BMJ Open

# Ethnic variation in outcome of people hospitalised during the first Covid-19 epidemic wave in Wales, UK: An analysis of national surveillance data using Onomap, a name-based ethnicity classification tool



| | |
|---|---|
| Journal: | *BMJ Open* |
| Manuscript ID | bmjopen-2020-048335 |
| Article Type: | Original research |
| Date Submitted by the Author: | 22-Dec-2020 |
| Complete List of Authors: | Thomas, Daniel; Public Health Wales; Cardiff Metropolitan University, School of Health Sciences<br>Orife, Oghogho; Public Health Wales, Communicable Disease Surveillance Centre<br>Plimmer, Amy; Public Health Wales, Communicable Disease Surveillance Centre<br>Williams, Christopher; Public Health Wales<br>Karani, George; Cardiff Metropolitan University<br>Evans, Meirion; Public Health Wales<br>Longley, Paul; UCL<br>Janiec, Janusz; Narodowy Instytut Zdrowia Publicznego<br>Saltus, Roiyah; University of South Wales<br>Shankar, Ananda; Public Health Wales |
| Keywords: | COVID-19, EPIDEMIOLOGY, Public health < INFECTIOUS DISEASES |
| | |

*I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.*

*The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.*

*Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.*

# Ethnic variation in outcome of people hospitalised during the first Covid-19 epidemic wave in Wales (UK): An analysis of national surveillance data using Onomap, a name-based ethnicity classification tool

Daniel Rh Thomas,[1,2] Oghogho Orife,[1] Amy Plimmer, [1] Christopher Williams,[1] George Karani,[2] Meirion R Evans,[1] Paul A Longley,[3] Janusz Janiec,[4,] Roiyah Saltus,[5] A Giri Shankar[6]

1. Public Health Wales, Communicable Disease Surveillance Centre, Cardiff, Wales, UK
2. Cardiff Metropolitan University School of Health Sciences, Cardiff, Wales, UK
3. University College London, London, UK
4. Narodowy Instytut Zdrowia Publicznego, Warsaw, Poland
5. University of South Wales, Pontypridd, Wales, UK
6. Public Health Wales, Health Protection Division, Cardiff, Wales, UK

Correspondence: Daniel Rh Thomas (daniel.thomas@wales.nhs.uk)

Keywords: Covid-19; ethnicity; outcomes; epidemiology

## Abstract

### Objective

To identify risk of severe outcome in Black, Asian and minority ethnic populations hospitalised with Covid-19 during the first epidemic wave.

### Design

Descriptive analysis of 76,503 SARS-CoV-2 tests carried out in Wales to 31 May 2020. Cohort study of 4,046 individuals hospitalised with confirmed Covid-19 between 1st March and 31st May. In both analyses, ethnicity was assigned using a name-based classifier.

### Setting

Wales (UK)

### Primary and secondary outcomes

Admission to an intensive care unit following hospitalisation with a positive SARS-CoV-2 PCR test. Death within 28 days of a positive SARS-CoV-2 PCR test.

### Results

Using a name-based ethnicity classifier, we found that proportion of the Black, Asian and ethnic minority population tested for SARS-CoV-2 and proportion positive were higher in those classified as 'White'. Hospitalised Black, Asian and minority ethnic cases were younger (median age 53 compared to 76 years; $p<0.01$) and more likely to be admitted to intensive care. Bangladeshi (adjusted odds ratio: 9.80, 95%CI 1.21- 79.40) and 'White – Other than British or Irish' (aOR: 1.99, 95%CI: 1.15- 3.44) ethnic groups were most likely to be admitted to ICU. In Wales, older age (aOR for over 70 years: 10.29, 95%CI: 6.78–15.64) and male gender (aOR: 1.38, 95%CI: 1.19–1.59), but not ethnicity, were associated with death in hospitalised patients.

### Conclusions

This study adds to the growing evidence that ethnic minorities are disproportionately affected by Covid-19. During the first Covid-19 epidemic wave in Wales, although ethnic minority populations were less likely to be tested and

less likely to be hospitalised, those that did attend hospital were younger and more likely to be admitted to intensive care. Primary, secondary and tertiary prevention should target Black, Asian and minority ethnic communities in Wales.

## Strengths and limitations of this study

- Secondary analysis of data obtained through routine national Covid-19 surveillance.
- Ethnicity was assigned using a name-based classifier. This approach has both strengths and limitations. Studies relying on clinician reported ethnicity contain high proportions of missing and poor quality data. We were able to assign ethnicity to nearly all participants. Whilst sensitivity and specificity of the classifier varies in specific ethnic groups, and is poor in black British and people of mixed ethnicity, its performance is quantifiable and classification bias can be taken into account when interpreting findings.

Age, gender and deprivation were taken into account in the analysis, but individual data on history of chronic disease was poorly recorded, and treatment histories once hospitalised were not available.

## Introduction

There is growing evidence that Black, Asian and other minority ethnic (BAME) people living in Europe are at increased risk of infection with SARS-CoV-2 and, if infected, are more likely to have severe disease.[1] In the United Kingdom, the Intensive Care National Audit and Research Centre first raised concerns that BAME people were over-represented amongst Covid-19 patients admitted to intensive care.[2] These findings were reported widely in the media and discussed in opinion pieces. [3]–[7] In Wales, the First Minister established an advisory group to examine the issue and provide recommendations to reduce ethnic inequality in Covid-19 outcomes. [8]

Investigating ethnic health inequalities is hampered by the poor recording of ethnicity in clinical data. This is the case for Covid-19 notifications and laboratory reports in Wales. In order to rapidly investigate ethnic variation in Covid-19 epidemiology, we applied Onomap, a name-based ethnicity classification tool developed by the Department of Geography at University College London, [9] to routinely collected, named Covid-19 laboratory test data, held by Public Health Wales Communicable Disease Surveillance Centre.

## Methods

### Participants

We obtained routine surveillance data on 77,555 SARS-CoV-2 PCR tests carried out by Public Health Wales and authorised as at 1300 hrs, 31 May 2020 from Microbiology Datastore, a repository of test results recorded in the all-Wales Laboratory Information Management System.

Data were also obtained on records of 4,046 hospitalised patients (people admitted to hospital within 14 days of a positive SARS-CoV-2 test or individuals who tested positive for SARS-CoV-2 whilst in hospital) as at 1700 hrs, 31 May 2020 available in IC-Net, an infection prevention and control information management system. These data contained information on whether an individual was admitted to intensive care (ICU).

These individual person data on hospitalised cases were linked to records of 1,309 Covid-19 in-hospital deaths (Covid-19 cases who died in hospital, and had a positive test result for SARS-CoV-2 28 days or less prior to the date of death or 7 days after death) reported to Public Health Wales' Covid-19 mortality surveillance scheme to 1700hrs, 31 May, as at 28 June 2020.

### Ethnicity

Ethnicity was categorised using the name-based ethnicity classifier, Onomap, a software tool developed by geographers at University College London, and the 2001 Census classification of ethnicity.[10] We collapsed the Census categories further into: 'White British or Irish', 'White Other', 'Asian or British Asian', 'Black or Black British', 'other ethnicity' and 'unclassified', with a further aggregation to create a 'BAME' field, containing all ethnicities other than 'White British', 'White Irish', or 'White Other'. Unclassified observations were excluded.

### Deprivation

Small (Lower Super Output) areas in Wales were assigned a deprivation score using the Welsh Index of Multiple Deprivation[11] and areas were ordered into quintiles based on the distribution of these scores, ranging from least to most

deprived. Each individual was then assigned to a deprivation quintile based on their Lower Super Output Areas of residence.

**Statistical analysis**

Proportions of population tested, with 95% confidence intervals, were calculated for White and BAME groups using population data from the most recent Office for National Statistics Labour Force Survey.[12] Proportion testing positive with 95% confidence intervals were calculated by dividing number positive by number tested for the same time period.

Using the cohort of 4,046 hospitalised patients, we carried out a logistic regression to calculate odds ratios for the outcomes: (a) admitted to intensive care and (b) mortality, with 95% confidence intervals, for ethnic groups, in each case using 'White British or Irish' ethnicity as the baseline comparator. Independent variables were gender and age group. Multivariable analyses were then used to calculate odds ratios for ethnic groups whilst controlling for gender, age group and local area deprivation. Differences in the distributions of previously reported risk factors for fatal outcomes (age, gender, deprivation medical history) [13], [14] were investigated further in White and BAME groups. The Mann Whitney two-sample test was used to compare differences in the age distribution of BAME and White deaths. Odds ratios with 95% confidence intervals were calculated to compare proportion male and proportion with underlying health conditions amongst deceased BAME and White individuals. All analysis was carried out using Stata 14. [15]

**Validation of Onomap's performance**

The performance of Onomap was assessed using three data sets containing reliable self-reported or healthcare professional-reported ethnicity. These data were: A list of people attending a mosque in Wales who were offered screening for hepatitis C (n=189), a list of tuberculosis patients notified by doctors in Wales (n=3267) and a list of patients attending an infectious disease clinic in Poland (n=3184). Using these data as a 'gold standard', sensitivity and specificity were calculated to measure Onomap's performance in correctly classifying specific ethnicities.

## Ethical and privacy considerations

Ethical oversight of the project was provided by Public Health Wales NHS Trust R&D Division. As this work was carried out as part of the health protection response to a public health emergency in Wales, using routinely collected surveillance data, Public Health Wales R&D Division advised that NHS research ethics approval was not required. The use of named patient data in the investigation of communicable disease outbreaks and surveillance of notifiable disease is permitted under Public Health Wales' Establishment Order. Data were held and processed under Public Health Wales' information governance arrangements, in compliance with the Data Protection Act, Caldicott Principles and Public Health Wales guidance on the release of small numbers. No data identifying protected characteristics of an individual were released outside Public Health Wales. Validation work was from a project that had previously received permission from the Confidentiality Advisory Group to process patient data on tests for viral hepatitis carried out by laboratories in Wales, and research ethics approval from West of Scotland REC4 (Application title: Incidence of infectious disease in BME groups using Onomap; CAG reference: 16/CAG/0133; IRAS project ID: 210327 REC reference: 16/WS/018).

## Patient and Public Involvement statement

Patients or the public were involved in the design and conduct of our research and the work has been shared with the Welsh Government BAME COVID Advisory group, which contains community and stakeholder groups on a number of occasions. This research has also been presented to the Race Council Cymru.

## Funding

No external funding was sought. The study was done with existing Public Health Wales resources.

## Results

### Ethnicity classification

Onomap estimated the ethnicity of 98.1% (13,789/14,054) of tested individuals, 99.2% (4,013/4,046) of those hospitalised, 97.4% (305/313) of those admitted to intensive care, and 99.6% (1,304/1,309) of those who died following admission to hospital.

### Testing and hospitalisation

By classifying ethnicity using names, we estimate that 3.7% (n=2,896) tests were of Black, Asian and other minority (BAME) individuals (Table 1). Using the most recent Statistics Wales population estimates for ethnic groups in Wales,[12] this represents 1,580 tests per 100, 000 population in BAME, compared to 2,512 tests per 100,000 population in White ethnic groups.

