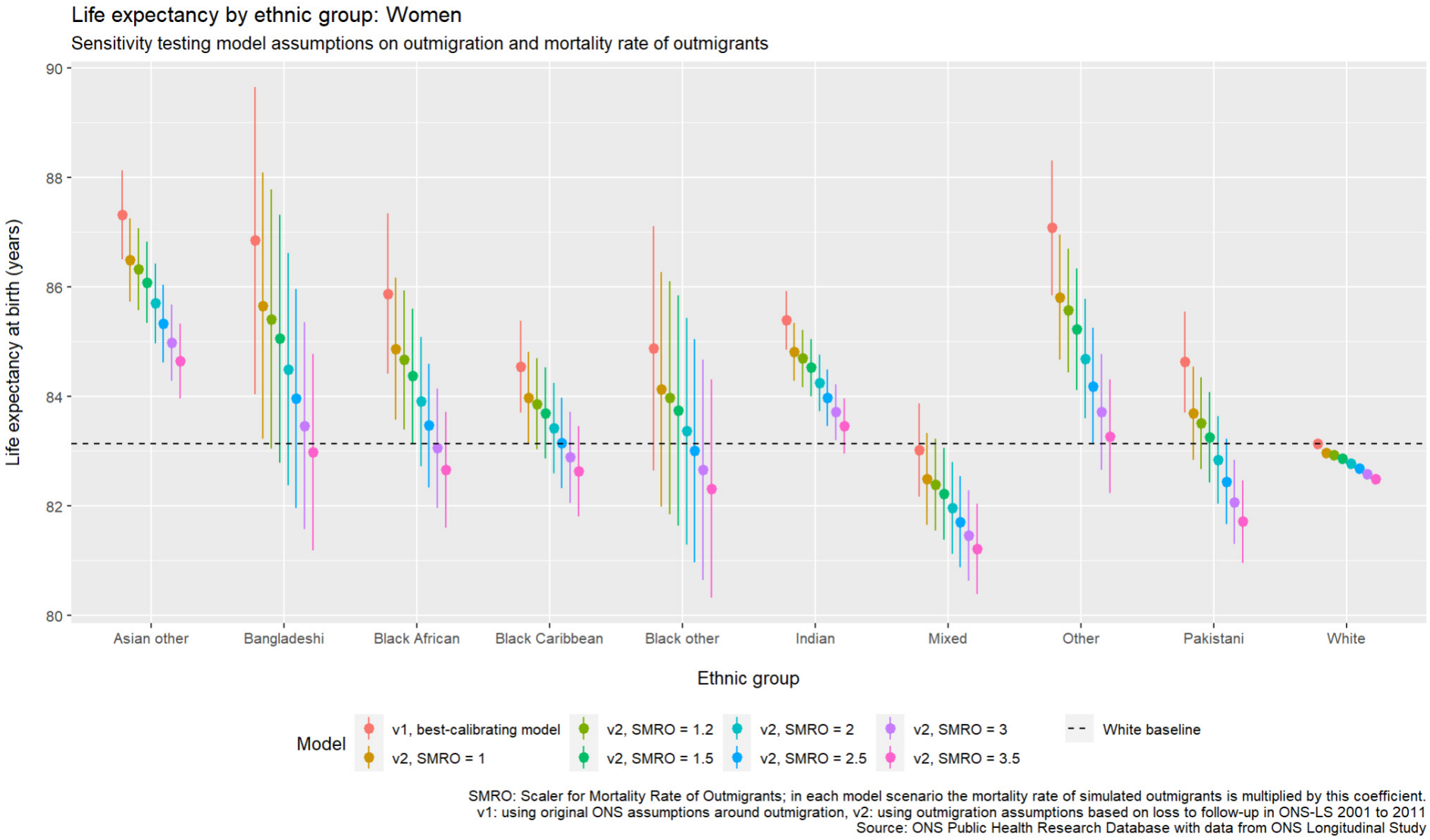


Figure 4: LE estimates at birth by ethnic group, testing sensitivity of life expectancy estimates to outmigration assumptions (women)

