Document finalized on April 24, 2020
Submitted for publication on May 1, 2020

# Ethnicity and Outcomes of COVID-19 Patients in England

Ahmed M. Alaa[1], Zhaozhi Qian[2], Jem Rashbass[3], Keith Gomes Pinto[3], Jonathan Benger[3], and Mihaela van der Schaar[2]

[1]University of California, Los Angeles, CA 90095, USA
[2]University of Cambridge Centre for Mathematical Sciences, Cambridge CB3 0WA, UK
[3]NHS Digital, Trevelyan Square, Leeds, LS1 6AE, UK

## Abstract

Background
The rate of COVID-19 infections, hospitalizations and deaths are hypothesized to be disproportionately high among members of the Black, Asian, Mixed-race and Ethnic minorities (BAME) community in England. We analyzed observational data to determine whether people from BAME backgrounds are more susceptible to severe disease.

Methods
We linked five routine data sets to follow a cohort of patients who had been diagnosed with COVID-19 to examine hospital admission, intensive care unit admission and death by self-reported ethnicity. We compared the outcomes of four broad ethnic groups using descriptive statistics and a multivariate mixed-effects Cox regression model adjusting for the effect of age, deprivation and comorbidities on outcome.

Findings
Of 78,443 patients diagnosed with COVID-19, 72,358 (92%) had information on ethnicity. Compared to the overall population, individuals from a BAME background were more likely to be diagnosed with COVID-19, more likely to be admitted to hospital and intensive care, and more likely to die. Importantly, the median age of BAME patients was significantly lower than white patients (Asian 51 years; Mixed/Other 57 years; Black 57 years; White 69 years). Despite these age differences the total burden of comorbidities was similar across all ethnicities, though respiratory diseases were more prevalent in White patients, and cardiovascular and endocrine disease were more prevalent in BAME patients. The Cox regression model demonstrated a range of risk factors, with age dominant. BAME background emerged as an independent risk factor, particularly Asian ethnicity.

Interpretation
Increased prevalence of COVID-19 amongst individuals from a BAME background may be explained by the geographical distribution of COVID in England, deprivation and occupational exposure. Our findings suggest that BAME patients, and particularly those with an Asian background, are at an elevated risk of mortality. Whilst awaiting further research we recommend that ethnicity be considered, alongside other factors, when assessing an individual's COVID-19 risk.

Funding
None

## Introduction

The COVID-19 pandemic has had a detrimental impact on global healthcare systems and economies. On 31st December 2019 the first official report to the World Health Organization (WHO) office in China described 44 cases of pneumonia of unknown etiology detected in Wuhan City, China.(1) As of 30th April 2020, more than 3.2 million cases of COVID-19 have been reported across 186 countries and territories, resulting in 228,000 deaths.(2)

According to data from the Centers for Disease Control and Prevention, COVID-19 has a disproportionate impact on individuals of Black ethnicity. A recent report found an overrepresentation of Blacks among hospitalized patients (33%) in the United States compared to the general population (18%).(3) The death rate among Blacks/African Americans was also substantially higher (92.3 deaths per 100,000), when compared to Hispanic/Latino (74.3), White (45.2) and Asian (34.5) ethnicities.(4) However, it is difficult to extrapolate this to other countries given differences in populations, social factors and healthcare systems.

In the United Kingdom (UK), a report from the Intensive Care National Audit & Research Centre (ICNARC) further highlighted this potential issue. Data collected from patients on Intensive Care Units (ICUs) across England, Northern Ireland and Wales showed that an increased proportion of confirmed COVID-19 patients admitted to critical care were from Black, Asian, Mixed-race and Ethnic minorities (BAME) backgrounds, compared to the general population of the UK. An updated report, on April 23rd, 2020, continued to highlight that Asians (15.4%) and Blacks (10.7%) form a higher percentage of those admitted in comparison to their representation in the UK population. (5) According to data published by the Office for National Statistics (ONS), Asians (8.4%), followed by Blacks (3.8%)

formed the largest ethnic minority groups in England in 2018.(6)

The potential that COVID-19 might tend to cause a more severe illness in those with a BAME background has understandably caused significant concern, especially among health and care workers, a substantial proportion of whom are from ethnic minority backgrounds. (7) This was compounded by the fact that 10 of the first 11 doctors in the UK to die from COVID-19 were identified as being from ethnic minorities. (8)

This report describes an analysis of early data from the outbreak of COVID-19 in the UK, and sets it in the context of other work that seeks to understand the effect of ethnicity on COVID-19 disease severity and outcomes. This has important implications for health and care staff, and wider public health across the UK and globally. (9)

Our objective was to analyze the clinical outcomes of patients diagnosed with COVID-19 in England according to their self-declared ethnicity to determine whether individuals from a Black, Asian, Mixed-race and Ethnic minorities (BAME) background are particularly vulnerable to COVID-19, compared to White ethnicities, and identify potential factors that could explain any increased susceptibility to the disease.

## Methods

In our analysis, we linked five different data sets to construct a comprehensive record of the disease outcomes of a cohort of patients diagnosed with COVID-19 in England. These data sets are described below.

**SGSS data.** COVID-19 testing data was obtained from the Second Generation Surveillance System (SGSS), covering the period to 16th April 2020. The data set contains 78,443 people with positive COVID-19 laboratory tests.

**CHESS data.** Data for COVID19 hospitalizations was obtained from the COVID-19 Hospitalization in England Surveillance System (CHESS) established by Public Health England (PHE). This data covered the period from February 8th to April 14th 2020 and was submitted from 94 of 152 (62%) of acute hospital trusts in England (7,714 hospital admissions, including 3,092 Intensive Care Unit (ICU) admissions).

**PDS data.** We utilized the Personal Demographics Service (PDS) data from NHS Digital, which recorded 15,090 deaths due to COVID-19 up to April 20th, 2020.

**HES data.** We used data from Hospital Episode Statistics (HES) in order to obtain information regarding each patient's self-declared ethnicity. If a patient had multiple HES records with conflicting ethnicity, we chose the one with highest frequency. We divided the population of patients diagnosed with COVID-19 in England into four broad ethnic groups: White, Asian, Black and Other/Mixed ethnic background.

**Primary care data.** Finally, we used primary care prescription medicine data recorded between July 2019 and January 2020 to extract information on each patient's likely preexisting medical conditions. Comorbidities were inferred from the British National Formulary (BNF) chapter code for all medications prescribed to each patient.