Of 14,054 people tested positive for SARS-CoV-2 in Wales to 31 May 2020, Onomap classified 13,092 in White ethnic groups and 697 in BAME groups. Proportions with positive test results were similar for both groups: 447 per 100,000 of the White group tested positive and 380 per 100,000 in the BAME group. Trends in those tested positive should be interpreted with caution as they most likely reflect testing policy as well as incidence.

Of all those testing positive, a smaller proportion (15.1%) of those tested in the BAME group attended hospital compared to the White group (29.9%: see Table 2). However, the trend was reversed in people aged 50 to 59 years: 23.8% of positive BAME individuals aged 50-59 years attended hospital, compared to 16.3% of White individuals testing positive. The median age of hospitalised BAME individuals was 53 years compared to 76 years for White individuals ($p<0.01$; Mann Whitney 2 sample test).

### Admission to intensive care

Of those attending hospital, a much higher proportion (21.9%) of BAME individuals were admitted to intensive care compared to White individuals

(7.2%). Proportions of hospitalised patients admitted to intensive care (ICU) were highest amongst the 'Asian and British Asian - Indian, Pakistani and Bangladeshi' (29.0%) and 'White – other' (20.8%) groups. The median age of BAME patients admitted to ICU was 51 years compared to 58 years for White individuals ($p<0.01$; Mann Whitney 2 sample test). Amongst hospitalised patients aged between 50-59 years, 27.6% of BAME patients were admitted to ICU compared to 21% of White patients. More patients died in hospital without being admitted to ICU. Of all those attending hospital, 10.5% of patients identified as BAME died compared to 33.1% of White patients (Table 2).

We successfully linked all records of 4,046 people hospitalised with Covid-19, those admitted to ICU, and those who died in hospital, all as at 31 May 2020, using NHS numbers. Intensive care was more likely in hospitalised males (aOR: 2.03, 95% CI: 1.55-2.65) and in younger patients (Table 3, Figure 1). When specific ethnicities were examined, being admitted to ICU was more likely in 'White Other', 'Asian and British Asian – Bangladeshi', 'Asian and British Asian – Indian', and 'Asian and British Asian – Pakistani' ethnic groups. After adjusting for gender, age and social deprivation, 'White Other' (aOR: 1.99, 95% CI: 1.15-3.44), and 'Bangladeshi' (aOR: 9.8, 95%CI: 1.21-79.40), ethnic groups remained significantly more likely to be admitted to ICU.

Hospitalised cases living in the most deprived areas of Wales were significantly more likely to attend ICU (OR: 1.70; 95% CI 1.17-2.45). However, this effect did not remain significant after adjusting for age, sex and ethnicity (aOR: 1.37, 95% CI: 0.93-2.02).

**Mortality**

Likelihood of dying was significantly higher for hospitalised males. This effect remained after adjusting for age and ethnicity (aOR: 1.38, 95%CI: 1.19-1.59) (Table 3, Figure 1). No increase was observed in risk of death with increasing deprivation. There was a strong association between increasing age and death from Covid-19 which remained after adjusting for gender, ethnicity and social deprivation (aOR for aged 70 years and over: 10.29 (95% CI: 6.78-15.64). However, there was no evidence from this study that BAME groups were more

likely to die from Covid-19 than White-British or Irish groups, even after adjusting for gender, age and social deprivation (Table 3).

To investigate further, we compared the differences in the distribution of previously reported risk factors for fatal outcome[13] in White and BAME groups who had died. BAME people who died in Wales with Covid-19 were younger than White people who died (BAME median age 71 compared to 79 for White people; p=0.06, Mann-Whitney 2 sample test). Underlying chronic disease was recorded for 50% of deaths. There was a higher proportion of BAME people 72.7 (95% CI: 39.0-94.0) that had a history of underlying chronic disease compared to White people 49.6 (95% CI: 46.8- 52.3).

**Validation of Onomap**

Onomap returned predicted ethnicity for 97% of the 6640 names in the four data sets. Sensitivity and specificity was calculated for each ethnic group. Onomap generally had a high specificity, that is: it was unlikely to return a false ethnicity in people self-reporting a given ethnicity (Table 4). Specificity was 77% for white ethnicities, indicating that a proportion of people in BAME groups will be misclassified as white. In terms of its sensitivity, that is its likelihood of detecting all people self-reporting as an ethnic group, Onomap was poor for some ethnic groups, most notably for those self-reporting as black or British black. In other words, many people self-reporting as black or black British will be misclassified, most likely as white.

## Discussion

This was a rapid analysis of routinely available national surveillance data using name-based ethnicity classification software. It adds to the emerging evidence of increased risk of severe Covid-19 outcomes in ethnic minorities in Western Europe and the United States.

We found risk of severe Covid-19, indicated by admission to ICU, to be higher in many ethnic minorities living in Wales, and significantly higher in those of Bangladeshi ethnicity and in White ethnic groups, other than British or Irish. Bangladeshi communities have been identified in other studies as being at particular risk of the effects of Covid-19.[16] The second group we identified, 'White-other', will contain a range of nationalities, but in Wales, recent migrants from Eastern Europe will comprise a significant proportion. The risk associated with the latter group has not been previously reported, and is an important finding given recent outbreaks reported in factory settings in Wales where many European migrants are employed.

The finding that certain minority ethnic groups are at higher risk of being admitted to intensive care, but are no more likely to die than the White British and Irish group, was also found in the CO-CIN cohort study involving 23,577 Covid-19 patients attending hospitals in the UK.[16] This slightly counterintuitive finding may be a genuine finding or may be the result of classification bias. Firstly, if genuine, differences in the age distribution of cases in White and BAME groups are likely to be a factor. During early 2020, Covid-19 mortality was observed overwhelmingly in the elderly, with White men over 70 years disproportionally affected. These patients may have been less likely to have been admitted to ICU for treatment. Black, Asian and minority ethnic populations in Wales are generally younger [16], and lower median ages were observed in hospitalised BAME individuals. The finding that despite being generally younger, BAME individuals were more likely to be admitted to ICU is an important finding.

The lack of increased risk of mortality is at odds with some studies from England that have found that Covid-19 deaths have been disproportionally high in BAME

groups. Of course, it is always possible that Wales, with a more deprived general population relative to England, a lower density of BAME people and smaller urban conurbations, presents a less unequal risk setting. On the other hand, there may be methodological issues affecting this finding. It is likely that Onomap underestimates the absolute number of BAME individuals, particularly for Black groups. The misclassification of Black as White may have led to an under estimation of relative risk. It is also possible that date of onset to death is longer in younger people and that our study did not take sufficient account of this.

Onomap has been used widely as a tool in public health, for example in studies investigating variation in influenza mortality,[17] hepatitis B infection [18] and HPV vaccination uptake.[19] However, Onomap has limitations, and all findings should be interpreted in light of these. We previously validated the tool using data containing self-reported or healthcare professional-reported ethnicity. Onomap performs well for most ethnicities, but has a low sensitivity for Black or Black British individuals. Risks identified for Black and Black British groups are therefore likely to be underestimated. Kandt and Longley have published a comparison of Onomap with 2011 Census data.[20]

Wales is less ethnically diverse than many other areas of the United Kingdom, but its black, Asian and other minority ethnic (BAME) population has increased in recent years. In 2001, the Census recorded 2.1% of the population as BAME. This increased to 4.4% in the 2011 Census. Most recent estimates from the Labour Force Survey indicates that this has grown to 5.9%.[11] The Welsh Government has established an Advisory Group to investigate issues around Covid-19 in BAME groups and has published a series of recommendations. In Wales, an occupational risk assessment tool has been developed with the aim of reducing risk of infection in those most vulnerable to severe infection.[21] This tool, developed initially for the health care sector, is for all ethnicities, but includes a weighting to account for the emerging evidence of increased risk in BAME individuals.

One of the recommendations of the Welsh Government Advisory Group is to improve recording of ethnicity in routine health data, and a data improvement

plan is urgently required so ethnicity can be included in routine public health surveillance. There is an urgent need for all European countries carrying out Covid-19 surveillance to report trends by ethnicity, in order to inform local infection prevention and control policy and practice. Ethnic variation should also be considered in the design of interventions, and in crisis communication.

This is a complex topic and it is still unclear whether ethnic variation in poor outcomes is the result of higher incidence of infection or greater severity of disease. Minority ethnic groups are more likely to live in urban areas, to have public facing jobs, are more likely to live in crowded housing and live in multi-generational households.[22] Further research is needed to quantify risk of infection and risk of severe outcomes in ethnic minorities, and better understand the underlying processes behind any disparities. However, there is now probably enough evidence to act, and effort should now be focussed on designing innovative interventions for primary, secondary and tertiary prevention of Covid-19 in minority ethnic groups.[23]

## Acknowledgements

Onomap was purchased from Publicprofiler Ltd. The authors acknowledge the many laboratory and surveillance staff in Public Health Wales involved in developing and maintaining routine Covid-19 surveillance in Wales. Victoria McClure assisted with extracting data from IC-Net.

### Conflict of interest

Paul Longley is Director of Publicprofiler Ltd.

### Authors' contributions

Daniel Thomas designed the study, contributed to the analysis and wrote the manuscript. Oghogho Orife contributed to the analysis and commented on the manuscript. Amy Plimmer, George Karani, Meirion Evans, Janusz Janiec and Roiyah Saltus commented on the manuscript and contributed to the validation work. Paul Longley commented on the methodology and results, including use of the names classification tool. Chris Williams and Giri Shankar commented on the design and analysis and manuscript.

### Exclusive Licence

("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

## References


[1] D. Pan *et al.*, "The impact of ethnicity on clinical outcomes in COVID-19: A systematic review," *EClinicalMedicine*, vol. 23, p. 100404, Jun. 2020, doi: 10.1016/j.eclinm.2020.100404.

[2] "ICNARC – Reports," 2020. [Online]. Available: https://www.icnarc.org/Our-Audit/Audits/Cmp/Reports. [Accessed: 14-Jun-2020].

[3] Guardian, "Ethnic minorities dying of Covid-19 at higher rate, analysis shows | World news | The Guardian," 2020.

[4] R. Cifuentes, "All in it together? The impact of Coronavirus on BAME people in Wales. | Bevan Foundation," 2020.

[5] M. Pareek *et al.*, "Ethnicity and COVID-19: an urgent public health research priority," *The Lancet*, vol. 395, no. 10234. Lancet Publishing Group, pp. 1421–1422, 02-May-2020, doi: 10.1016/S0140-6736(20)30922-3.

[6] BMA. Press Release, "Review into COVID-19 impact on BAME communities must be backed by real-time data and include measures to address problem now, says BMA," 2020. [Online]. Available: https://www.bma.org.uk/bma-media-centre/review-into-covid-19-impact-on-bame-communities-must-be-backed-by-real-time-data-and-include-measures-to-address-problem-now-says-bma. [Accessed: 14-Jun-2020].