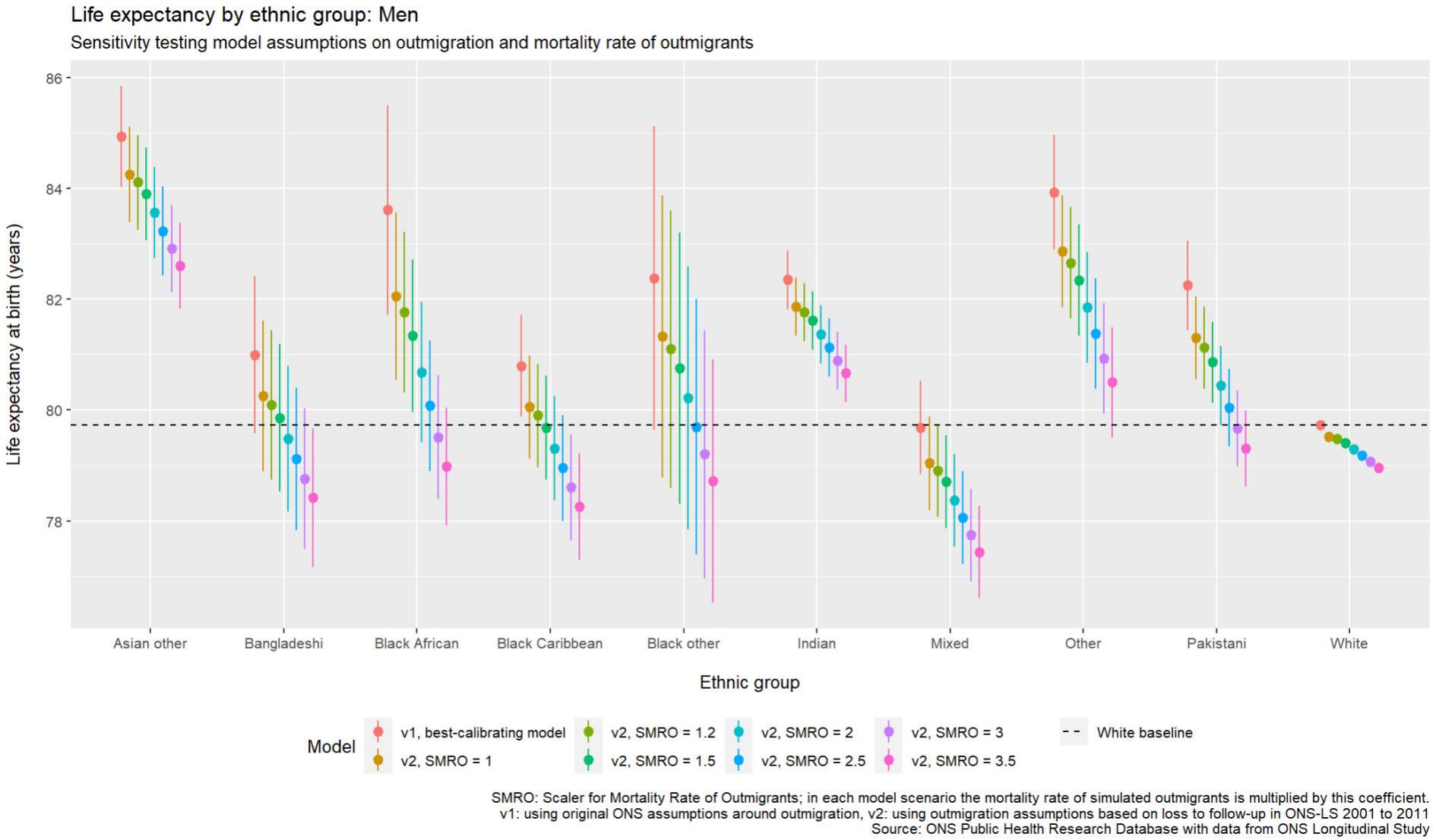


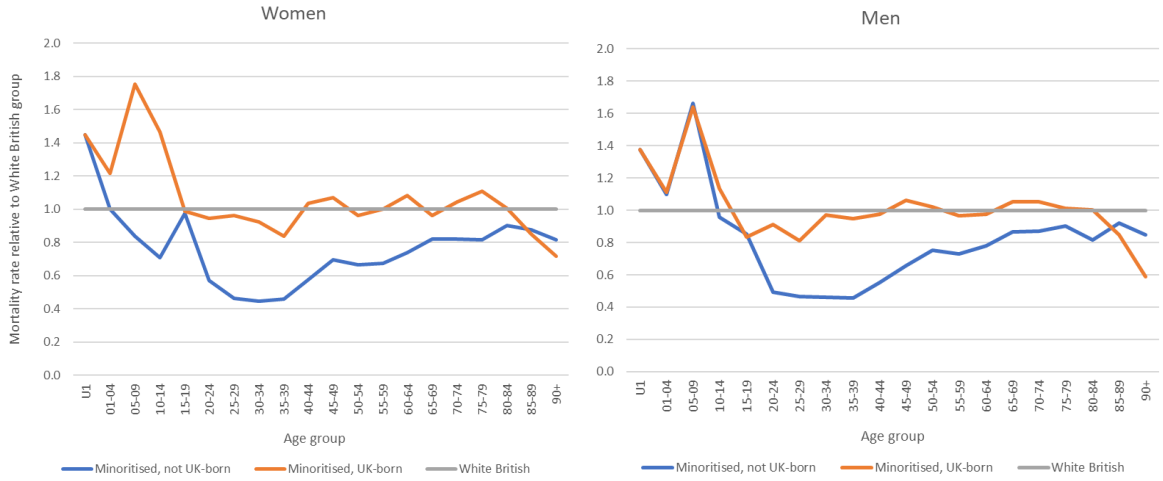
Figure 5: LE estimates at birth by ethnic group, testing sensitivity of life expectancy estimates to outmigration assumptions (men)

LE ESTIMATES FOR UK-BORN AND FOREIGN-BORN COHORTS

To understand the likely impact of underestimating outmigration among those born outside the UK, we graphed mortality rate for minoritised groups born inside the UK, minoritised groups born outside the UK, and the White British group.

As in previous steps, we excluded 2011 due to April and May 2011 having mortality that appeared anomalous with the monthly figures seen in 2012-2020. We used the mortality rate for the White British group as the reference, and observed how mortality rates differed with age for combined minoritised ethnic groups born inside and outside the UK. Figure 6 shows the results, demonstrating how mortality rates are higher for both minoritised groups before age 20. For the UK-born minoritised cohort, mortality rates were then consistent with age until approximately age 80, where they declined sharply.

Figure 6 demonstrates how life expectancy results vary substantially by country of birth, showing that mortality rates were substantially lower for the foreign-born minoritised cohort from approximately age 20 until death. A potential explanation for this is missing data in the foreign-born minoritised cohort leading to statistical immortality. A similar observation was made in research by Bhopal and colleagues, whereby the observed mortality advantage among minoritised ethnic groups was lower for the UK or Irish-born cohort than it was for migrants [14]. It should be noted that much of the literature that has referenced the ONS experimental life expectancy statistics does not draw a distinction between migrants and UK-born minoritised cohorts, reporting simply on there being higher life expectancy among minoritised ethnic groups in the UK. Despite our making this observation, we would suggest that the life expectancy differences between UK-born and foreign-born minoritised groups in the UK may be indicative of methodological shortcomings, rather than actual differences.