Using the data sets above, we constructed one comprehensive data set for a cohort of 72,358 COVID-19 patients, including information on whether each patient was hospitalized, whether they were admitted to the ICU, their comorbidities, in addition to basic demographic information (age, sex, geographical region and postcode), and their outcome on April 20th, 2020. The data extraction process is shown in Figure 1.

Descriptive statistics for age and the ethnic groups of interest were collected from the data sets above. In addition, a multivariate mixed-effects Cox regression model was fit to adjust for the effect of age, deprivation (derived from postcode) and comorbidities on outcome, with patient-level clinical predictors as fixed effects, and the patient's geographical location as a random effect.

## Results

Of 78,443 patients with a positive COVID-19 test in the SGSS testing data, 72,358 (92%) had information on ethnicity and formed the population of interest (Figure 1). By comparing the representation of patients from different ethnic groups at the different stages of COVID-19 disease (Figure 2), we found that BAME patients constitute a disproportionately large fraction of hospitalizations and ICU admissions, especially in younger age groups. Conversely, in the group of patients aged 70 years and older, most of the patients were from the White population. Table 1 provides further summary statistics for each ethnic group. The proportion of patients who died in the Black and Other ethnic groups exceeded the proportion of these groups in the general population as estimated by the ONS in 2018,[6] whilst the proportion of deaths in the Asian and White ethnic groups was lower.

**Age disparities among ethnic groups.** Figure 3 provides a detailed breakdown of the age distribution of COVID-19 patients within each ethnic group. A significant difference in the age distributions of the White and BAME populations is apparent (Panel A). We found that the median age of diagnosed patients in the Asian ethnic group was 51 years, whilst that in the Black and Other groups was 57 years, compared to a median age of 69 years in the White patient population. In Panels C and B of Figure 3, we compare the representation of each ethnic group within specific age groups and compare these with the ethnic composition of the baseline demographic based on the ONS data for 2018. Compared to the baseline demographic, we observed an

excess in the proportion of patients from BAME backgrounds, in keeping with Table 1, and particularly prominent in the younger Asian population (Figure 3, Panel B). On the other hand, the elderly population (more than 70 years old) was dominated by the white population.

**Adjusting for baseline demographics.** While BAME patients appear to be over-represented in those diagnosed with COVID-19 in relation to their representation in the overall demographics of England (Figure 3, panels B and C), it is important to note that the COVID-19 population is more concentrated in the more ethnically diverse regions of England. That is, while the population of greater London comprises only around 14% of the overall population of England, more than 24% of COVID-19 cases in our dataset were associated with the greater London area. By comparing the rates of diagnosis in each ethnic group with their corresponding regional demographics (Figure 4), we found that the diagnosis of COVID-19 generally reflects the underlying ethnic distribution in each region, however White populations are less likely to be diagnosed with COVID-19, compared to their regional representation, while Black and Mixed populations are more likely to be diagnosed.

**Prevalence of comorbidities.** The higher prevalence of COVID-19 in younger BAME patients could potentially be explained by differences in comorbidities. Tables 2 and 3 show the overall prevalence and number of comorbidities by ethnic group, with additional detail in Figure 6. BAME patients displayed a higher prevalence of cardiovascular disease, endocrine disease and hypertension compared to age-matched white patients, whilst White patients were more likely to suffer from respiratory disease.

**Ethnicity and risk elevation.** Figure 7 shows the hazard ratios and 95% confidence intervals for all the clinical predictors under consideration in the fitted multi-variate Cox model for mortality. An Asian ethnic background significantly elevates mortality risk even when adjusting for age, deprivation and comorbidities. This is explored further in Figure 8 in which this analysis is repeated, separating the Asian ethnic group into Indian, Pakistani and Bangladeshi in comparison to other Asian ethnicity, and demonstrating that the Indian, Pakistani and Bangladeshi ethnic group has a higher mortality risk than other Asian individuals. To further illustrate the elevated risk within the Asian subgroup, Figure 9 shows the Kaplan-Meier estimates for the probability of death plotted against the number of days since diagnosis stratified by various age groups. We plot these curves for both patients in the overall population and patients who have CHESS records (hospitalized). Asian patients display an elevated risk in older age groups (more than 70 years old) for both the hospitalized and non-hospitalized subgroups, and for the hospitalized age group between 60 and 70 years old.

## Discussion

There are clear differences in the BAME and White populations diagnosed with COVID-19 in England. In keeping with other studies,(3,4) our findings confirm that BAME individuals are over-represented in COVID-19 deaths, compared to the general population. However, they are also over-represented in positive tests, hospitalizations and particularly ICU admissions.

The reason that BAME individuals are over-represented in positive tests may relate to geographical variation in the prevalence of COVID disease. Major outbreaks have occurred in London and the West Midlands, where the population is much more ethnically diverse than the country as a whole (Figure 4), and crude comparisons to the whole population of England are therefore flawed. Other factors that may be important are deprivation, since BAME populations tend to be more deprived (Figure 5), and occupation. “Front line” occupations such as those in health and social care and transport are more likely to be exposed to COVID-19, and are also more ethnically diverse than the general population.[7]

The most striking difference between BAME and White patients with COVID-19 is the younger age distribution in BAME patients. This age differential can be explained by: (i) the difference in the underlying age distribution within each ethnic group in the population, (ii) differences in social deprivation which tends to be higher in BAME populations (Figure 5), and (iii) differences in the age of onset of chronic diseases and co-morbidities that themselves influence outcome (Figure 6).

Differences in the age profile between populations also explains the higher rates of ICU admission in BAME patients reported previously.[5] Individuals with the best chance of benefit are selected for ICU care, and this tends to favour younger patient populations. Conversely, older patients with COVID-19 may not be considered appropriate for hospital admission, with a preference for care to be delivered at or closer to home.

Significant concerns have been raised regarding a disproportionate death rate amongst BAME health and social care staff in England,[10] and this may also be influenced by the observed differences in age profile. The median age for White patients in our analysis is older than the usual retirement age in England, whilst the median age for BAME individuals is well within the working population.

Comorbidities are a well-recognized risk factor for severe COVID-19 disease.[11] In keeping with other studies we found age to be the dominant risk factor, but also identified endocrine disease, cardiovascular disease, respiratory disease and hypertension to be independent risk factors. In this study “endocrine disease” is almost exclusively diabetes, and it is important to note that this and cardiovascular disease were both the most important comorbidities and most prevalent amongst the BAME population diagnosed with COVID-19.