[7] NHS Confederation, "The impact of COVID-19 on BME communities and health and care staff," 2020. [Online]. Available: https://www.nhsconfed.org/resources/2020/04/the-impact-of-covid19-on-bme-communities-and-staff. [Accessed: 14-Jun-2020].

[8] Welsh Government, "Wales BAME Covid-19 health advisory group takes a cross-Government approach | GOV.WALES," 2020.

[9] "Public Profiler," 2020. [Online]. Available: https://www.publicprofiler.org/. [Accessed: 14-Jun-2020].

[10] Office for National Statistics, "[ARCHIVED CONTENT] Ethnic Group Statistics: a guide for the collection and classification of ethnicity data: Office for National Statistics," 2003.

[11] StatsWales, "Welsh Index of Multiple Deprivation." [Online]. Available: https://statswales.gov.wales/Catalogue/Community-Safety-and-Social-Inclusion/Welsh-Index-of-Multiple-Deprivation. [Accessed: 14-Jun-2020].

[12] StatsWales, "Ethnicity by area and ethnic group." [Online]. Available: https://statswales.gov.wales/Catalogue/Equality-and-Diversity/Ethnicity/ethnicity-by-area-ethnicgroup. [Accessed: 14-Jun-2020].

[13] E. Williamson *et al.*, "OpenSAFELY: factors associated with COVID-19-related hospital death in the linked electronic health records of 17 million adult NHS patients.," *medRxiv*, p. 2020.05.06.20092999, May 2020, doi: 10.1101/2020.05.06.20092999.

[14] A. K. Clift *et al.*, "Living risk prediction algorithm (QCOVID) for risk of hospital admission and mortality from coronavirus 19 in adults: national derivation and validation cohort study," *BMJ*, vol. 371, Oct. 2020, doi: 10.1136/bmj.m3731.

[15] StataCorp LLC, "Stata 14 | Stata." [Online]. Available: https://www.stata.com/stata14/. [Accessed: 14-Jun-2020].

[16] E. Harrison, A. Docherty, C. Semple, and CO-CIN, "CO-CIN: Investigating associations between ethnicity and outcome from COVID-19 - report to SAGE, 25 April 2020 - GOV.UK," 2020. [Online]. Available: https://www.gov.uk/government/publications/co-cin-investigating-associations-between-ethnicity-and-outcome-from-covid-19-report-to-sage-25-april-2020. [Accessed: 14-Jun-2020].

[17] H. Zhao, R. J. Harris, J. Ellis, and R. G. Pebody, "Ethnicity, deprivation and mortality due to 2009 pandemic influenza A(H1N1) in England during the 2009/2010 pandemic and the first post-pandemic season," *Epidemiol. Infect.*, vol. 143, no. 16, pp. 3375–3383, Dec. 2015, doi: 10.1017/S0950268815000576.

[18] M. Binka *et al.*, "Differing profiles of people diagnosed with acute and chronic hepatitis B virus infection in British Columbia, Canada," *World J. Gastroenterol.*, vol. 24, no. 11, pp. 1216–1227, Mar. 2018, doi: 10.3748/wjg.v24.i11.1216.

[19] K. G. Pollock *et al.*, "Evidence of decreased HPV vaccine acceptance in Polish communities within Scotland," *Vaccine*, vol. 37, no. 5, pp. 690–692, Jan. 2019, doi: 10.1016/j.vaccine.2018.10.097.

[20] J. Kandt and P. A. Longley, "Ethnicity estimation using family naming practices," *PLoS One*, vol. 13, no. 8, p. e0201774, Aug. 2018, doi: 10.1371/journal.pone.0201774.

[21] Welsh Government, "COVID-19 workforce risk assessment tool | GOV.WALES," 2020. [Online]. Available: https://gov.wales/covid-19-workforce-risk-assessment-tool. [Accessed: 14-Jun-2020].

[22] Office for National Statistics, "2011 Census General Report - Office for National Statistics," 2011. [Online]. Available: https://www.ons.gov.uk/census/2011census/howourcensusworks/howdidwedoin2011/2011censusgeneralreport. [Accessed: 14-Jun-2020].

[23] Z. Haque, L. Becares, and N. Treloar, "A Runnymede Trust and ICM Survey Over-Exposed and Under-Protected The Devastating Impact of COVID-19 on Black and Minority Ethnic Communities in Great Britain Runnymede: Intelligence for a Multi-ethnic Britain," 2020.

**Table 1. Proportions of the population tested for SARS-CoV-2 in Wales and associated proportions of positive test results, White and BAME individuals**

| Ethnicity | Estimated population[1] | Number of tests for SARS-Cov-2 | Tests per 100 000 pop (95%CI) | Positive test results | Positive tests per 100 000 pop (95%CI) |
|---|---|---|---|---|---|
| White Ethnicities[2] | 2,930,200 | 73,607 | 2,512 (2,494-2,530) | 13092 | 447 (439-455) |
| Non-White ethnicities[2] | 183,300 | 2,896 | 1,579 (1,523-1,638) | 697 | 380 (353-410) |
| All Ethnicities | 3,113,500 | 76,503 | 2,457 (2,440-2,474) | 13,789 | 443 (436-450) |

[1] 2019 estimates of White and non-White populations in Wales taken from ONS Labour Force Survey, 2019

[2] Onomap estimates of ethnicity

**Table 2. Ethnicity breakdown of individuals tested for SARS-CoV-2 in Wales and proportions hospitalised, attending intensive care units (ICU), and deceased**

| Ethnicity[1] | Positive tests | Hospital-ised | % hospitalised (95%CI) | | Attending ICU | % hospitalised attending ICU (95% CI) | | Deceased citing Covid-19 | % hospitalised and deceased citing Covid-19 (95%CI) | |
|---|---|---|---|---|---|---|---|---|---|---|
| All ethnicities | 14,054 | 4,046 | 28.8 | (28.0-29.5) | 313 | 7.7 | (6.9-8.6) | 1,309 | 32.4 | (30.9-33.8) |
| White British and Irish | 12,565 | 3812 | 30.3 | (29.5-31.2) | 262 | 6.9 | (6.1-7.7) | 1,274 | 33.4 | (31.9-34.9) |
| White – other | 527 | 96 | 18.2 | (15.0-21.8) | 20 | 20.8 | (13.2-30.3) | 19 | 19.8 | (12.4-29.2) |
| **All white ethnicities** | **13,092** | **3908** | **29.9** | **(29.1-30.6)** | **282** | **7.2** | **(6.4-8.1)** | **1,293** | **33.1** | **(31.6-34.6)** |
| Asian and British Asian (Indian, Pakistani, Bangladeshi) | 371 | 62 | 16.7 | (13.1-20.9) | 18 | 29.0 | (18.1-41.9) | <10 | - | |
| Black and Black British | 54 | <10 | - | | <10 | - | | <10 | - | |
| Chinese | 40 | <10 | - | | <10 | - | | <10 | - | |
| Other | 232 | 30 | 12.9 | (8.9-17.9) | <10 | - | | <10 | - | |
| **All non-white ethnicities** | **697** | **105** | **15.1** | **(12.5-17.9)** | **23** | **21.9** | **(14.4-31.0)** | **11** | **10.5** | **(5.3-18.0)** |
| Unknown | 265 | 12 | 4.5 | (2.4-7.7) | 8 | 66.7 | (34.9-90.1) | 5 | 41.7 | (15.2-72.3) |

[1] Onomap estimates of ethnicity

**Table 3. Personal characteristics associated with severe outcomes for Welsh residents hospitalised with Covid-19, to 31 May 2020**

| Personal characteristic | | Hospitalised | Received intensive care | | | | Died | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Univariate analysis | | Multivariate analysis | | Univariate analysis | | Multivariate analysis | |
| Gender | | | OR | (95%CI) | OR | (95%CI) | OR | (95%CI) | OR | (95%CI) |
| | Female | 1920 | - | | - | | - | | - | |
| | Male | 2126 | **1.90** | **(1.49-2.43)** | **2.03** | **(1.55-2.65)** | **1.28** | **(1.12-1.47)** | **1.38** | **(1.19-1.59)** |
| Age | | | | | | | | | | |
| | 0-49 | 428 | - | | - | | - | | - | |
| | 50-59 | 466 | 1.24 | (0.89-1.72) | 1.27 | (0.89-1.80) | **2.43** | **(1.49-3.95)** | **2.29** | **(1.40-3.74)** |
| | 60-69 | 569 | **0.64** | **(0.45-0.91)** | **0.64** | **(0.45-0.93)** | **5.93** | **(3.80-9.25)** | **5.11** | **(3.26-8.03)** |
| | 70 and over | 2583 | **0.11** | **(0.07-0.15)** | **0.11** | **(0.08-0.17)** | **11.40** | **(7.55-17.20)** | **10.29** | **(6.78-15.64)** |
| Ethnicity | | | | | | | | | | |
| | White: British | 3,716 | - | | - | | - | | - | |
| | White: Irish | 96 | 0.43 | (0.14-1.37) | 0.25 | (0.06-1.06) | 0.73 | (0.47-1.16) | 0.68 | (0.42-1.10) |
| | White: other | 96 | **3.51** | **(2.11-5.84)** | **1.99** | **(1.15-3.44)** | **0.49** | **(0.29- 0.81)** | 0.82 | (0.46-1.44) |
| | Asian and British Asian: Bangladeshi | 5 | **8.89** | **(1.48-53.49)** | **9.8** | **(1.21-79.40)** | 0.49 | (0.06-4.43) | 0.61 | (0.06-6.29) |
| | Asian and British Asian: Chinese | 6 | 2.67 | (0.31-22.93) | 1.78 | (0.19-16.85) | 0.40 | (0.05-3.39) | 0.97 | (0.10-9.36) |
| | Asian and British Asian: Indian | 16 | **6.07** | **(2.09-17.59)** | 2.49 | (0.83-7.53) | 0.46 | (0.13-1.60) | 1.06 | (0.27-4.14) |
| | Asian and British Asian: Pakistani | 41 | **4.89** | **(2.42-9.88)** | 1.91 | (0.89-4.09) | **0.16** | **(0.05-0.51)** | 0.44 | (0.13-1.50) |
| | Black and Black British: African | 7 | 2.22 | (0.27-18.55) | 0.93 | (0.10-8.59) | 0.33 | (0.04-2.74) | - | - |
| | Any other ethnic group | 30 | 1.54 | (0.46-5.12) | 0.59 | (0.17-2.01) | **0.15** | **(0.03-0.62)** | 0.25 | (0.06-1.12) |
| | Unknown ethnicity | 33 | **4.27** | **(1.91-9.56)** | **2.99** | **(1.25-7.18)** | **0.35** | **(0.14-0.92)** | 0.59 | (0.22-1.61) |
| Deprivation | | | | | | | | | | |
| | Most deprived | 939 | **1.70** | **(1.17-2.45)** | **1.37** | (0.93-2.02) | 0.95 | (0.77-1.16) | 1.10 | (0.88-1.36) |
| | Quintile 2 | 867 | 1.25 | (0.85-1.84) | 1.03 | (0.69-1.56) | 0.93 | (0.76-1.15) | 1.06 | (0.85-1.32) |