Source: ONS Public Health Research Database. Mortality rate of not-UK born minoritised ethnic group girls aged 1-4 years was suppressed due to small sample size, and has been set to equal that of the White British group for charting purposes.

Figure 6: Age-specific mortality rates relative to White British group for minoritised ethnic groups (UK-born and born outside UK)

Figure 6 also shows that, despite the similarities between the mortality rate of the UK-born minoritised ethnic cohort and the White British group, life expectancy at birth was higher for the UK-born minoritised cohort (83.9 Vs 83.3 for women; 80.8 Vs 80.0 for men; source: ONS Public Health Research Database). This is likely an artefact of the previously discussed sensitivity of life expectancy estimates to the eldest age groups, due to the population structure of the minoritised ethnic group. The sample size for the 85-89 and 90+ age groups was relatively small for the UK-born minoritised cohort, meaning that we were unable to produce estimates by ethnic group and country of birth. It is also important to note that the aggregated nature of the UK-born minoritised ethnic category masks the heterogeneity in ethnic group at each age range. For example, the ethnic group make-up of the eldest age groups in this data will be influenced by migration patterns that occurred during the 1920s; as such, ethnic groups whose migration trends into the UK are more recent will likely be under-represented in this cohort.

DISABILITY FREE LIFE EXPECTANCY (DFLE) ESTIMATES

Previous studies using data from the 2001 Census in Scotland [61], and England and Wales [62], have found significantly lower healthy life expectancy among people from minoritised ethnic groups. This is to be expected given well-documented ethnic inequalities in health. Cézard [61] found that the White Scottish population had lower life expectancy than many minoritised ethnic groups. Pakistani and Indian women, and Pakistani men had significantly higher life expectancy but significantly lower disability-free life expectancy (DFLE). Wohland and colleagues [62] used data from the 2001 Census for England and Wales, and found that most minoritised ethnic groups had lower life expectancy and DFLE than the White British group.

Table 10: Mortality rate of 80+ group between 2011-2019 in PHRD data, split by ethnic group and limiting long term illness.

	Minoritised groups	White British
LLTI: Not limited	39.3%	58.5%
LLTI: Limited a little	67.8%	79.7%
LLTI: Limited a lot	85.1%	93.2%

Source: ONS Public Health Research Database

There was a clear relationship between limiting long-term illness (LLTI) and mortality in our data, as evidenced by table 10. We subsequently created DFLE estimates using the Sullivan method as described by Jagger et al. [63], with disability defined as having a long-term limiting illness or having “bad” or “very bad” general health. The results can be seen in table 11, showing lower DFLE for many minoritised ethnic groups. In some cases, the differences in DFLE were large, particularly among women, Bangladeshi women (54.3), Pakistani women (55.3) and White Gypsy/Traveller (49.3) having much lower DFLE than White British women (65.0). It should be noted, however, that the sample size was low for the Gypsy/Traveller group, particularly at older ages.

This analysis formalises the disconnect between mortality and health seen among minoritised ethnic groups in England and Wales. Here we use the same dataset used to create our estimates of life expectancy to demonstrate that poorer health does not translate to lower estimates of life expectancy for minoritised ethnic groups. The counter-intuitive nature of this observation has been noted in other studies. Research using statistical models to estimate life expectancy among ethnic groups has placed minoritised ethnic groups as having lower life expectancies than the white majority. In their attempts to estimate life expectancy in the UK at geographical area level [64], Morris and colleagues did not have access data that mapped ethnicity to death registrations, so used the known relationship between mortality and deprivation, age, and area-level characteristics to model mortality across ethnic groups. This enabled the question: given the deprivation, age and area-level characteristics of an ethnic group, what would we expect the life expectancy of this group to be, when based on the mortality risk associated with those characteristics seen in the wider population? Modelling mortality in this way produced lower life expectancy estimates among people from minoritised ethnic groups. Separately, Rees and colleagues produced ethnicity-specific estimates of mortality based on Census data on long-term limiting illness and also found lower life expectancy for many minoritised ethnic groups [65]. Although the ecological assumptions inherent in such approaches make for a weaker design which should be acknowledged, it demonstrates the counter-intuitive nature of the observed migrant mortality advantage.