The younger demographic, prevalence of comorbidities at an earlier age, higher deprivation and the fact that ethnically diverse regions in England exhibited higher infection rates, may explain much of the disproportionate rates of diagnosis, hospitalization and ICU admissions among younger BAME patients. However, an Asian ethnic background, particularly Indian, Pakistani or Bangladeshi, was still found to be a significant independent risk factor for mortality even after adjusting for age, deprivation and co-morbidities. Of note, endocrine diseases (i.e. diabetes) and cardiovascular disease are also strong independent predictors of mortality and more prevalent in the Asian ethnic group. It is therefore possible that some of the excess mortality associated with a BAME background is attributable to diabetes and cardiovascular disease that is present but not diagnosed, or present but not detected in this analysis. Alternatively, the elevated risk within the Asian ethnic group, and to some extent other BAME individuals, may be attributable to other confounding clinical, social or genetic factors that we have not been able to measure in this study.

This research utilizes a large data set with complete data linkage, and our findings are both consistent with other studies and current observations in the COVID-19 pandemic. Study weakness are: exclusion of 8% of patients for whom no ethnicity was recorded; the use of self-reported ethnicity, and broad category groupings that may not fully reflect an individual's underlying genetic profile; incomplete data in the CHESS collection; the use of postcode and primary care prescription data to infer deprivation and comorbidities, respectively. However, the most significant weakness is the use of observational and routinely collected data that makes it possible for our findings to be influenced by unmeasured confounders. We were unable to account for smoking or obesity in our Cox regression model, and there may be other factors that will need to be considered as additional evidence emerges. Nevertheless, whilst awaiting further research we recommend that ethnicity, and particularly Asian ethnicity, be considered in any assessment of a person's individual risk of hospitalization and death following infection with COVID-19.

## Conclusion

Increased prevalence of COVID-19 amongst individuals with a BAME background is at least partially explained by the geographical distribution of COVID in England, deprivation and occupational exposure. BAME patients with COVID-19 have a younger median age than White patients, and a higher burden of comorbidity, particularly cardiovascular and endocrine diseases. Our findings suggest that BAME patients, and particularly those with an Asian background, are at an elevated risk of mortality, however we cannot exclude the possibility that our observational data have been influenced by unmeasured confounders. Whilst awaiting further research, we suggest that ethnicity be considered as an independent risk factor when assessing an individual's COVID-19 risk.

## References


1. World Health Organization. Novel Coronavirus (2019-nCoV) SITUATION REPORT - 1 21 JANUARY 2020. [Online] 04 30, 2020. [Cited: 04 30, 2020.] https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200121-sitrep-1-2019-ncov.pdf.

2. Johns Hopkins University. COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. [Online] 04 30, 2020. [Cited: 04 30, 2020.] https://gisanddata.maps.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6.

3. Centers for Disease Control and Prevention. COVID-19 in Racial and Ethnic Minority Groups. Coronavirus Disease 2019 (COVID -19). [Online] 04 22, 2020. [Cited: 04 30, 2020.] https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/racial-ethnic-minorities.html.

4. Centers for Disease Control and Prevention. Hospitalization Rates and Characteristics of Patients Hospitalized with Laboratory-Confirmed Coronavirus Disease 2019 — COVID-NET, 14 States, March 1–30, 2020. Morbidity and Mortality Weekly Report (MMWR). [Online] 04 17, 2020. [Cited: 04 30, 2020.] https://www.cdc.gov/mmwr/volumes/69/wr/mm6915e3.htm?s_cid=mm6915e3_w.

5. Intensive Care National Audit & Research Centre. ICNARC report on COVID-19 in critical care 24 April 2020. [Online] 04 24, 2020. [Cited: 04 30, 2020.] https://www.icnarc.org/Our-Audit/Audits/Cmp/Reports.

6. Office for National Statistics. Population denominators by ethnic group, regions and countries: England and Wales, 2011 to 2018. [Online] 07 27, 2018. [Cited: 04 30, 2020.] https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/adhocs/008780populationdenominatorsbyethnicgroupregionsandcountriesenglandandwales2011to2017.

7. NHS workforce. Ethnicity Facts and Figures. [Online] 01 06, 2020. [Cited: 04 30, 2020.] https://www.ethnicity-facts-figures.service.gov.uk/workforce-and-business/workforce-diversity/nhs-workforce/latest.

8. British Medical Association. BMA council chair's message to members. COVID-19. [Online] 04 10, 2020. [Cited: 04 30, 2020.] https://www.bma.org.uk/advice-and-support/covid-19/what-the-bma-is-doing/bma-council-chair-message.

9. Pareek M, Bangash MN, Pareek N, Pan D, Sze S, Minhas JS, Hanif W, Khunti K. Ethnicity and COVID-19: an urgent public health research priority. [Online] 04 21, 2020. [Cited:

04 30, 2020]. https://doi.org/10.1016/S0140-6736(20)30922-3.

10. NHS Confederation. The impact of COVID-19 on BME communities and health and care staff. [Online] 04 23, 2020. [Cited: 04 30, 2020.] https://www.nhsconfed.org/resources/2020/04/the-impact-of-covid19-on-bme-communities-and-staff.

11. Richardson S, Hirsch JS, Narasimhan M, et al. Presenting Characteristics, Comorbidities, and Outcomes Among 5700 Patients Hospitalized With COVID-19 in the New York City Area. [Online] 04 22, 2020. [Cited 04 30, 2020]. https://jamanetwork.com/journals/jama/fullarticle/2765184.

| | Population (2018) | Diagnosis | Hospitalization | ICU Admission | Deaths |
|---|---|---|---|---|---|
| White | 47,010,735 (84%) | 53,820 (74.6%) | 5,149 (79.6%) | 1,629 (68.7%) | 11,286 (79.4.6%) |
| Asian | 4,686,280 (8.4%) | 7,100 (9.8%) | 581 (9.0%) | 346 (14.6%) | 989 (6.9%) |
| Black | 2,104,834 (3.8%) | 4,778 (6.6%) | 323 (5.0%) | 155 (6.5%) | 811 (5.7%) |
| Other | 2,175,473 (3.8%) | 6,515 (9.0%) | 414 (6.4%) | 239 (10.0%) | 1,127 (7.9%) |