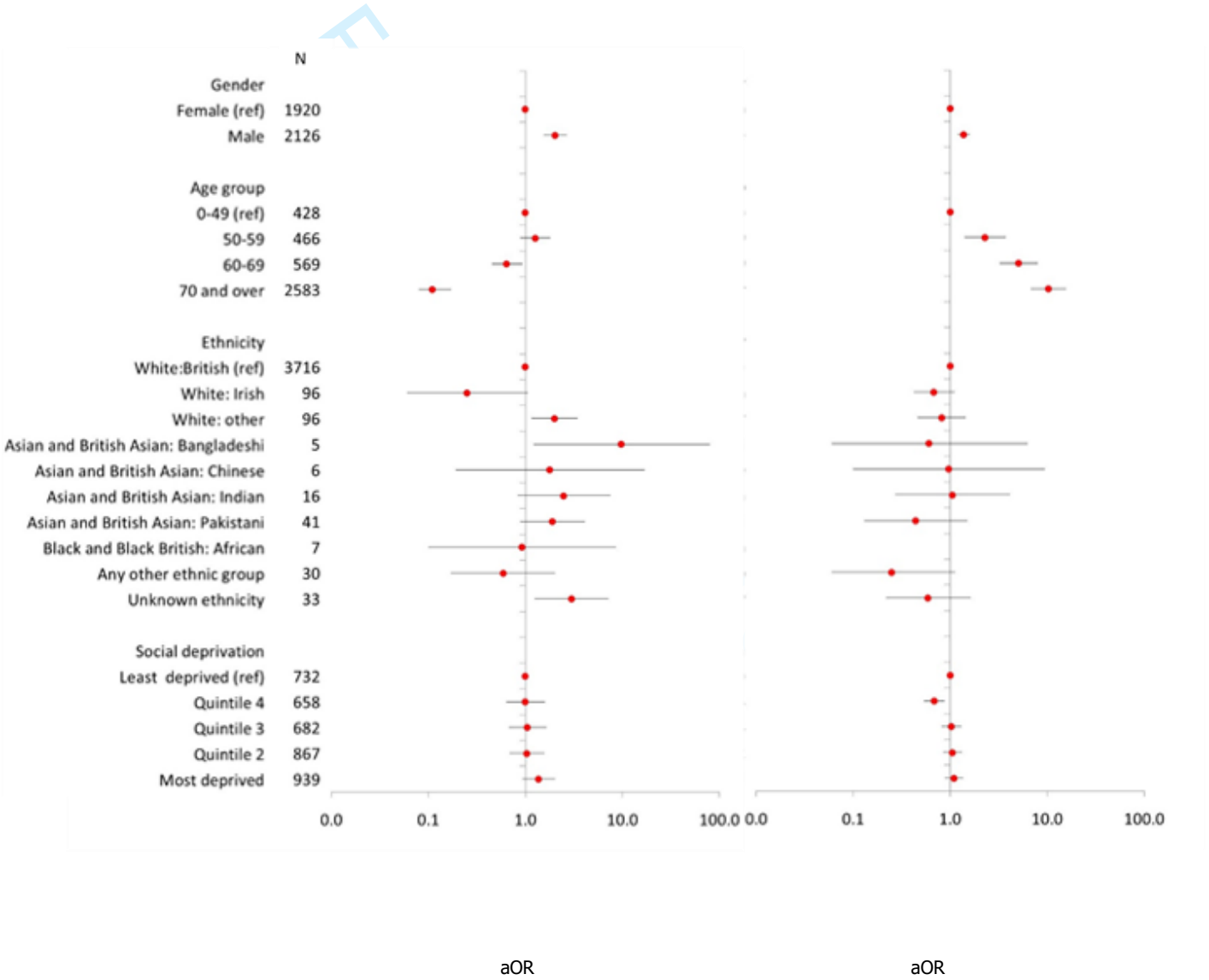
| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Quintile 3 | 682 | 1.08 | (0.71-1.64) | 1.05 | (0.68-1.65) | 1.01 | (0.81-1.26) | 1.03 | (0.82-1.30) |
| | Quintile 4 | 658 | 1.09 | (0.72-1.67) | 1.00 | (0.64-1.58) | **0.68** | **(0.54-0.86)** | **0.69** | **(0.54-0.87)** |
| | Least Deprived | 732 | - | | - | | - | | - | |

Table 4. Validation of Onomap. Estimated sensitivity and specificity of Onomap by ethnic group. Calculated by measuring the performance of Onomap to predict ethnicity in three clinical data sets[1] already containing self-reported or healthcare professional-reported ethnicity

| Ethnicity | Ethnicity reported by participant | Ethnicity predicted by Onomap | Ethnicity correctly predicted | Sensitivity | Specificity |
|---|---|---|---|---|---|
| White British or Irish | 1681 | 1811 | | | |
| Other White | 3235 | 3418 | | | |
| **Total White** | 4916 | 5229 | 4844 | **98.5%** | **77.7%** |
| | | | | | |
| Indian | 364 | 239 | | | |
| Pakistani | 313 | 348 | | | |
| Bangladeshi | 96 | 88 | | | |
| Chinese | 55 | 18 | | | |
| Other Asian | 9 | 118 | | | |
| **Total Asian or Asian British** | 837 | 811 | 609 | **72.8%** | **96.5%** |
| | | | | | |
| Black- African | 344 | 142 | | | |
| Black - Caribbean | 10 | 1 | | | |
| Other Black | 23 | 0 | | | |
| **Total Black or Black British** | 377 | 143 | 112 | **29.7%** | **99.5%** |
| | | | | | |
| Arabic | 39 | 279 | | | |
| Other | 9 | 4 | | | |
| **Other Ethnic Group** | 45 | 283 | 24 | **53.3%** | **96.1%** |
| | | | | | |
| **Mixed** | 234 | 0 | - | | |
| | | | | | |
| **Unclassified/Unknown** | 231 | 174 | 8 | **3.5%** | **97.4%** |
| | | | | | |
| **Total** | 6640 | 6640 | **5589** | **87.4%** | **96.1%** |

[1] Three data sets which included self-reported or healthcare professional-reported ethnicity were used to validate Onomap: A list of individuals attending a mosque in Wales who were offered screening for hepatitis C (n=189), a list of tuberculosis patients notified by doctors in Wales (n=3267) and a list of patients attending an infectious disease clinic in Poland (n=3184).

Figure 1. Determinants of: 1. Being admitted to intensive care unit (ICU); and 2. In-hospital mortality in 4,046 individuals hospitalised with Covid-19 in Wales to 31 May 2020, as at 28 June 2020. Adjusted odds ratios (aOR) with 95% confidence intervals are given for male gender, compared to female, older age groups compared to those aged less than 50 years, and Onomap estimated ethnicities, compared to 'White British'. Odds ratios greater than one represent an increased risk; odds ratios less than one represent a decreased risk. 95% confidence intervals not crossing one reflect that the odds ratio is statistically significant.



STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

| | Item No | Recommendation | |
|---|---|---|---|
| **Title and abstract** | 1 | (*a*) Indicate the study's design with a commonly used term in the title or the abstract | ✓ |
| | | (*b*) Provide in the abstract an informative and balanced summary of what was done and what was found | ✓ |
| **Introduction** | | | |
| Background/rationale | 2 | Explain the scientific background and rationale for the investigation being reported | ✓ |
| Objectives | 3 | State specific objectives, including any prespecified hypotheses | ✓ |
| **Methods** | | | |
| Study design | 4 | Present key elements of study design early in the paper | ✓ |
| Setting | 5 | Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection | ✓ |
| Participants | 6 | (*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up | ✓ |
| | | (*b*) For matched studies, give matching criteria and number of exposed and unexposed | |
| Variables | 7 | Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable | ✓ |
| Data sources/ measurement | 8* | For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group | ✓ |
| Bias | 9 | Describe any efforts to address potential sources of bias | ✓ |
| Study size | 10 | Explain how the study size was arrived at | No sample taken. Used Welsh population |
| Quantitative variables | 11 | Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why | ✓ |
| Statistical methods | 12 | (*a*) Describe all statistical methods, including those used to control for confounding | ✓ |
| | | (*b*) Describe any methods used to examine subgroups and interactions | ✓ |
| | | (*c*) Explain how missing data were addressed | ✓ |
| | | (*d*) If applicable, explain how loss to follow-up was addressed | Not applicable |
| | | (*e*) Describe any sensitivity analyses | ✓ |
| **Results** | | | |
| Participants | 13* | (a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed | ✓ |
| | | (b) Give reasons for non-participation at each stage | ✓ |

| | | | |
|---|---|---|---|
| | | (c) Consider use of a flow diagram | ✓ |
| Descriptive data | 14* | (a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders | ✓ |
| | | (b) Indicate number of participants with missing data for each variable of interest | ✓ |
| | | (c) Summarise follow-up time (eg, average and total amount) | ✓ |
| Outcome data | 15* | Report numbers of outcome events or summary measures over time | ✓ |
| Main results | 16 | (*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included | ✓ |
| | | (*b*) Report category boundaries when continuous variables were categorized | ✓ |
| | | (*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period | ✓ |
| Other analyses | 17 | Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses | ✓ |
| **Discussion** | | | |
| Key results | 18 | Summarise key results with reference to study objectives | ✓ |
| Limitations | 19 | Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias | ✓ |
| Interpretation | 20 | Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence | ✓ |
| Generalisability | 21 | Discuss the generalisability (external validity) of the study results | ✓ |
| **Other information** | | | |
| Funding | 22 | Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based | ✓ |

*Give information separately for exposed and unexposed groups.