Table 11: Disability-free life expectancy for women and men by ethnic group

Women													
Age group	Asian other	Bangladeshi	Black African	Black Caribbean	Black other	Indian	Mixed	Other	Pakistani	White British	White other	White Gypsy / Traveller	White Irish
U1	67.9	54.3	65.0	63.2	63.7	62.9	62.7	61.3	55.3	65.0	68.5	49.3	64.9
01-04	67.2	53.6	64.4	62.7	63.1	62.2	62.0	60.6	54.7	64.2	67.7	48.4	64.1
05-09	63.3	49.7	60.5	58.8	59.2	58.3	58.1	56.7	50.9	60.3	63.8	44.5	60.2
10-14	58.4	45.0	55.7	54.0	54.4	53.4	53.3	52.0	46.1	55.5	58.9	39.8	55.4
15-19	53.6	40.2	50.8	49.2	49.7	48.6	48.5	47.2	41.4	50.7	54.1	35.2	50.8
20-24	48.8	35.5	46.0	44.5	45.0	43.7	43.8	42.5	36.7	46.0	49.3	30.5	46.1
25-29	43.9	30.8	41.3	40.0	40.3	38.9	39.2	37.7	32.1	41.3	44.5	26.4	41.4
30-34	39.2	26.1	36.6	35.4	35.8	34.1	34.6	32.9	27.5	36.7	39.7	22.6	36.7
35-39	34.4	21.6	31.9	30.9	31.2	29.4	30.3	28.5	23.0	32.2	34.9	19.0	32.1
40-44	29.7	17.4	27.4	26.6	26.8	24.7	26.0	24.2	18.8	27.9	30.2	15.7	27.6
45-49	25.2	13.6	23.1	22.5	22.9	20.3	22.0	20.2	14.9	23.7	25.7	12.6	23.3
50-54	20.8	10.7	18.9	18.6	19.2	16.2	18.1	16.6	11.4	19.6	21.5	10.4	19.1
55-59	16.7	8.9	14.9	14.8	15.6	12.5	14.6	13.3	8.9	15.8	17.4	8.2	15.3
60-64	13.0	7.5	11.3	11.2	12.2	9.3	11.4	10.3	6.8	12.3	13.6	6.0	12.1
65-69	9.6	5.8	8.3	8.0	9.3	6.4	8.5	7.6	4.9	9.1	10.1	4.9	9.1
70-74	6.6	4.6	5.7	5.9	6.8	4.4	5.8	5.2	3.4	6.1	6.9	2.9	6.4
75-79	4.5	3.9	3.9	4.0	5.1	3.0	3.6	3.4	2.7	3.7	4.3	1.9	4.0
80-84	2.9	3.1	2.4	2.8	3.3	2.1	2.7	2.2	2.3	2.0	2.5	1.3	2.3
85-89	2.0	2.2	2.0	2.0	1.8	1.5	1.4	2.0	1.5	1.0	1.6	0.8	1.4
90+	1.8	1.0	1.0	1.4	1.7	1.3	1.1	1.6	0.9	0.6	1.3	1.3	0.8

Men													
Age group	Asian other	Bangladeshi	Black African	Black Caribbean	Black other	Indian	Mixed	Other	Pakistani	White British	White other	White Gypsy / Traveller	White Irish
U1	68.9	58.9	68.1	63.9	64.4	65.9	62.8	63.8	60.7	64.2	68.2	46.6	62.8
01-04	68.2	58.2	67.5	63.3	63.8	65.2	62.1	63.1	60.2	63.4	67.4	45.8	62.0
05-09	64.3	54.4	63.7	59.5	60.0	61.3	58.2	59.3	56.4	59.6	63.5	42.1	58.1
10-14	59.5	49.6	59.0	54.9	55.3	56.4	53.5	54.5	51.7	54.8	58.7	37.6	53.5
15-19	54.7	44.9	54.3	50.2	50.7	51.6	48.9	49.8	47.0	50.2	53.9	33.2	48.7
20-24	49.9	40.1	49.5	45.5	46.2	46.9	44.3	45.1	42.4	45.6	49.2	28.8	44.2
25-29	45.1	35.4	44.9	41.0	41.7	42.1	39.7	40.5	37.8	41.0	44.5	24.7	39.5
30-34	40.4	30.7	40.3	36.6	37.2	37.3	35.2	35.9	33.1	36.5	39.7	20.9	34.9
35-39	35.7	26.2	35.7	32.2	32.9	32.6	30.7	31.4	28.6	32.0	35.0	17.6	30.3
40-44	31.1	21.8	31.1	27.9	28.6	28.0	26.5	27.0	24.3	27.7	30.3	14.6	26.0
45-49	26.6	17.5	26.6	23.8	24.6	23.5	22.6	22.9	20.0	23.4	25.8	11.3	21.8
50-54	22.3	13.8	22.2	19.8	20.7	19.3	18.7	19.0	16.1	19.4	21.5	9.1	17.8
55-59	18.3	10.5	18.0	16.0	16.9	15.5	15.0	15.4	12.7	15.5	17.5	6.7	14.1
60-64	14.5	8.1	14.1	12.4	13.5	11.9	11.8	12.0	9.9	12.0	13.7	4.9	10.8
65-69	11.1	5.5	10.6	9.1	10.7	8.8	8.9	9.0	7.3	9.0	10.2	3.8	8.0
70-74	8.2	4.2	7.7	6.8	8.7	6.1	6.9	6.2	5.5	6.2	7.1	2.4	5.6
75-79	5.9	4.0	5.7	4.7	6.4	4.1	4.6	4.3	4.1	4.0	4.6	2.1	3.8
80-84	4.3	3.6	4.0	3.3	4.8	2.6	2.9	3.0	3.0	2.4	2.6	1.2	2.3
85-89	3.4	2.5	2.5	2.2	4.1	1.4	2.1	1.6	2.5	1.3	1.6	0.7	1.4
90+	3.0	1.9	2.5	1.4	5.6	1.0	1.9	1.2	2.7	0.9	1.3	- 0.1	0.9

Source: ONS Public Health Research Database

CONCLUSION

In this study, we set out to improve current estimates of ethnic differences in mortality rates, life expectancy and healthy life expectancy. As part of this, we aimed to provide accurate and up to date estimates of mortality rates and life expectancy (LE) for ethnic groups in England and Wales. We found that the data necessary to provide accurate, definitive estimates of life expectancy is unavailable, with the key data limitation being the inability to reliably measure outmigration from England and Wales resulting in a large proportion of people from minoritised ethnic groups having unknow outcomes and consequently becoming "statistically immortal".