Table 1: Ethnic composition of patients at each stage of the COVID-19 trajectory.

| | Respiratory | Cardiovascular | Endocrine | Malignant & Immunosup. |
|---|---|---|---|---|
| White | 38.7% | 63.3% | 45.9% | 13.4% |
| Asian | 24.4% | 49.6% | 41.7% | 6.5% |
| Black | 27.6% | 61.8% | 48.2% | 11.9% |
| Other | 33.6% | 56.8% | 46.6% | 11.4% |

Table 2: Prevalence of comorbidities in different ethnic groups.

| N Comorb. / Ethnicity | 0 | 1 | 2 | >=3 |
|---|---|---|---|---|
| White | 26.15% | 20.96% | 23.21% | 29.6% |
| Asian | 38.67% | 20.12% | 23.34% | 17.86% |
| Black | 27.78% | 21.84% | 27.72% | 22.66% |
| Other | 32.16% | 17.92% | 23.64% | 26.28% |

Table 3: Number of comorbidities in different ethnic groups.

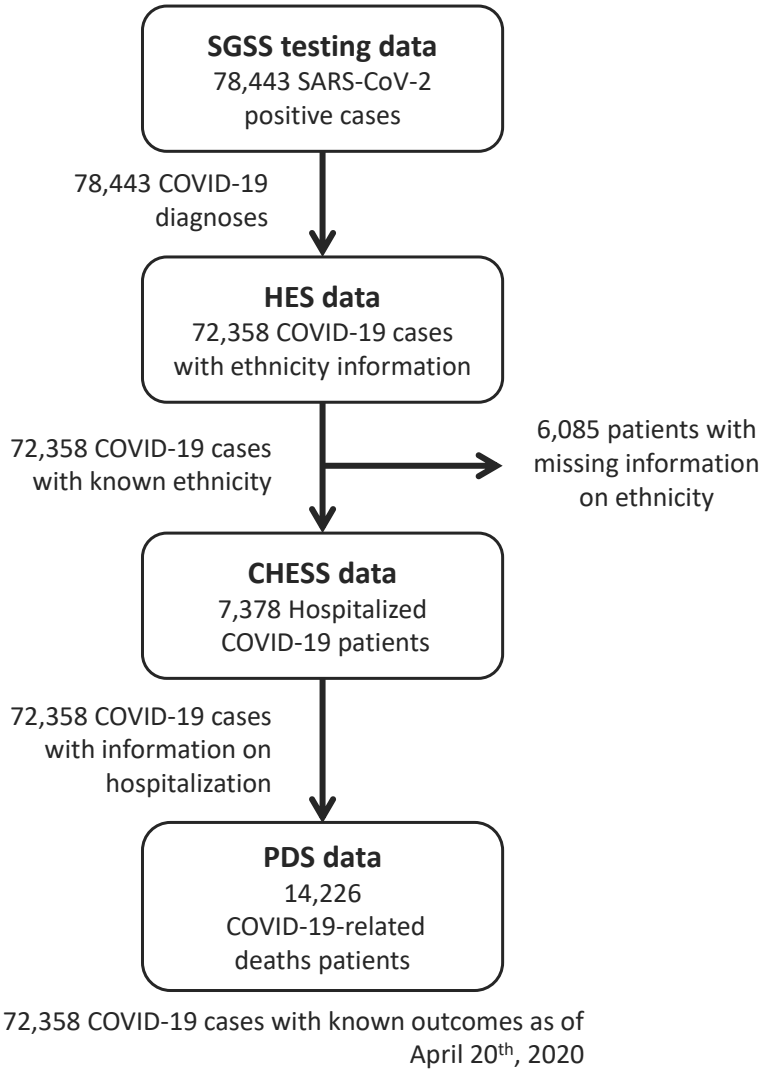


Figure 1: Data linkage and patient inclusion.