**Note:** An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.

BMJ Open

# Ethnic variation in outcome of people hospitalised during the first COVID-19 epidemic wave in Wales (UK): An analysis of national surveillance data using Onomap, a name-based ethnicity classification tool



| | |
|---|---|
| Journal: | *BMJ Open* |
| Manuscript ID | bmjopen-2020-048335.R1 |
| Article Type: | Original research |
| Date Submitted by the Author: | 21-Apr-2021 |
| Complete List of Authors: | Thomas, Daniel; Public Health Wales; Cardiff Metropolitan University, School of Health Sciences<br>Orife, Oghogho; Public Health Wales, Communicable Disease Surveillance Centre<br>Plimmer, Amy; Public Health Wales, Communicable Disease Surveillance Centre<br>Williams, Christopher; Public Health Wales<br>Karani, George; Cardiff Metropolitan University<br>Evans, Meirion; Public Health Wales<br>Longley, Paul; UCL<br>Janiec, Janusz; Narodowy Instytut Zdrowia Publicznego<br>Saltus, Roiyah; University of South Wales<br>Shankar, Ananda; Public Health Wales |
| Primary Subject Heading: | Epidemiology |
| Secondary Subject Heading: | Public health |
| Keywords: | COVID-19, EPIDEMIOLOGY, Public health < INFECTIOUS DISEASES |
| | |

*I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.*

*The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.*

*Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.*

# Ethnic variation in outcome of people hospitalised during the first COVID-19 epidemic wave in Wales (UK): An analysis of national surveillance data using Onomap, a name-based ethnicity classification tool

Daniel Rh Thomas,[1,2] Oghogho Orife,[1] Amy Plimmer, [1] Christopher Williams,[1] George Karani,[2] Meirion R Evans,[1] Paul A Longley,[3] Janusz Janiec,[4,] Roiyah Saltus,[5] A Giri Shankar[6]

1. Public Health Wales, Communicable Disease Surveillance Centre, Cardiff, Wales, UK
2. Cardiff Metropolitan University School of Health Sciences, Cardiff, Wales, UK
3. University College London, London, UK
4. Narodowy Instytut Zdrowia Publicznego, Warsaw, Poland
5. University of South Wales, Pontypridd, Wales, UK
6. Public Health Wales, Health Protection Division, Cardiff, Wales, UK

Correspondence: Daniel Rh Thomas (daniel.thomas@wales.nhs.uk)

Keywords: Covid-19; ethnicity; outcomes; epidemiology

## Abstract

### Objective

To identify ethnic differences in proportion positive for SARS-CoV-2, and proportion hospitalised, proportion admitted to intensive care, and proportion died in hospital with COVID-19 during the first epidemic wave in Wales.

### Design

Descriptive analysis of 76,503 SARS-CoV-2 tests carried out in Wales to 31 May 2020. Cohort study of 4,046 individuals hospitalised with confirmed COVID-19 between 1st March and 31st May. In both analyses, ethnicity was assigned using a name-based classifier.

### Setting

Wales (UK)

### Primary and secondary outcomes

Admission to an intensive care unit following hospitalisation with a positive SARS-CoV-2 PCR test. Death within 28 days of a positive SARS-CoV-2 PCR test.

### Results

Using a name-based ethnicity classifier, we found a higher proportion of Black, Asian and ethnic minority people tested for SARS-CoV-2 by PCR tested positive, compared to those classified as White. Hospitalised Black, Asian and minority ethnic cases were younger (median age 53 compared to 76 years; $p<0.01$) and more likely to be admitted to intensive care. Bangladeshi (adjusted odds ratio: 9.80, 95%CI 1.21- 79.40) and 'White – Other than British or Irish' (aOR: 1.99, 95%CI: 1.15- 3.44) ethnic groups were most likely to be admitted to ICU. In Wales, older age (aOR for over 70 years: 10.29, 95%CI: 6.78–15.64) and male gender (aOR: 1.38, 95%CI: 1.19–1.59), but not ethnicity, were associated with death in hospitalised patients.

### Conclusions

This study adds to the growing evidence that ethnic minorities are disproportionately affected by COVID-19. During the first COVID-19 epidemic

wave in Wales, although ethnic minority populations were less likely to be tested and less likely to be hospitalised, those that did attend hospital were younger and more likely to be admitted to intensive care. Primary, secondary and tertiary COVID-19 prevention should target ethnic minority communities in Wales.

## Strengths and limitations of this study

- Secondary analysis of data obtained through routine national COVID-19 surveillance.
- Studies relying on clinician reported ethnicity contain high proportions of missing and poor quality data.
- Using a proven name-based classifier, we were able to assign ethnicity to nearly all participants.
- Whilst sensitivity and specificity of the classifier varies in specific ethnic groups, and is poor in black British and people of mixed ethnicity, its performance is quantifiable and classification bias can be taken into account when interpreting findings.
- Age, gender and deprivation were taken into account in the analysis, but individual data on history of chronic disease was poorly recorded, and treatment histories once hospitalised were not available.

## Introduction

There is growing evidence that Black, Asian and other minority ethnic (BME) people living in Europe are at increased risk of infection with SARS-CoV-2 and, if infected, are more likely to have severe disease.[1] In the United Kingdom, the Intensive Care National Audit and Research Centre first raised concerns that BME people were over-represented amongst COVID-19 patients admitted to intensive care.[2] These findings were reported widely in the media and discussed in opinion pieces. [3]–[7] In Wales, the First Minister established an advisory group to examine the issue and provide recommendations to reduce ethnic inequality in COVID-19 outcomes. [8] Whilst focusing on COVID-19, this group has recognised the underlying inequalities Black, Asian and Minority Ethnic people experience in their lives, which are likely to have impacted in ethnic disparities in COVID-19.

Investigating ethnic health inequalities is hampered by the poor recording of ethnicity in clinical data. This is the case for COVID-19 notifications and laboratory reports in Wales. In order to rapidly investigate ethnic variation in COVID-19 epidemiology, we applied Onomap, a name-based ethnicity classification tool developed by the Department of Geography at University College London that has been found effective in 30 other published studies in healthcare, epidemiology and public health [9]. This was applied to routinely collected, named COVID-19 laboratory test data, held by Public Health Wales Communicable Disease Surveillance Centre.

## Methods

### Participants

We obtained routine surveillance data on 77,555 SARS-CoV-2 PCR tests carried out by Public Health Wales and authorised as at 1300 hrs, 31 May 2020 from Microbiology Datastore, a repository of test results recorded in the all-Wales Laboratory Information Management System.

Data were also obtained on records of 4,046 hospitalised patients (people admitted to hospital within 14 days of a positive SARS-CoV-2 test or individuals who tested positive for SARS-CoV-2 whilst in hospital) as at 1700 hrs, 31 May 2020 available in IC-Net, an infection prevention and control information management system. These data contained information on whether an individual was admitted to intensive care (ICU).

These individual person data on hospitalised cases were linked to records of 1,309 Covid-19 in-hospital deaths (Covid-19 cases who died in hospital, and had a positive test result for SARS-CoV-2 28 days or less prior to the date of death or 7 days after death) reported to Public Health Wales' Covid-19 mortality surveillance scheme to 1700hrs, 31 May, as at 28 June 2020.

### Ethnicity

Ethnicity was categorised using the name-based ethnicity classifier, Onomap, a software tool developed by geographers at University College London, and the 2001 Census classification of ethnicity.[10] We collapsed the Census categories further into: 'White British or Irish', 'White Other', 'Asian or British Asian', 'Black or Black British', 'other ethnicity' and 'unclassified', with a further aggregation to create a 'BME' field, containing all ethnicities other than 'White British', 'White Irish', or 'White Other'. Unclassified observations were excluded.

### Deprivation

Small (Lower Super Output) areas in Wales were assigned a deprivation score using the Welsh Index of Multiple Deprivation[11] and areas were ordered into quintiles based on the distribution of these scores, ranging from least to most

deprived. Each individual was then assigned to a deprivation quintile based on their Lower Super Output Areas of residence.

**Statistical analysis**

Proportions of population tested, with 95% confidence intervals, were calculated for White and BME groups using population data from the most recent Office for National Statistics Labour Force Survey.[12] Proportion testing positive with 95% confidence intervals were calculated by dividing number positive by number tested for the same time period. Using logistic regression we calculated odds ratios for testing positive for ethnic groups, after adjusting for age, sex and deprivation quintile.

Using the cohort of 4,046 hospitalised patients, we carried out a logistic regression to calculate odds ratios for the outcomes: (a) admitted to intensive care and (b) mortality, with 95% confidence intervals, for ethnic groups, in each case using 'White British or Irish' ethnicity as the baseline comparator. Independent variables were gender and age group. Multivariable analyses were then used to calculate odds ratios for ethnic groups whilst controlling for gender, age group and local area deprivation. Differences in the distributions of previously reported risk factors for fatal outcomes (age, gender, deprivation medical history) [13], [14] were investigated further in White and BME groups. The Mann Whitney two-sample test was used to compare differences in the age distribution of BME and White deaths. Odds ratios with 95% confidence intervals were calculated to compare proportion male and proportion with underlying health conditions amongst deceased BAME and White individuals. All analysis was carried out using Stata 14. [15]

**Validation of Onomap's performance**

The performance of Onomap was assessed using three data sets containing reliable self-reported or healthcare professional-reported ethnicity. These data were: A list of people attending a mosque in Wales who were offered screening for hepatitis C (n=189), a list of tuberculosis patients notified by doctors in Wales (n=3267) and a list of patients attending an infectious disease clinic in Poland (n=3184). Using these data as a 'gold standard', sensitivity and

specificity were calculated to measure Onomap's performance in correctly classifying specific ethnicities.

## Ethical and privacy considerations

Ethical oversight of the project was provided by Public Health Wales NHS Trust R&D Division. As this work was carried out as part of the health protection response to a public health emergency in Wales, using routinely collected surveillance data, Public Health Wales R&D Division advised that NHS research ethics approval was not required. The use of named patient data in the investigation of communicable disease outbreaks and surveillance of notifiable disease is permitted under Public Health Wales' Establishment Order. Data were held and processed under Public Health Wales' information governance arrangements, in compliance with the Data Protection Act, Caldicott Principles and Public Health Wales guidance on the release of small numbers. No data identifying protected characteristics of an individual were released outside Public Health Wales. Validation work was from a project that had previously received permission from the Confidentiality Advisory Group to process patient data on tests for viral hepatitis carried out by laboratories in Wales, and research ethics approval from West of Scotland REC4 (Application title: Incidence of infectious disease in BME groups using Onomap; CAG reference: 16/CAG/0133; IRAS project ID: 210327 REC reference: 16/WS/018).

## Patient and Public Involvement statement

Patients or the public were involved in the design and conduct of our research and the work has been shared with the Welsh Government BAME COVID Advisory group, which contains community and stakeholder groups on a number of occasions. This research has also been presented to the Race Council Cymru.

## Funding

No external funding was sought. The study was done with existing Public Health Wales resources.

## Results

### Ethnicity classification

Onomap estimated the ethnicity of 98.1% (13,789/14,054) of tested individuals, 99.2% (4,013/4,046) of those hospitalised, 97.4% (305/313) of those admitted to intensive care, and 99.6% (1,304/1,309) of those who died following admission to hospital.

### Testing and hospitalisation

By classifying ethnicity using names, we estimate that 3.7% (n=2,896) tests were of Black, Asian and other minority (BME) individuals. Using the most recent Statistics Wales population estimates for ethnic groups in Wales,[12] this represents 1,580 tests per 100, 000 population in BME, compared to 2,512 tests per 100,000 population in White ethnic groups.

Whilst White groups were more likely to be tested for SARS-CoV-2, BME groups were more likely to test positive. Of 14,054 people tested positive for SARS-CoV-2 in Wales to 31 May 2020, Onomap classified 13,092 in White ethnic groups and 697 in BME groups. Ethnic groups most likely to test positive were: Chinese, Indian, Pakistani, Asian-Other and White-Other (Figure 1).

Of all those testing positive, a smaller proportion (15.1%) of those tested in the BME group attended hospital compared to the White group (29.9%: see Table 1). However, the trend was reversed in people aged 50 to 59 years: 23.8% of positive BME individuals aged 50-59 years attended hospital, compared to 16.3% of White individuals testing positive. The median age of hospitalised BME individuals was 53 years compared to 76 years for White individuals ($p<0.01$; Mann Whitney 2 sample test).