Given this crucial limitation, we instead demonstrated the sensitivity of the existing life expectancy estimates to potential sources of error in the methodology and data used by the ONS. We demonstrated that life expectancy estimates for some minoritised ethnic groups changed by a number of years when we introduced 20 extra deaths into the 90+ age group, whereas the estimates for other ethnic groups (particularly White British) remained stable. We demonstrated that outmigration is likely under-estimated for the non-UK born cohort in the ONS model of life expectancy. We then tested an alternative estimate of outmigration by using loss to follow-up from the ONS Longitudinal Study, and saw a reduction in life expectancy for many minoritised ethnic groups. We then applied higher mortality rates to outmigrants using coefficients based upon research from France that demonstrated higher mortality risk among returnees, observing a change of up to 5 years in life expectancy at birth for some minoritised ethnic groups. Conversely, we saw a change of less than 0.5 years for the White British group.

The loss to follow-up figures from the ONS Longitudinal Study demonstrate that dropout/loss to follow-up (and hence statistical immortality) is likely to be higher among minoritised ethnic groups. We observed that this missingness is particularly high among minoritised ethnic groups at older ages, i.e. in the age groups where missing data can most effect life expectancy at birth estimates. We showed that in the dataset used by the ONS in their estimates, the mortality rate generally declines over time across the study period, and this decline is sharper among minoritised ethnic groups compared with the White British group. This may be due to statistical immortality increasing over time, as the likelihood of having outmigrated (and that outmigration event being unaccounted for) increases over time. As such, this indicates that unrecorded outmigration is likely higher among minoritised ethnic groups.

We also aimed to extend the original ONS estimates of life expectancy. We explored the difference in mortality between people from minoritised ethnic groups who were born in the UK, and those who were born outside the UK. We observed that, using the ONS methodology, the age-specific mortality rates of UK-born people from minoritised ethnic groups is not largely dissimilar to that of the White British group. This suggests that the apparent advantage in life expectancy among minoritised ethnic groups is isolated to migrants from minoritised ethnic groups; however, our attempts to verify this were limited due to the small sample size seen in the UK-born cohorts of most minoritised ethnic groups at the eldest ages. We extended our analysis

to disaggregate ethnic groups, and were able to calculate estimates for the Arab, Chinese, White British, White Gypsy/Traveller, White Irish and White Other groups, ensuring that the aggregation of ethnic groups in the ONS methodology did not mask differences between the constituent groups. We found particularly poor life expectancy and disability-free life expectancy outcomes for the Gypsy/Traveller ethnic group. Finally, we produced estimates of disability-free life expectancy, showing that disability-free life expectancy contradicts the life-expectancy estimates. Given that limiting long-term illness showed strong association with mortality in our data, this calls into question the longer life expectancy seen among minoritised ethnic groups in the original LE estimates

In conclusion, we have demonstrated that the ONS experimental estimates of life expectancy by ethnic group are particularly sensitive to missing data and methodological assumptions around these missing data. We advise that, given the limitations in the data available in England and Wales to study ethnic inequalities in mortality, and the particular sensitivity of life expectancy estimates to shortcomings in this data, that life expectancy estimates should not be used to comment on ethnic inequalities in health. We note that the ONS' life expectancy estimates are being routinely referenced without suitable warnings, and we would advise against researchers, practitioners, or policy makers using these statistics in their publications.

RECOMMENDATIONS FOR FUTURE WORK

Improvement in the accuracy of life expectancy estimates by ethnic group depends on improvement in the quality and availability of suitable data.

Work currently being carried out by the ONS into improving estimates of outmigration has the potential to increase the accuracy of these estimates. The admin-based approach to producing migration statistics using Registration and Population Interaction Database (RAPID) data [52] may provide a more reliable estimate of outmigration, by looking at indicators of activity across a range of government services. This would reduce the amount of error due to statistical immortality. We note that the ONS currently do not produce estimates of outmigration by ethnic group. It is of critical importance that the ONS enable future estimates of outmigration, whether produced using an admin-based approach, to be produced by ethnic group.

Another data-driven approach that would improve estimates would be to link 2011 Census data to 2021 Census data at an individual-level. We would then be able to observe whether someone has been lost to follow-up, which we could then use as a likely flag of outmigration. However, this would not account for the outcomes of people who choose not to participate in the Census. In addition, recording of mortality statistics could be improved include recording ethnic group information on death certificates. This would allow us to know the proportion of unmatched deaths among each ethnic group.

The above only addresses one source of error in quantifying the mortality rate of migrants. To account for increased mortality among returnees, an alternative approach would be required that can track the mortality outcomes of people who have returned to their country of origin. If pensions data were available as in the paper by Guillot et al. [18], we would be able to fully correct for the additional mortality among people who returned to their country of origin at any stage in life. Indeed, Guillot et al. state that it is essential to have visibility of this cohort in order to fully correct for salmon bias, as observing only those people who outmigrate during the study period is not sufficient.

Finally, we note that at present the ONS are engaged in a process on consultation on moving away from a Census to a more admin-based approach to creating statistics on the population of the UK [66]. We implore that accurate statistics detailed by ethnic groups should be a key deliverable for any such re-orientation of the ONS' methods.