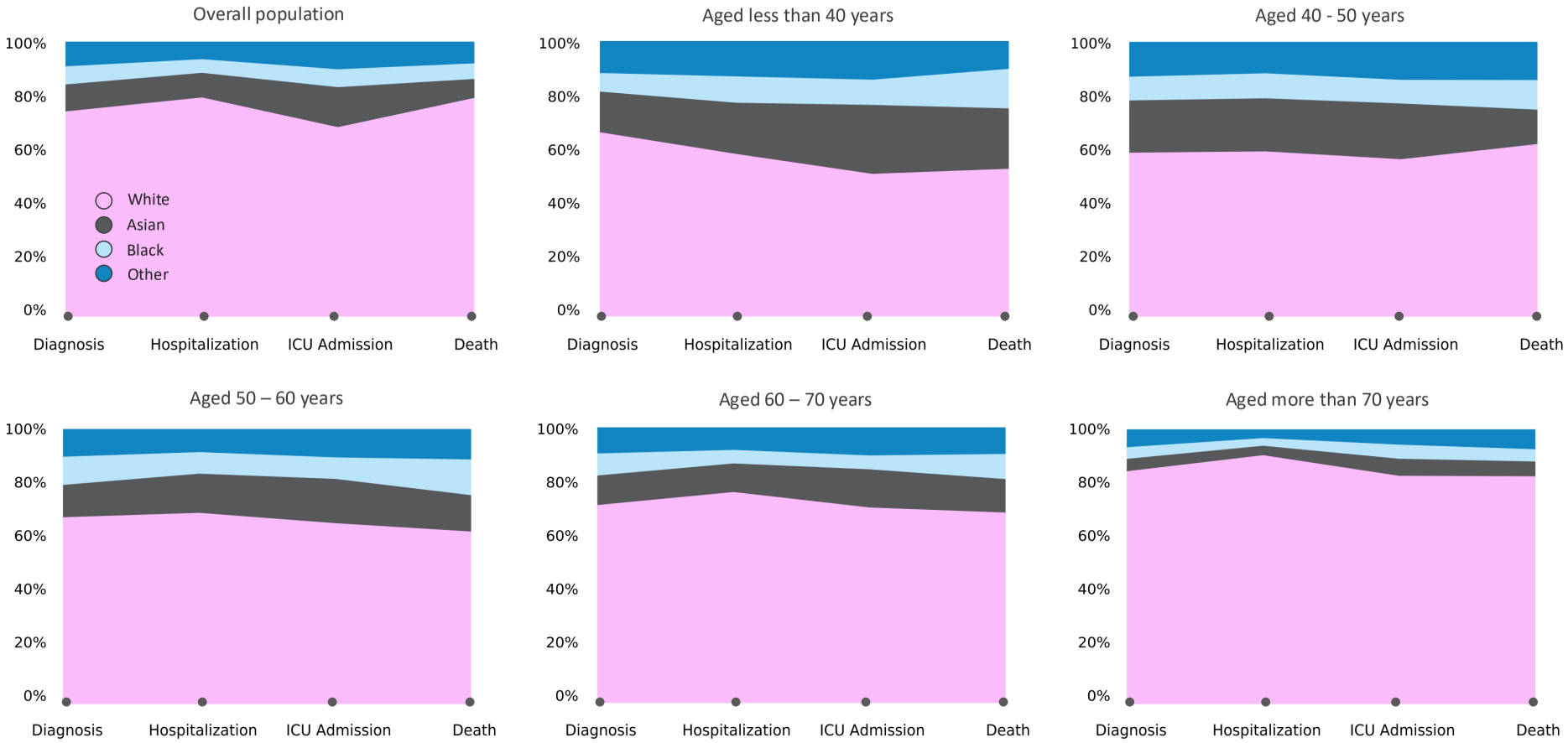


Figure 2: Representation of different ethnic groups for various age ranges across different stages of the COVID-19 trajectory.

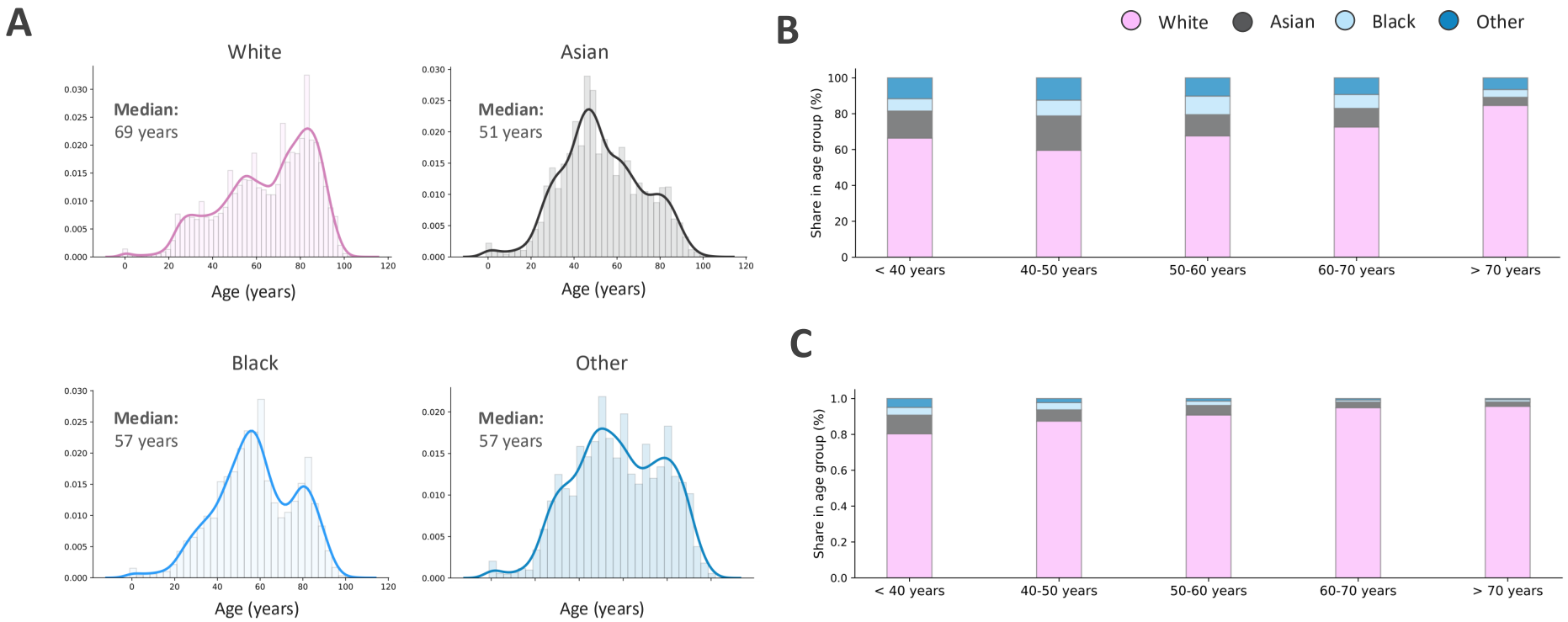


Figure 3: **(A)** Age distribution for COVID-19 patients within the different ethnic groups. **(B)** Representation of the different ethnic groups across various age ranges within the COVID-19 population. **(C)** Representation of the different ethnic groups across various age ranges in the population of England based on the 2018 population estimates.