### Admission to intensive care

Of those attending hospital, a much higher proportion (21.9%) of BME individuals were admitted to intensive care compared to White individuals (7.2%). Proportions of hospitalised patients admitted to intensive care (ICU)

were highest amongst the 'Asian and British Asian - Indian, Pakistani and Bangladeshi' (29.0%) and 'White – other' (20.8%) groups. The median age of BME patients admitted to ICU was 51 years compared to 58 years for White individuals ($p<0.01$; Mann Whitney 2 sample test). Amongst hospitalised patients aged between 50-59 years, 27.6% of BME patients were admitted to ICU compared to 21% of White patients. More patients died in hospital without being admitted to ICU. Of all those attending hospital, 10.5% of patients identified as BME died compared to 33.1% of White patients (Table 1).

We successfully linked all records of 4,046 people hospitalised with Covid-19, those admitted to ICU, and those who died in hospital, all as at 31 May 2020, using NHS numbers. Intensive care was more likely in hospitalised males (aOR: 2.03, 95% CI: 1.55-2.65) and in younger patients (Table 2, Figure 2). When specific ethnicities were examined, being admitted to ICU was more likely in 'White Other', 'Asian and British Asian – Bangladeshi', 'Asian and British Asian – Indian', and 'Asian and British Asian – Pakistani' ethnic groups. After adjusting for gender, age and social deprivation, 'White Other' (aOR: 1.99, 95% CI: 1.15-3.44), and 'Bangladeshi' (aOR: 9.8, 95%CI: 1.21-79.40), ethnic groups remained significantly more likely to be admitted to ICU.

Hospitalised cases living in the most deprived areas of Wales were significantly more likely to attend ICU (OR: 1.70; 95% CI 1.17-2.45). However, this effect did not remain significant after adjusting for age, sex and ethnicity (aOR: 1.37, 95% CI: 0.93-2.02).

**Mortality**

Likelihood of dying was significantly higher for hospitalised males. This effect remained after adjusting for age and ethnicity (aOR: 1.38, 95%CI: 1.19-1.59) (Table 2, Figure 2). No increase was observed in risk of death with increasing deprivation. There was a strong association between increasing age and death from Covid-19 which remained after adjusting for gender, ethnicity and social deprivation (aOR for aged 70 years and over: 10.29 (95% CI: 6.78-15.64). However, there was no evidence from this study that BME groups were more likely to die from Covid-19 than White-British or Irish groups, even after adjusting for gender, age and social deprivation (Table 2).

To investigate further, we compared the differences in the distribution of previously reported risk factors for fatal outcome[13] in White and BME groups who had died. BME people who died in Wales with Covid-19 were younger than White people who died (BME median age 71 compared to 79 for White people; p=0.06, Mann-Whitney 2 sample test). Underlying chronic disease was recorded for 50% of deaths. There was a higher proportion of BME people 72.7 (95% CI: 39.0-94.0) that had a history of underlying chronic disease compared to White people 49.6 (95% CI: 46.8- 52.3).

**Validation of Onomap**

Onomap returned predicted ethnicity for 97% of the 6640 names in the four data sets. Sensitivity and specificity was calculated for each ethnic group. Onomap generally had a high specificity, that is: it was unlikely to return a false ethnicity in people self-reporting a given ethnicity (Table 3). Specificity was 77% for white ethnicities, indicating that a proportion of people in BME groups will be misclassified as white. In terms of its sensitivity, that is its likelihood of detecting all people self-reporting as an ethnic group, Onomap was poor for some ethnic groups, most notably for those self-reporting as black or British black. In other words, many people self-reporting as black or black British will be misclassified, most likely as white.

## Discussion

This was a rapid analysis of routinely available national surveillance data using name-based ethnicity classification software, carried out in response to the emerging epidemic in Wales. It adds to the evidence of increased risk of severe Covid-19 outcomes in ethnic minorities in Western Europe and the United States.

BME people living in Wales were less likely to be tested for SARS-CoV-2 in the first COVID-19 wave, but of those tested, people in Chinese, Indian, Pakistani and White-Other groups were more likely to test positive. It should be noted that testing in the first wave was mainly of people hospitalised and those working in health and social care, so trends in testing and in proportion positive need to be interpreted with caution.

We found risk of severe Covid-19, indicated by admission to ICU, to be higher in many ethnic minorities living in Wales, and significantly higher in those of Bangladeshi ethnicity and in White ethnic groups, other than British or Irish. Bangladeshi communities have been identified in other studies as being at particular risk of the effects of Covid-19.[16] The second group we identified, 'White-other', will contain a range of nationalities, but in Wales, recent migrants from Eastern Europe will comprise a significant proportion. The risk associated with the latter group has not been previously reported, and is an important finding given recent outbreaks reported in factory settings in Wales where many European migrants are employed. That the White-other group is at increased risk of severe COVID-19 gives weight to the hypothesis that ethnic disparities are socio-economic in basis.

The finding that certain minority ethnic groups are at higher risk of being admitted to intensive care, but are no more likely to die than the White British and Irish group, was also found in the CO-CIN cohort study involving 23,577 Covid-19 patients attending hospitals in the UK.[16] This slightly counterintuitive finding may be a genuine finding or may be the result of classification bias. Firstly, if genuine, differences in the age distribution of cases in White and BME groups are likely to be a factor. During early 2020, Covid-19 mortality was

observed overwhelmingly in the elderly, with White men over 70 years disproportionally affected. These patients may have been less likely to have been admitted to ICU for treatment. Black, Asian and minority ethnic populations in Wales are generally younger [16], and lower median ages were observed in hospitalised BME individuals. The finding that despite being generally younger, BME individuals were more likely to be admitted to ICU is an important finding.

The lack of increased risk of mortality is at odds with some studies from England that have found that Covid-19 deaths have been disproportionally high in BAME groups. Of course, it is always possible that Wales, with a more deprived general population relative to England, a lower density of BME people and smaller urban conurbations, presents a less unequal risk setting. On the other hand, there may be methodological issues affecting this finding. It is likely that Onomap underestimates the absolute number of BAME individuals, particularly for Black groups. The misclassification of Black as White may have led to an under estimation of relative risk. It is also possible that date of onset to death is longer in younger people and that our study did not take sufficient account of this. The absence of individual-level data on comorbidities is a limitation of this study. Further work is currently being carried out using linkage of COVID-19 notification data with other routine health records to further understand risks associated with hospitalisation. Also, we only had access to deaths that occurred in hospital. It is possible that there may have been ethnic differences in the proportion of people dying outside of hospital.

Onomap has been used widely as a tool in public health, for example in studies investigating variation in influenza mortality,[17] hepatitis B infection [18] and HPV vaccination uptake.[19] However, Onomap has limitations, and all findings should be interpreted in light of these. We previously validated the tool using data containing self-reported or healthcare professional-reported ethnicity. Onomap performs well for most ethnicities, but has a low sensitivity for Black or Black British individuals. Risks identified for Black and Black British groups are therefore likely to be underestimated. Kandt and Longley have published a comparison of Onomap with self-reported ethnicity in the 2011 Census.[20] Notwithstanding apparent success in 30 reported studies in public health,

healthcare and epidemiology (and wider application in equity audits in, *inter alia*, housing allocation, management science and social media), the reliability and limitations of such methods should be acknowledged and understood. In the absence of good ethnicity recording in routine health records, it does facilitate scientific investigation with margins of error that are understood. Moreover, many of the existing studies where individual person ethnicity is available have missing data, and are not without their own classification bias. Anecdotally, members of minority ethnic groups are more likely to defer from reporting their ethnicities, and clinician-based classification is understood to be unreliable.

One of the recommendations of the Welsh Government Advisory Group is to improve recording of ethnicity in routine health data, and a data improvement plan is urgently required so ethnicity can be included in routine public health surveillance. There is an urgent need for all European countries carrying out Covid-19 surveillance to report trends by ethnicity, in order to inform local infection prevention and control policy and practice. Ethnic variation should also be considered in the design of interventions, and in crisis communication.

Wales is less ethnically diverse than many other areas of the United Kingdom, but its BME population has increased in recent years. In 2001, the Census recorded 2.1% of the population as BME. This increased to 4.4% in the 2011 Census. Most recent estimates from the Labour Force Survey indicates that this has grown to 5.5%.[12] The Welsh Government has established an Advisory Group to investigate issues around Covid-19 in BME groups and has published a series of recommendations. In Wales, an occupational risk assessment tool has been developed with the aim of reducing risk of infection in those most vulnerable to severe infection.[21] This tool, developed initially for the health care sector, is for all ethnicities, but includes a weighting to account for the emerging evidence of increased risk in BME individuals. A recent report by the Race Disparity Unit in England [22] provides a summary of the actions being undertaken In England to reduce ethnic variation in COVID-19, including community engagement initiatives, economic support for work sectors that over-represent minority ethnic groups, and asymptomatic testing pilots. Comparing first and early second wave data, early analysis provides evidence that disparities appear to have improved for some ethnic groups including Black

Africans, Black Caribbean, Chinese and Indians but have worsened for Pakistanis and Bangladeshis. [23,24] In England, as a result of the findings from the QCOVID risk model,[14] the list of people shielding has been updated, using a new predictive risk model which combines factors including ethnicity, and the postcode where people live and its link with deprivation.

COVID-19 is now a vaccine preventable disease, and vaccination is being rolled out across the UK. There are concerns that vaccination uptake may be lowest in areas with high numbers of minority ethnic populations. Office for National Statistics report that from early December 2020 to early January 2021, less than half (49%) of Black or Black British adults reported that they were likely to have the vaccine. [25] The latest OpenSAFELY data reports that approximately 60% of black people over 70 have been vaccinated compared to 75% for South Asians and 90% of white people. [26] Initiatives are being undertaken to improve vaccine uptake in ethnic minority groups in Wales, and latest data indicate that progress is being made in reducing variation. [27]

This is a complex topic and it is still unclear whether ethnic variation in poor outcomes is the result of higher incidence of infection or greater severity of disease. Minority ethnic groups are more likely to live in urban areas, to have public facing jobs, are more likely to live in crowded housing and live in multi-generational households.[28] Further research is needed to quantify risk of SARS-CoV-2 infection and risk of severe outcomes in ethnic minority communities, and better understand the underlying processes behind any disparities. However, there is now enough evidence to act, and effort should be focussed on continuing to design innovative interventions for primary, secondary and tertiary prevention of Covid-19 in minority ethnic groups.[29]

## Acknowledgements

Onomap was purchased from Publicprofiler Ltd. The authors acknowledge the many laboratory and surveillance staff in Public Health Wales involved in developing and maintaining routine Covid-19 surveillance in Wales. Victoria McClure assisted with extracting data from IC-Net.

### Conflict of interest

Paul Longley is Director of Publicprofiler Ltd.