ADDITIONAL MATERIALS

APPENDIX A: DEATH REGISTRATIONS IN PHRD 2011-2022

	Source: Unweighted sum of deaths from PHRD											
	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
Jan		40,868	44,739	40,724	53,814	44,405	51,544	54,322	45,881	46,469	69,983	44,506
Feb		40,320	39,426	36,887	41,918	41,414	41,878	44,097	40,654	40,412	47,306	37,631
Mar	4,322	39,730	44,465	38,476	42,788	44,299	40,949	46,332	38,778	49,287	39,527	41,028
Apr	32,170	37,780	40,934	36,314	38,847	40,086	37,377	39,020	39,518	74,430	35,845	40,542
May	34,306	37,372	37,120	35,861	38,246	37,835	38,778	37,449	38,118	46,972	36,989	37,561
Jun	32,944	34,131	33,584	34,479	35,543	35,295	35,484	35,284	35,413	36,541	35,179	35,614
Jul	33,457	34,806	33,834	35,530	34,696	37,224	35,695	36,398	36,339	35,372	39,481	38,534
Aug	33,060	34,348	33,377	35,070	35,811	36,331	36,553	35,082	36,075	36,472	37,786	36,248
Sep	33,078	34,024	33,762	34,869	35,621	34,350	36,498	35,466	36,209	36,521	38,467	34,679
Oct	35,223	37,415	36,390	38,106	38,912	39,446	39,666	38,259	40,365	42,917	42,210	38,321
Nov	35,482	37,042	36,549	38,365	38,011	40,699	40,345	38,879	41,856	47,083	42,290	3,964
Dec	41,438	43,785	40,626	48,434	41,613	46,063	49,140	42,734	47,434	53,235	46,190	

Source: ONS Public Health Research Database

APPENDIX B: LOSS TO FOLLOW-UP BETWEEN 2001 AND 2011 IN ONS-LS, BY ETHNIC GROUP AND AGE GROUP

Ethnic group	Age group											
	0-18	19-29	30-39	40-49	50-59	50+	60-69	60-74	70+	70-79	75+	80+
Asian Other	18%	33%	22%	19%	16%	-	19%	-	20%	-	-	-
Bangladeshi	18%	18%	12%	16%	14%	-	30%	-	22%	-	-	-
Black African	27%	38%	27%	26%	18%	-	-	24%	-	-	34%	-
Black Caribbean	21%	26%	20%	18%	19%	-	-	19%	-	-	13%	-
Black Other	26%	30%	21%	27%		21%	-	-	-	-	-	-
Chinese	30%	49%	27%	18%	22%	-	-	23%	-	-	17%	-
Indian	12%	18%	14%	11%	13%	-	15%	-	-	21%	-	15%
Mixed	20%	29%	24%	17%	14%	-	-	18%	-	-	15%	-
Other	37%	47%	43%	28%	25%	-	-	26%	-	-	32%	-
Pakistani	19%	23%	16%	16%	16%	-	19%	-	-	23%	-	28%
White British	14%	14%	11%	8%	7%	-	7%	-	-	6%	-	3%
White Irish	31%	43%	25%	18%	17%	-	13%	-	-	11%	-	8%
White Other	36%	58%	40%	27%	21%	-	20%	-	-	13%	-	8%

Source: ONS Longitudinal Study

REFERENCES

1 Office for National Statistics. Ethnic differences in life expectancy and mortality from selected causes in England and Wales: 2011 to 2014. 2021. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/articles/ethnicdifferencesinlifeexpectancyandmortalityfromselectedcausesinenglandandwales/2011to2014

2 Bécares L. Which ethnic groups have the poorest health? In: Jivraj A, Simpson L, eds. *Ethnic Identity and Inequalities in Britain: The Dynamics of Diversity*. Bristol, UK: : Policy Press 2015. 123–40. doi:10.2307/j.ctt1t89504.14

3 Hill L, Ndugga N, Artiga S. Key data on health and health care by race and ethnicity. KFF. 2023.

4 Blom N, Huijts T, Kraaykamp G. Ethnic health inequalities in Europe. The moderating and amplifying role of healthcare system characteristics. *Soc Sci Med* 2016;**158**. doi:10.1016/j.socscimed.2016.04.014

5 Aslam S, Patsalides MA, Stoma S, *et al.* Aortic valve intervention rates in patients of different ethnicity with severe aortic stenosis in Leicestershire, UK. *Open Hear* 2023;**10**. doi:10.1136/openhrt-2023-002266

6 Ling S, Xu G, Zaccardi F, *et al.* Kidney function and long-term risk of end stage kidney disease and mortality in a multi-ethnic population. *Kidney Int Reports* Published Online First: 2023. doi:10.1016/j.ekir.2023.06.014

7 Tjepkema M, Christidis T, Olaniyan T, *et al.* Mortality inequalities of Black adults in Canada. *Heal Reports* 2023;**34**. doi:10.25318/82-003-x202300200001-eng

8 Public Health England. Health Profile for England 2021. 2021. https://fingertips.phe.org.uk/static-reports/health-profile-for-england/hpfe_report.html

9 Grover N. Life expectancy lower for white and mixed ethnic people than Black and Asian groups – study. Guard. 2021.https://www.theguardian.com/society/2021/jul/26/life-expectancy-lower-for-white-and-mixed-ethnic-people-than-black-and-asian-groups-study

10 Anson J. The migrant mortality advantage: A 70 month follow-up of the Brussels population. Eur. J. Popul. 2004;**20**. doi:10.1007/s10680-004-0883-1

11 Boulogne R, Jougla E, Breem Y, *et al.* Mortality differences between the foreign-born and locally-born population in France (2004-2007). *Soc Sci Med* 2012;**74**. doi:10.1016/j.socscimed.2012.01.002

12 Scott AP, Timæus IM. Mortality differentials 1991-2005 by self-reported ethnicity: Findings from the ONS Longitudinal Study. *J Epidemiol Community Health* 2013;**67**. doi:10.1136/jech-2012-202265

13 Di Napoli A, Rossi A, Alicandro G, *et al.* Salmon bias effect as hypothesis of the lower mortality rates among immigrants in Italy. *Sci Rep* 2021;**11**. doi:10.1038/s41598-021-87522-2

14 Bhopal RS, Gruer L, Cezard G, *et al.* Mortality, ethnicity, and country of birth on a national scale, 2001–2013: A retrospective cohort (Scottish Health and Ethnicity Linkage Study).