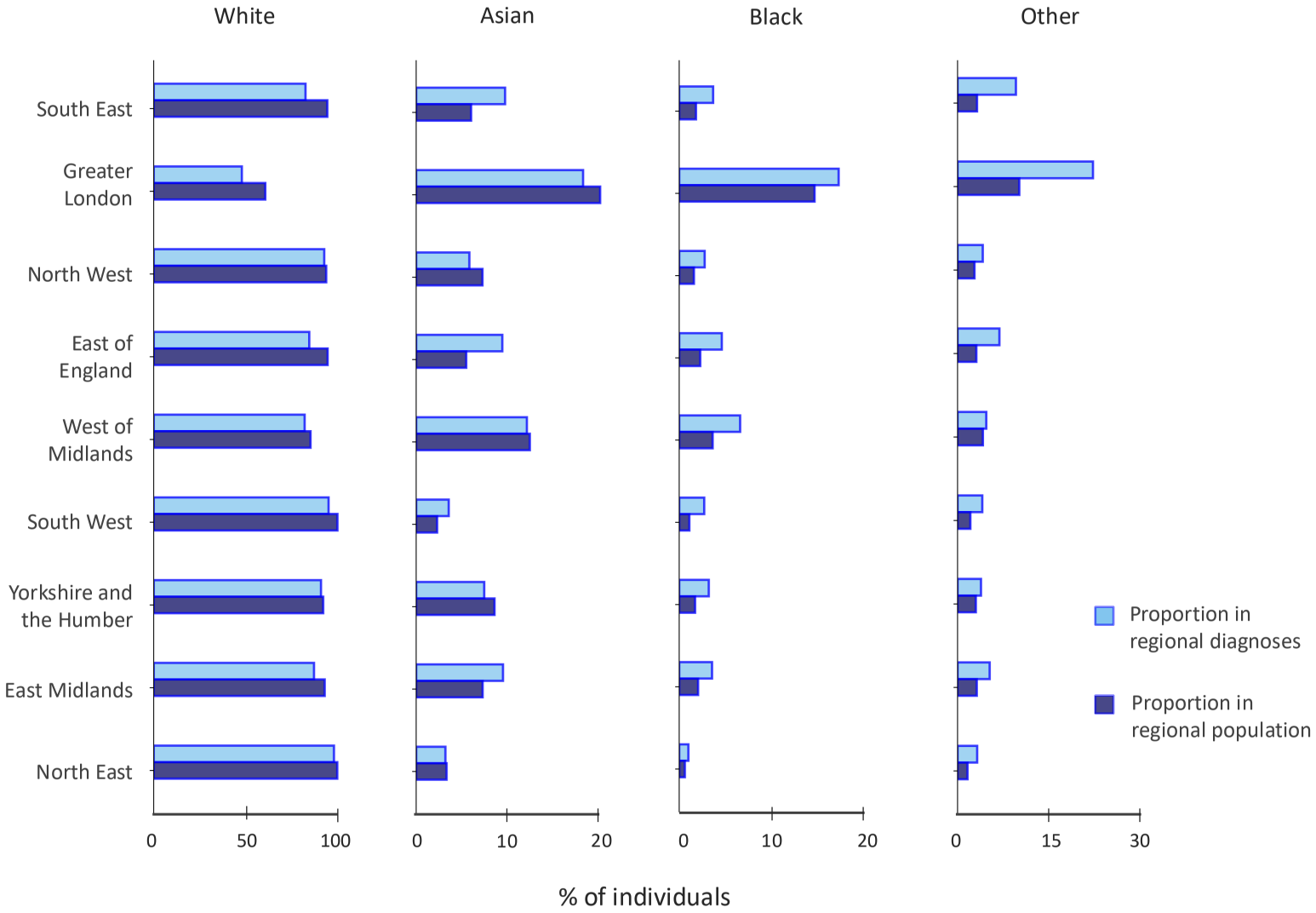


Figure 4: Regional proportion of COVID-19 diagnoses within the different ethnic groups compared to the regional demographics based on the 2018 population estimates.

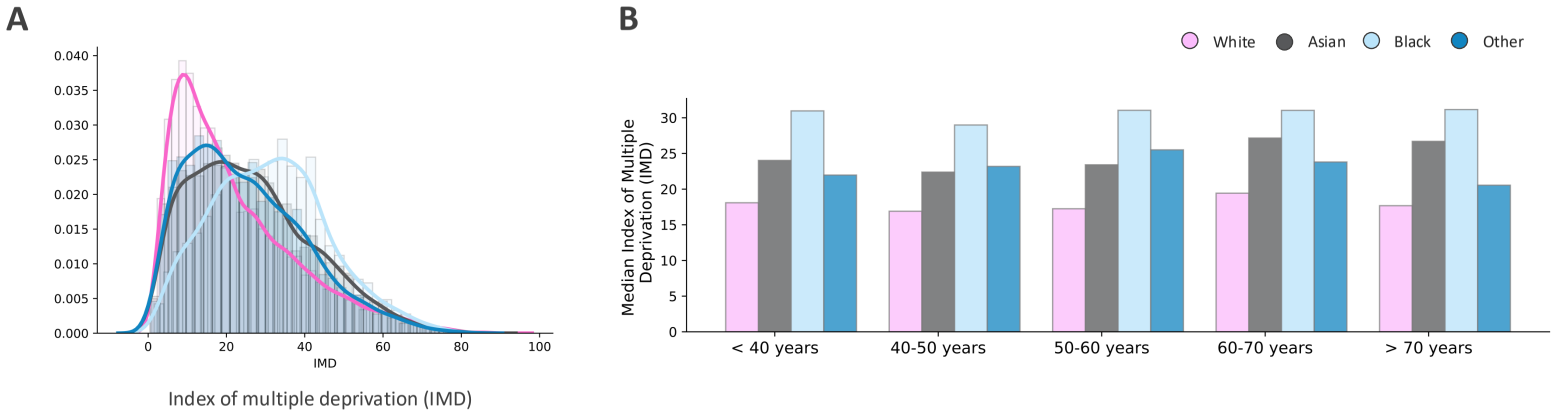


Figure 5: **(A)** Histogram for the index of multiple deprivation (IMD) among patients in various ethnic groups. **(B)** Median IMD in various age ranges within the different ethnic groups.