### Authors' contributions

Daniel Thomas designed the study, contributed to the analysis and wrote the manuscript. Oghogho Orife contributed to the analysis and commented on the manuscript. Amy Plimmer, George Karani, Meirion Evans, Janusz Janiec and Roiyah Saltus commented on the manuscript and contributed to the validation work. Paul Longley commented on the methodology, referee comments and results, including use of the names classification tool. Chris Williams and Giri Shankar commented on the design and analysis and manuscript.

### Exclusive Licence

affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

**Data availability**

Data are available upon reasonable request- Individual data not available but aggregated data may be made available.

Figure 1. Determinants of having a positive SARS-CoV-2 PCR test. Adjusted odds ratios (aOR) with 95% confidence intervals are given for male gender, compared to female, older age groups compared to those aged less than 50 years, small area deprivation quintile comparing with least deprived, and Onomap estimated ethnicities, compared to 'White British'. Odds ratios greater than one represent an increased risk; odds ratios less than one represent a decreased risk. 95% confidence intervals not crossing one reflect that the odds ratio is statistically significant.

Figure 2. Determinants of: 1. Being admitted to intensive care unit (ICU); and 2. In-hospital mortality in 4,046 individuals hospitalised with Covid-19 in Wales to 31 May 2020, as at 28 June 2020. Adjusted odds ratios (aOR) with 95% confidence intervals are given for male gender, compared to female, older age groups compared to those aged less than 50 years, small area deprivation quintile comparing with least deprived, and Onomap estimated ethnicities, compared to 'White British'. Odds ratios greater than one represent an increased risk; odds ratios less than one represent a decreased risk. 95% confidence intervals not crossing one reflect that the odds ratio is statistically significant.

## References


[1] D. Pan *et al.*, "The impact of ethnicity on clinical outcomes in COVID-19: A systematic review," *EClinicalMedicine*, vol. 23, p. 100404, Jun. 2020, doi: 10.1016/j.eclinm.2020.100404.

[2] "ICNARC – Reports," 2020. [Online]. Available: https://www.icnarc.org/Our-Audit/Audits/Cmp/Reports. [Accessed: 14-Jun-2020].

[3] Guardian, "Ethnic minorities dying of Covid-19 at higher rate, analysis shows | World news | The Guardian," 2020.

[4] R. Cifuentes, "All in it together? The impact of Coronavirus on BAME people in Wales. | Bevan Foundation," 2020.

[5] M. Pareek *et al.*, "Ethnicity and COVID-19: an urgent public health research priority," *The Lancet*, vol. 395, no. 10234. Lancet Publishing Group, pp. 1421–1422, 02-May-2020, doi: 10.1016/S0140-6736(20)30922-3.

[6] BMA. Press Release, "Review into COVID-19 impact on BAME communities must be backed by real-time data and include measures to address problem now, says BMA," 2020. [Online]. Available: https://www.bma.org.uk/bma-media-centre/review-into-covid-19-impact-on-bame-communities-must-be-backed-by-real-time-data-and-include-measures-to-address-problem-now-says-bma. [Accessed: 14-Jun-2020].

[7] NHS Confederation, "The impact of COVID-19 on BME communities and health and care staff," 2020. [Online]. Available: https://www.nhsconfed.org/resources/2020/04/the-impact-of-covid19-on-bme-communities-and-staff. [Accessed: 14-Jun-2020].

[8] Welsh Government, "Wales BAME Covid-19 health advisory group takes a cross-Government approach | GOV.WALES," 2020.

[9] "Public Profiler," 2020. [Online]. Available: https://www.publicprofiler.org/. [Accessed: 14-Jun-2020].

[10] Office for National Statistics, "[ARCHIVED CONTENT] Ethnic Group Statistics: a guide for the collection and classification of ethnicity data: Office for National Statistics," 2003.

[11] StatsWales, "Welsh Index of Multiple Deprivation." [Online]. Available: https://statswales.gov.wales/Catalogue/Community-Safety-and-Social-Inclusion/Welsh-Index-of-Multiple-Deprivation. [Accessed: 14-Jun-2020].

[12] StatsWales, "Ethnicity by area and ethnic group." [Online]. Available: https://statswales.gov.wales/Catalogue/Equality-and-Diversity/Ethnicity/ethnicity-by-area-ethnicgroup. [Accessed: 14-Jun-2020].

[13] E. Williamson *et al.*, "Factors associated with COVID-19-related death using OpenSAFELY" *Nature* 584, 430–436 (2020). https://doi.org/10.1038/s41586-020-2521-4 .

[14] A. K. Clift *et al.*, "Living risk prediction algorithm (QCOVID) for risk of hospital admission and mortality from coronavirus 19 in adults: national derivation and validation cohort study," *BMJ*, vol. 371, Oct. 2020, doi: 10.1136/bmj.m3731.

[15] StataCorp LLC, "Stata 14 | Stata." [Online]. Available: https://www.stata.com/stata14/. [Accessed: 14-Jun-2020].

[16] E. Harrison, A. Docherty, C. Semple, and CO-CIN, "CO-CIN: Investigating associations between ethnicity and outcome from COVID-19 - report to SAGE, 25 April 2020 - GOV.UK," 2020. [Online]. Available: https://www.gov.uk/government/publications/co-cin-investigating-associations-between-ethnicity-and-outcome-from-covid-19-report-to-sage-25-april-2020. [Accessed: 14-Jun-2020].

[17] H. Zhao, R. J. Harris, J. Ellis, and R. G. Pebody, "Ethnicity, deprivation and mortality due to 2009 pandemic influenza A(H1N1) in England during the 2009/2010 pandemic and the first post-pandemic season," *Epidemiol. Infect.*, vol. 143, no. 16, pp. 3375–3383, Dec. 2015, doi: 10.1017/S0950268815000576.

[18] M. Binka *et al.*, "Differing profiles of people diagnosed with acute and chronic hepatitis B virus infection in British Columbia, Canada," *World J. Gastroenterol.*, vol. 24, no. 11, pp. 1216–1227, Mar. 2018, doi: 10.3748/wjg.v24.i11.1216.

[19] K. G. Pollock *et al.*, "Evidence of decreased HPV vaccine acceptance in Polish communities within Scotland," *Vaccine*, vol. 37, no. 5, pp. 690–692, Jan. 2019, doi: 10.1016/j.vaccine.2018.10.097.

[20] J. Kandt and P. A. Longley, "Ethnicity estimation using family naming practices," *PLoS One*, vol. 13, no. 8, p. e0201774, Aug. 2018, doi: 10.1371/journal.pone.0201774.

[21] Welsh Government, "COVID-19 workforce risk assessment tool | GOV.WALES," 2020. [Online]. Available: https://gov.wales/covid-19-workforce-risk-assessment-tool. [Accessed: 14-Jun-2020].

[22] Race Disparity Unit, "Second quarterly report on progress to address COVID-19 health inequalities- GOV.UK," 2021. [Online]. Available: https://www.gov.uk/government/publications/second-quarterly-report-on-progress-to-address-covid-19-health-inequalities. [Accessed: 28-Mar-2021].

[23] V. Nafilyan *et al.*, "Ethnic differences in COVID-19 mortality during the first two waves of the Coronavirus Pandemic: a nationwide cohort study of 29 million adults in England" *medRxiv*, p. 2021.02.03.21251004, Feb. 2021, doi: 10.1101/2021.02.03.21251004

[24] K. Bhaskaran *et al.*, "Short report: Ethnicity and COVID-19 death in the early part of the COVID-19 second wave in England: an analysis of OpenSAFELY data from 1st September to 9th November 2020" *medRxiv*, p. 2021.02.02.21250989, Feb. 2021, doi: 10.1101/2021.02.02.21250989

[25] Office for National Statistics, "Coronavirus and the social impacts on Great Britain: 29 January 2021," 2021. [Online]. Available: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthandwellbeing/bulletins/coronavirusandthesocialimpactsongreatbritain/29january2021#attitudes-to-covid-19-vaccination-by-different-sub-groups-of-the-population. [Accessed: 30-Jan-2021].

[26] OPENSAFELY, "NHS COVID-19 Vaccine Coverage Report," [Online]. Available: https://opensafely.org/research/2021/covid-vaccine-coverage/#weekly-report [Accessed: 15-Feb-2021]

[27] Public Health Wales, "Rapid COVID-19 Surveillance Dashboard" [Online]. Available: https://public.tableau.com/profile/public.health.wales.health.protection#!/vizhome/RapidCOVID-19virology-Public/Headlinesummary [Accessed: 28-Mar-2021]

[28] Office for National Statistics, "2011 Census General Report - Office for National Statistics," 2011. [Online]. Available: https://www.ons.gov.uk/census/2011census/howourcensusworks/howdidwedoin2011/2011censusgeneralreport. [Accessed: 14-Jun-2020].

[29] Z. Haque, L. Becares, and N. Treloar, "A Runnymede Trust and ICM Survey Over-Exposed and Under-Protected The Devastating Impact of COVID-19 on Black and Minority Ethnic Communities in Great Britain Runnymede: Intelligence for a Multi-ethnic Britain," 2020.

**Table 1. Ethnicity breakdown of individuals tested for SARS-CoV-2 in Wales and proportions hospitalised, attending intensive care units (ICU), and deceased**

| Ethnicity[1] | Positive tests | Hospital-ised | % hospitalised (95%CI) | | Attending ICU | % hospitalised attending ICU (95% CI) | | Deceased citing Covid-19 | % hospitalised and deceased citing Covid-19 (95%CI) | |
|---|---|---|---|---|---|---|---|---|---|---|
| All ethnicities | 14,054 | 4,046 | 28.8 | (28.0-29.5) | 313 | 7.7 | (6.9-8.6) | 1,309 | 32.4 | (30.9-33.8) |
| White British and Irish | 12,565 | 3812 | 30.3 | (29.5-31.2) | 262 | 6.9 | (6.1-7.7) | 1,274 | 33.4 | (31.9-34.9) |
| White – other | 527 | 96 | 18.2 | (15.0-21.8) | 20 | 20.8 | (13.2-30.3) | 19 | 19.8 | (12.4-29.2) |
| **All white ethnicities** | **13,092** | **3908** | **29.9** | **(29.1-30.6)** | **282** | **7.2** | **(6.4-8.1)** | **1,293** | **33.1** | **(31.6-34.6)** |
| Asian and British Asian (Indian, Pakistani, Bangladeshi) | 371 | 62 | 16.7 | (13.1-20.9) | 18 | 29.0 | (18.1-41.9) | <10 | - | |
| Black and Black British | 54 | <10 | - | | <10 | - | | <10 | - | |
| Chinese | 40 | <10 | - | | <10 | - | | <10 | - | |
| Other | 232 | 30 | 12.9 | (8.9-17.9) | <10 | - | | <10 | - | |
| **All non-white ethnicities** | **697** | **105** | **15.1** | **(12.5-17.9)** | **23** | **21.9** | **(14.4-31.0)** | **11** | **10.5** | **(5.3-18.0)** |
| Unknown | 265 | 12 | 4.5 | (2.4-7.7) | <10 | 66.7 | (34.9-90.1) | <10 | 41.7 | (15.2-72.3) |