PLoS Med 2018;**15**. doi:10.1371/journal.pmed.1002515

15 Helgesson M, Johansson B, Nordquist T, *et al*. Healthy migrant effect in the Swedish context: A register-based, longitudinal cohort study. *BMJ Open* 2019;**9**. doi:10.1136/bmjopen-2018-026972

16 Tarnutzer S, Bopp M. Healthy migrants but unhealthy offspring? A retrospective cohort study among Italians in Switzerland. *BMC Public Health* 2012;**12**. doi:10.1186/1471-2458-12-1104

17 Roelfs D, Shor E. The problematic nature of existing explanations for differential immigrant mortality: Insights from a comparative cross-national systematic review and meta-analysis. *Soc Sci Med* 2024;**349**. doi:10.1016/j.socscimed.2024.116897

18 Guillot M, Khlat M, Gansey R, *et al*. Return Migration Selection and Its Impact on the Migrant Mortality Advantage: New Evidence Using French Pension Data. *Demography* Published Online First: 2023. doi:10.1215/00703370-10938784

19 Pablos-Mendez A. Letter to the Editor. *JAMA* 1994;**271**:1237–8.

20 Abraído-Lanza AF, Dohrenwend BP, Ng-Mak DS, *et al*. The Latino mortality paradox: A test of the 'salmon bias' and healthy migrant hypotheses. *Am J Public Health* 1999;**89**. doi:10.2105/AJPH.89.10.1543

21 Namer Y, Razum O. Convergence Theory and the Salmon Effect in Migrant Health. In: *Oxford Research Encyclopedia of Global Public Health*. 2018. doi:10.1093/acrefore/9780190632366.013.17

22 Wallace M, Kulu H. Can the salmon bias effect explain the migrant mortality advantage in England and Wales? *Popul Space Place* 2018;**24**. doi:10.1002/psp.2146

23 Dunlavy A, Cederström A, Katikireddi SV, *et al*. Investigating the salmon bias effect among international immigrants in Sweden: a register-based open cohort study. *Eur J Public Health* 2022;**32**. doi:10.1093/eurpub/ckab222

24 Turra CM, Elo IT. The impact of salmon bias on the hispanic mortality advantage: New evidence from social security data. *Popul Res Policy Rev* 2008;**27**. doi:10.1007/s11113-008-9087-4

25 Office for National Statistics. Mortality from leading causes of death by ethnic group, England and Wales: 2012 to 2019. 2021.https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/articles/mortalityfromleadingcausesofdeathbyethnicgroupenglandandwales/latest

26 UK Government. Inclusive Britain: government response to the Commission on Race and Ethnic Disparities. London: 2022.

27 Schwandt H, Currie J, Bär M, *et al*. Inequality in mortality between Black and White Americans by age, place, and cause and in comparison to Europe, 1990 to 2018. Proc. Natl. Acad. Sci. U. S. A. 2021;**118**. doi:10.1073/pnas.2104684118

28 Office for National Statistics. People born outside the UK. 2018.https://www.ethnicity-facts-figures.service.gov.uk/uk-population-by-ethnicity/demographics/people-born-outside-the-uk/latest/

29 Aldridge RW, Nellums LB, Bartlett S, *et al*. Global patterns of mortality in international migrants: a systematic review and meta-analysis. *Lancet* 2018;**392**. doi:10.1016/S0140-6736(18)32781-8

30 Wallace M. Mortality among migrants and their descendants living in England and Wales. In: *Migration, Health and Survival: International Perspectives*. 2017. doi:10.4337/9781785365973.00015

31 Maskileyson D, Semyonov M, Davidov E. In search of the healthy immigrant effect in four west european countries. *Soc Incl* 2019;**7**. doi:10.17645/si.v7i4.2330

32 HM Government. Quarterly report on progress to address COVID-19 health inequalities. 2020. https://assets.publishing.service.gov.uk/media/5fce138b8fa8f54d5abcdf24/First_Covid_Disparities_report_to_PM___Health_Secretary_Final_22-10-20_-_Updated_December_2020.pdf

33 Wallace M, Darlington-Pollock F. Poor health, low mortality? Paradox found among immigrants in England and Wales. In: *Population, Space and Place*. 2022. doi:10.1002/psp.2360

34 Nazroo J, Bécares L, Kapadia D. Ethnic inequalities in mortality rates and life expectancy in England and Wales: Why we should treat experimental statistics with caution. NHS Race Heal. Obs. 2021.https://www.nhsrho.org/blogs/ethnic-inequalities-in-mortality-rates-and-life-expectancy-in-england-and-wales-why-we-should-treat-experimental-statistics-with-caution/

35 United Nations. World Population Prospects 2022. 2022. https://population.un.org/wpp/

36 Stopforth S, Kapadia D, Nazroo J, *et al*. Ethnic inequalities in health in later life, 1993-2017: The persistence of health disadvantage over more than two decades. *Ageing Soc* Published Online First: 2021. doi:10.1017/S0144686X2100146X

37 Bécares L. Which ethnic groups have the poorest health? In: *Ethnic Identity and Inequalities in Britain: The Dynamics of Diversity*. 2015. 123–40. doi:10.2307/j.ctt1t89504.14

38 Nazroo J. The Health of Britain's Ethnic Minorities. London: 1997.

39 Office for National Statistics. Guide to experimental statistics. https://www.ons.gov.uk/methodology/methodologytopicsandstatisticalconcepts/guidetoexperimentalstatistics