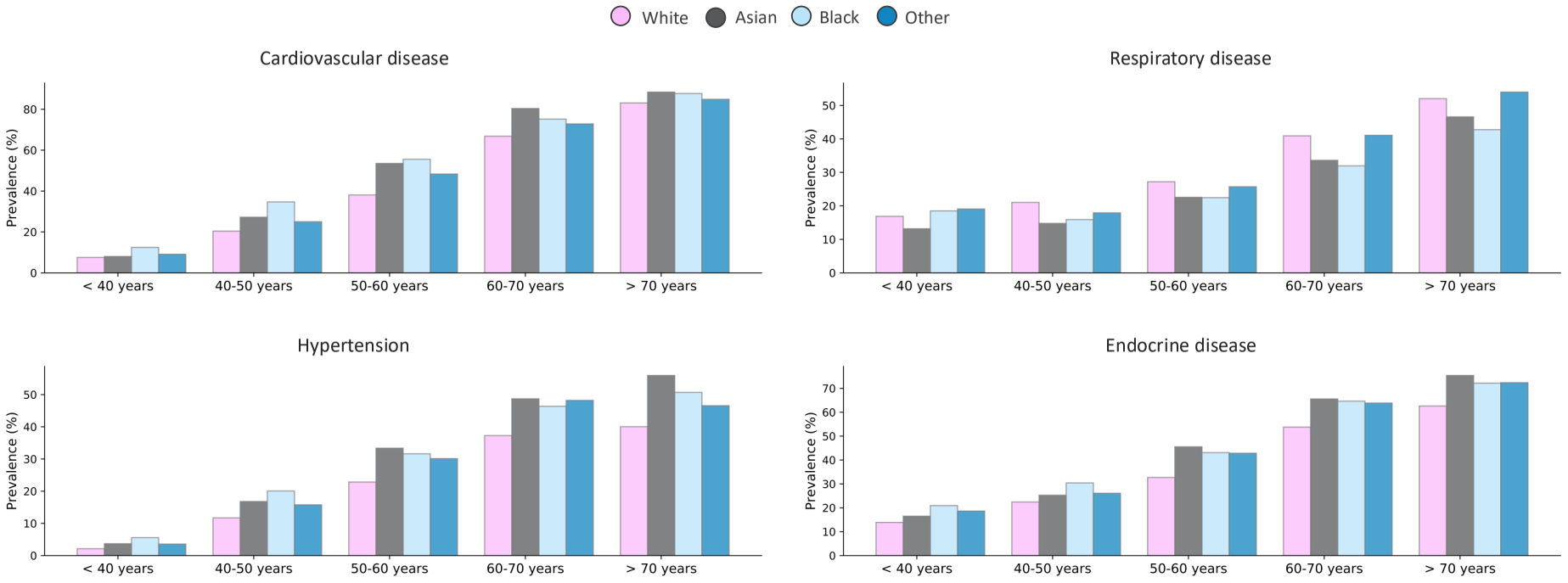


Figure 6: Prevalence of various comorbidities among different ethnic and age groups.

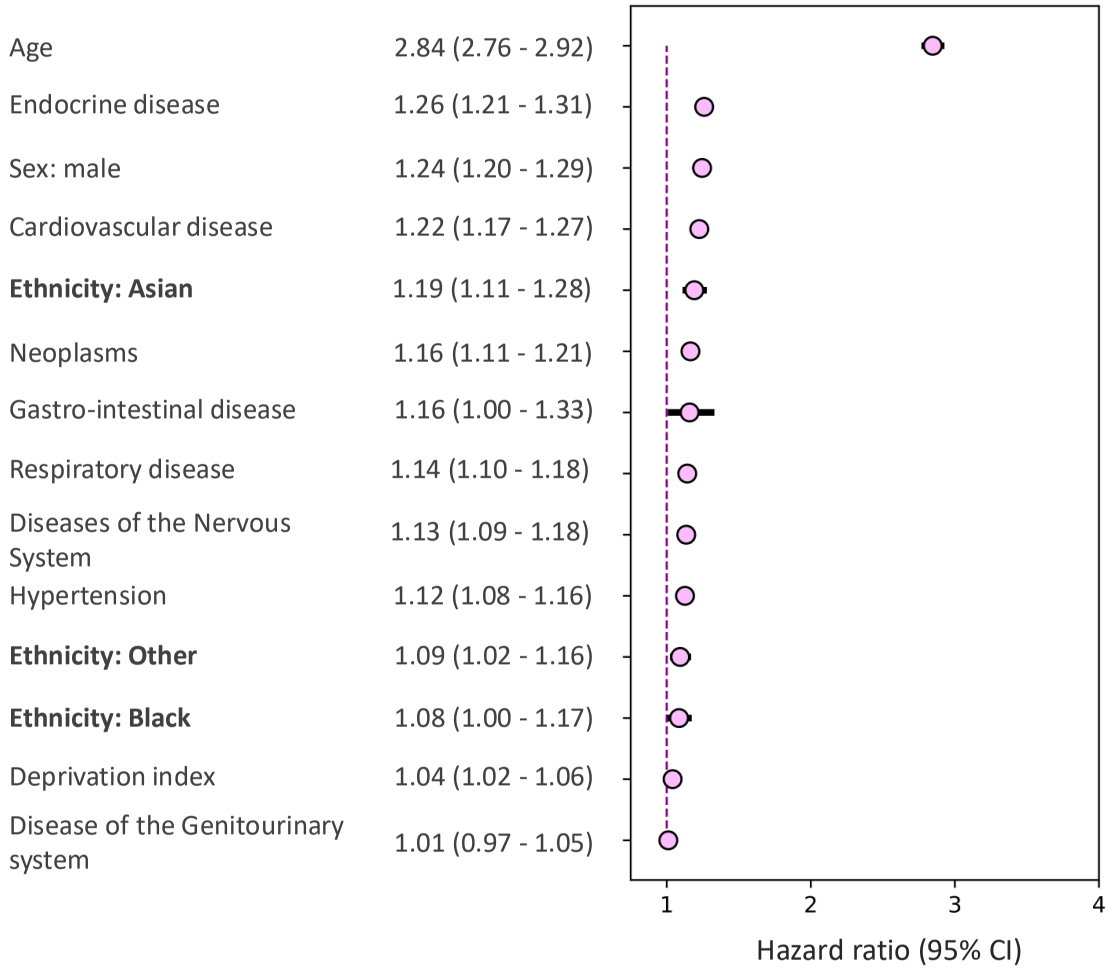


Figure 7: Coefficients of the Cox mixed-effects model for mortality.

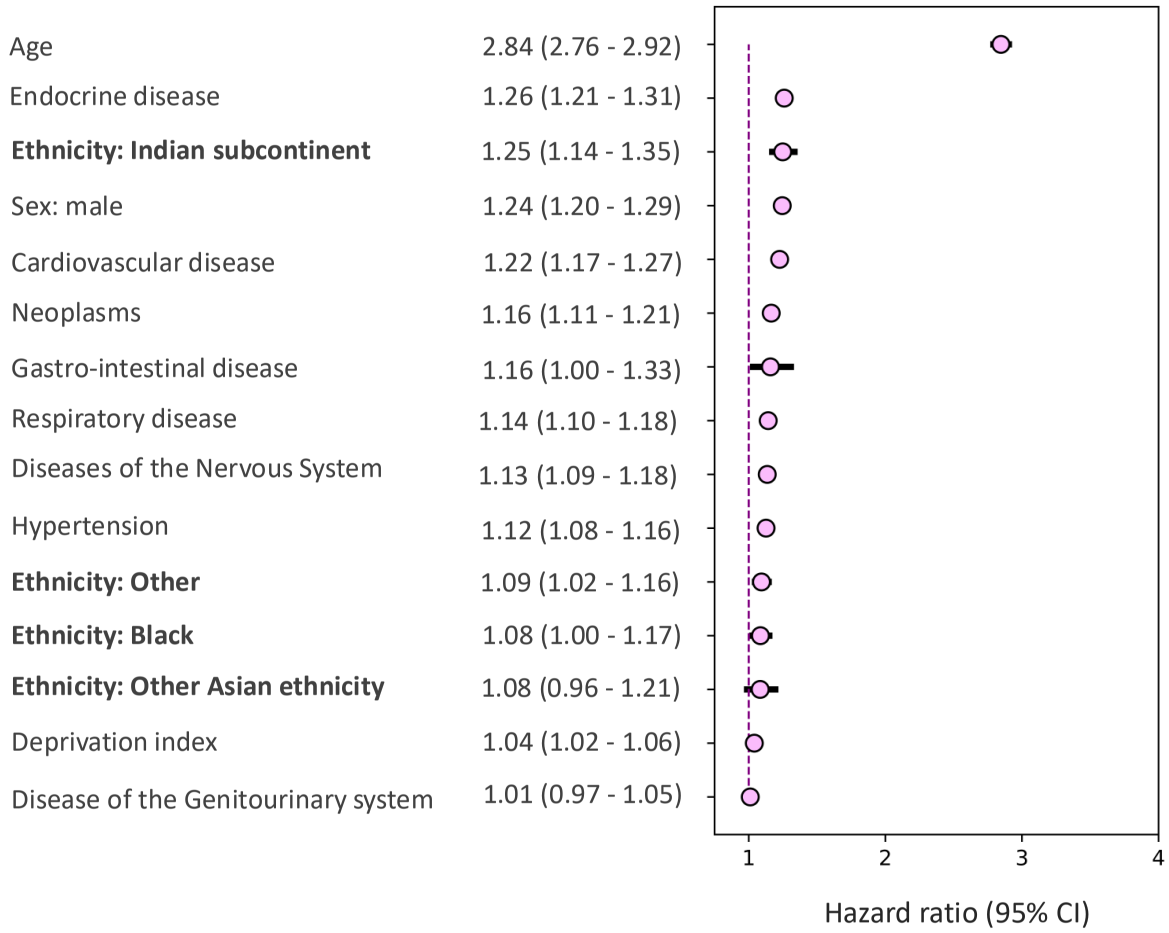


Figure 8: Coefficients of the Cox mixed-effects model for mortality with separate modeling of the Indian subcontinent ethnicity.

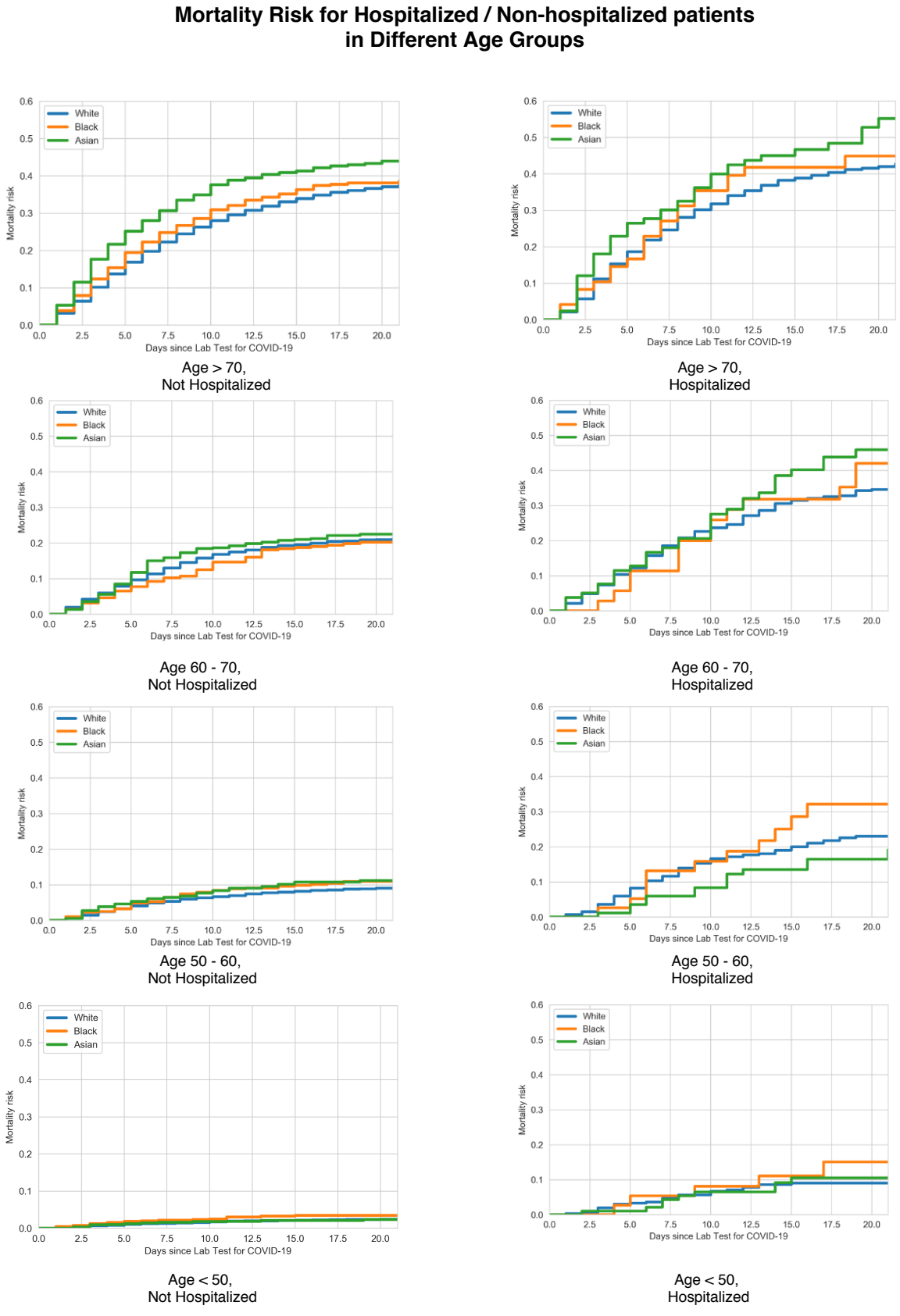


Figure 9: Death probabilities over time.