40 Phillips A. White and mixed ethnic groups 'don't live as long' as other groups, study suggests. Sky News. 2021.https://news.sky.com/story/white-and-mixed-ethnic-groups-dont-live-as-long-as-other-groups-study-suggests-12364473

41 Borrett A. Why are white Britons dying at higher rates than other ethnic groups? Financ. Times. 2023.https://www.ft.com/content/f51ee83d-8a9b-4eba-8a04-5609c70a74fa

42 Iacobucci G. Life expectancy: White people in England and Wales had worst outcomes before pandemic, data show. *BMJ* 2021;**374**. doi:10.1136/bmj.n1886

43 Mirza HS, Warwick R. Race and ethnicity. London: 2022. https://ifs.org.uk/publications/race-and-ethnicity

44 Office for National Statistics. Updating ethnic and religious contrasts in deaths involving the coronavirus (COVID-19), England: 24 January 2020 to 23 November 2022. 2023.https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/articles/updatingethniccontrastsindeathsinvolvingthecoronaviruscovid19englandandwales/24january2020to23november2022

45 Office for National Statistics. Personal communication - based on Census 2011, Death Registrations, Hospital Episode Statistics and General Practice Extraction Service Linked dataset, England and Wales (unpublished). 2023.

46 Office for National Statistics. ONS SRS Metadata Catalogue, dataset, ONS Longitudinal Study - England and Wales. 2019. doi:10.57906/z9xn-ng05

47 Office for National Statistics. Response Rates in the 2011 Census. 2012. https://webarchive.nationalarchives.gov.uk/ukgwa/20160108085252/http:/www.ons.gov.uk/ons/guide-method/census/2011/census-data/2011-census-user-guide/quality-and-methods/response-and-imputation-rates/index.html

48 Office for National Statistics, NHS Digital. ONS SRS Metadata Catalogue, dataset, Public Health Research Database - England and Wales. 2023. doi:10.57906/b8zx-6x19

49 Taylor H, Kapadia D, Bécares L, *et al*. Health and wellbeing. In: *Racism and Ethnic Inequality in a Time of Crisis: Findings from the Evidence for Equality National Survey*. Policy Press 2023. 78–95. doi:10.2307/jj.3279115.10

50 Office for National Statistics. International Passenger Survey: quality information in relation to migration flows. 2018.https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/internationalmigration/methodologies/internationalpassengersurveyqualityinformationinrelationtomigrationflows

51 Office for National Statistics. Migration Statistics Quarterly Report: November 2014. 2014.https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/internationalmigration/bulletins/migrationstatisticsquarterlyreport/2015-06-30#emigration-from-the-uk

52 Office for National Statistics. International migration: developing our approach for producing admin-based migration estimates. 2021.https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/internationalmigration/articles/internationalmigrationdevelopingourapproachforproducingadminbasedmigrationestimates/2021-04-16

53 Hattersley L. LS User Guide 18—International Migration Data in the Longitudinal Study. London: 1999.

54 Wallace M. Adult mortality among the descendants of immigrants in England and Wales: does a migrant mortality advantage persist beyond the first generation? *J Ethn Migr Stud* 2016;**42**. doi:10.1080/1369183X.2015.1131973

55 Office for National Statistics. Trend in life expectancy at birth and at age 65 by socio-economic position based on the National Statistics Socio-economic Classification, England and Wales: 1982—1986 to 2007—2011. 2015. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/bulletins/trendinlifeexpectancyatbirthandatage65bysocioeconomicposit

ionbasedonthenationalstatisticssocioeconomicclassificationenglandandwales/2015-10-21#detailed

56 White N. International Passenger Survey: quality information in relation to migration flows. Off. Natl. Stat. 2018.https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/internationalmigration/methodologies/internationalpassengersurveyqualityinformationinrelationtomigrationflows

57 Office for National Statistics. Long-term international migration, provisional: year ending June 2023. Published Online First: 2023.https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/internationalmigration/bulletins/longterminternationalmigrationprovisional/yearendingjune2023#emigration-out-of-the-uk

58 Blackwell L, Lynch K, Smith J, *et al.* Longitudinal study 1971–2001: completeness of census linkage. London: 2003.

59 Hopper N. Predicting patterns of household non-response in the 2011 Census. In: *Survey Methodology Bulletin No. 69 - September 2011, ONS*. 2011. 9–22.

60 Feliciano C. Immigrant selectivity effects on health, labor market, and educational outcomes. In: *Annual Review of Sociology*. 2020. doi:10.1146/annurev-soc-121919-054639

61 Cezard GI. Evidencing the gap between health expectancy and life expectancy for ethnic groups in Scotland. *Quetelet J* 2020;**7**. doi:10.14428/rqj2019.07.01.05

62 Wohland P, Rees P, Nazroo J, *et al.* Inequalities in healthy life expectancy between ethnic groups in England and Wales in 2001. *Ethn Heal* 2015;**20**:341–53. doi:10.1080/13557858.2014.921892

63 Jagger C, Oyen H Van, Robine J-M. Health Expectancy Calculation by the Sullivan Method: A Practical Guide. *Newcastle Univ Isntitute Ageing* 2014.

64 Morris M, Woods LM, Rachet B. A novel ecological methodology for constructing ethnic-majority life tables in the absence of individual ethnicity information. *J Epidemiol Community Health* 2015;**69**. doi:10.1136/jech-2014-204210

65 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**. doi:10.1016/j.socscimed.2009.08.015

66 Office for National Statistics. The future of population and migration statistics in England and Wales. 2023.https://consultations.ons.gov.uk/ons/futureofpopulationandmigrationstatistics/supporting_documents/Future of Population and Migration Statistics Consultation Document.pdf