[1] Onomap estimates of ethnicity

**Table 2. Personal characteristics associated with severe outcomes for Welsh residents hospitalised with Covid-19, to 31 May 2020**

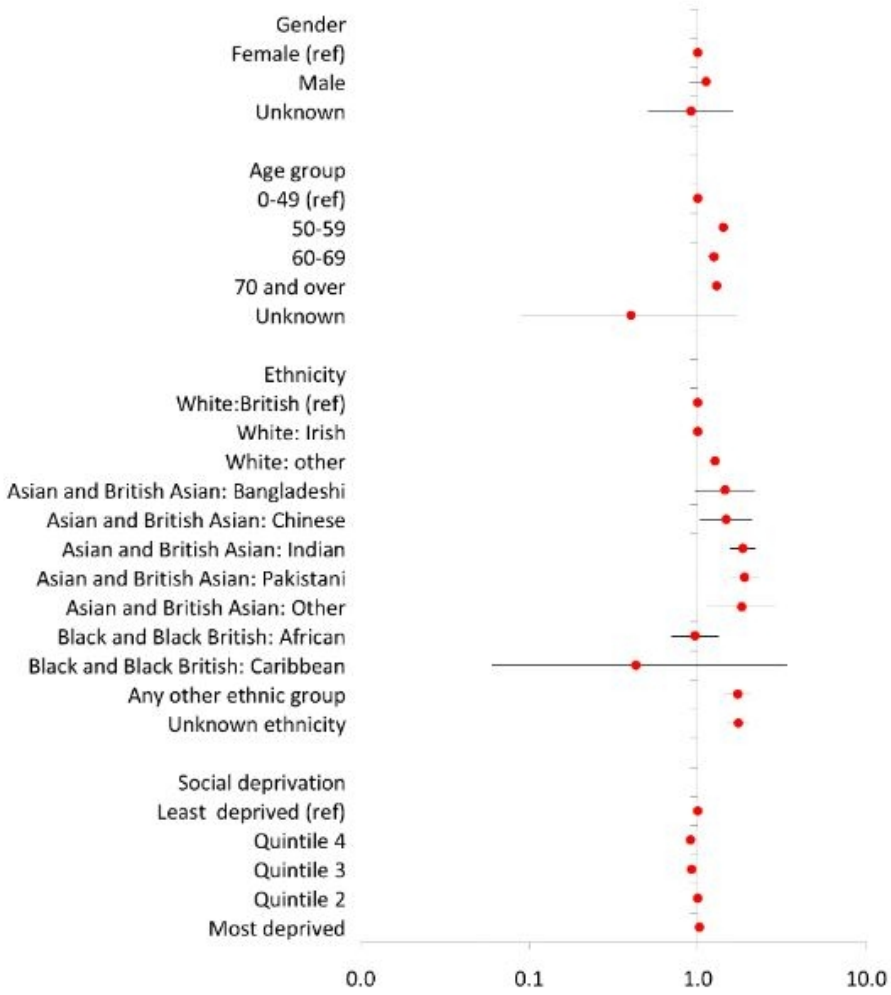
| Personal characteristic | | Hospitalised | Received intensive care | | | | Died | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Univariate analysis | | Multivariate analysis | | Univariate analysis | | Multivariate analysis | |
| Gender | | | OR | (95%CI) | OR | (95%CI) | OR | (95%CI) | OR | (95%CI) |
| | Female | 1920 | - | | - | | - | | - | |
| | Male | 2126 | **1.90** | **(1.49-2.43)** | **2.03** | **(1.55-2.65)** | **1.28** | **(1.12-1.47)** | **1.38** | **(1.19-1.59)** |
| Age | | | | | | | | | | |
| | 0-49 | 428 | - | | - | | - | | - | |
| | 50-59 | 466 | 1.24 | (0.89-1.72) | 1.27 | (0.89-1.80) | **2.43** | **(1.49-3.95)** | **2.29** | **(1.40-3.74)** |
| | 60-69 | 569 | **0.64** | **(0.45-0.91)** | **0.64** | **(0.45-0.93)** | **5.93** | **(3.80-9.25)** | **5.11** | **(3.26-8.03)** |
| | 70 and over | 2583 | **0.11** | **(0.07-0.15)** | **0.11** | **(0.08-0.17)** | **11.40** | **(7.55-17.20)** | **10.29** | **(6.78-15.64)** |
| Ethnicity | | | | | | | | | | |
| | White: British | 3,716 | - | | - | | - | | - | |
| | White: Irish | 96 | 0.43 | (0.14-1.37) | 0.25 | (0.06-1.06) | 0.73 | (0.47-1.16) | 0.68 | (0.42-1.10) |
| | White: other | 96 | **3.51** | **(2.11-5.84)** | **1.99** | **(1.15-3.44)** | **0.49** | **(0.29- 0.81)** | 0.82 | (0.46-1.44) |
| | Asian and British Asian: Bangladeshi | <10 | **8.89** | **(1.48-53.49)** | **9.8** | **(1.21-79.40)** | 0.49 | (0.06-4.43) | 0.61 | (0.06-6.29) |
| | Asian and British Asian: Chinese | <10 | 2.67 | (0.31-22.93) | 1.78 | (0.19-16.85) | 0.40 | (0.05-3.39) | 0.97 | (0.10-9.36) |
| | Asian and British Asian: Indian | 16 | **6.07** | **(2.09-17.59)** | 2.49 | (0.83-7.53) | 0.46 | (0.13-1.60) | 1.06 | (0.27-4.14) |
| | Asian and British Asian: Pakistani | 41 | **4.89** | **(2.42-9.88)** | 1.91 | (0.89-4.09) | **0.16** | **(0.05-0.51)** | 0.44 | (0.13-1.50) |
| | Black and Black British: African | <10 | 2.22 | (0.27-18.55) | 0.93 | (0.10-8.59) | 0.33 | (0.04-2.74) | - | - |
| | Any other ethnic group | 30 | 1.54 | (0.46-5.12) | 0.59 | (0.17-2.01) | **0.15** | **(0.03-0.62)** | 0.25 | (0.06-1.12) |
| | Unknown ethnicity | 33 | **4.27** | **(1.91-9.56)** | **2.99** | **(1.25-7.18)** | **0.35** | **(0.14-0.92)** | 0.59 | (0.22-1.61) |
| Deprivation | | | | | | | | | | |
| | Most deprived | 939 | **1.70** | **(1.17-2.45)** | **1.37** | (0.93-2.02) | 0.95 | (0.77-1.16) | 1.10 | (0.88-1.36) |
| | Quintile 2 | 867 | 1.25 | (0.85-1.84) | 1.03 | (0.69-1.56) | 0.93 | (0.76-1.15) | 1.06 | (0.85-1.32) |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Quintile 3 | 682 | 1.08 | (0.71-1.64) | 1.05 | (0.68-1.65) | 1.01 | (0.81-1.26) | 1.03 | (0.82-1.30) |
| | Quintile 4 | 658 | 1.09 | (0.72-1.67) | 1.00 | (0.64-1.58) | **0.68** | **(0.54-0.86)** | **0.69** | **(0.54-0.87)** |
| | Least Deprived | 732 | - | | - | | - | | - | |

Table 3. Validation of Onomap. Estimated sensitivity and specificity of Onomap by ethnic group. Calculated by measuring the performance of Onomap to predict ethnicity in three clinical data sets[1] already containing self-reported or healthcare professional-reported ethnicity

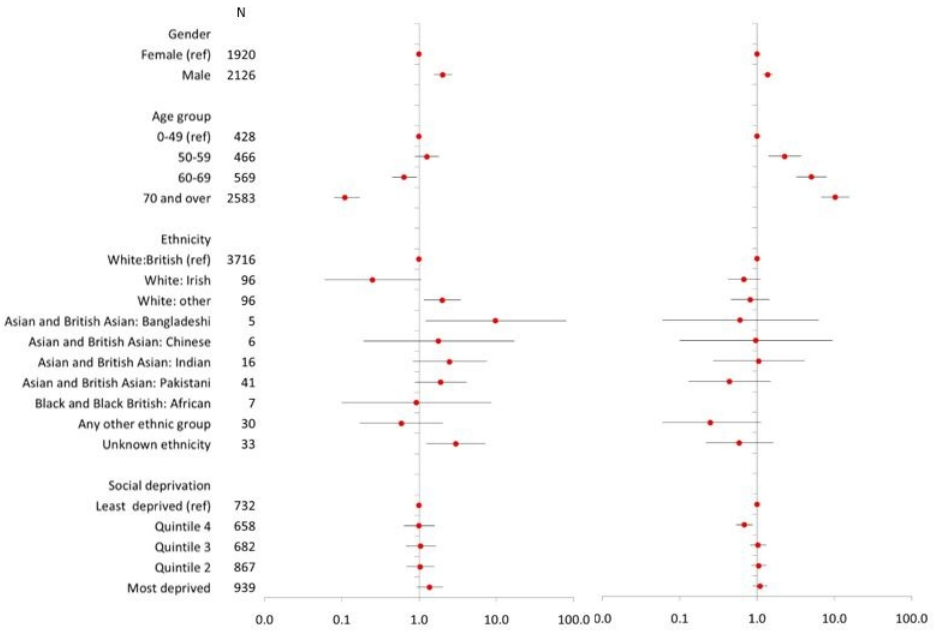
| Ethnicity | Ethnicity reported by participant | Ethnicity predicted by Onomap | Ethnicity correctly predicted | Sensitivity | Specificity |
|---|---|---|---|---|---|
| White British or Irish | 1681 | 1811 | | | |
| Other White | 3235 | 3418 | | | |
| **Total White** | 4916 | 5229 | 4844 | **98.5%** | **77.7%** |
| | | | | | |
| Indian | 364 | 239 | | | |
| Pakistani | 313 | 348 | | | |
| Bangladeshi | 96 | 88 | | | |
| Chinese | 55 | 18 | | | |
| Other Asian | <10 | 118 | | | |
| **Total Asian or Asian British** | 837 | 811 | 609 | **72.8%** | **96.5%** |
| | | | | | |
| Black- African | 344 | 142 | | | |
| Black - Caribbean | 10 | <10 | | | |
| Other Black | 23 | <10 | | | |
| **Total Black or Black British** | 377 | 143 | 112 | **29.7%** | **99.5%** |
| | | | | | |
| Arabic | 39 | 279 | | | |
| Other | <10 | <10 | | | |
| **Other Ethnic Group** | 45 | 283 | 24 | **53.3%** | **96.1%** |
| | | | | | |
| **Mixed** | 234 | <10 | - | | |
| | | | | | |
| **Unclassified/Unknown** | 231 | 174 | <10 | **3.5%** | **97.4%** |
| | | | | | |
| **Total** | 6640 | 6640 | **5589** | **87.4%** | **96.1%** |

[1] Three data sets which included self-reported or healthcare professional-reported ethnicity were used to validate Onomap: A list of individuals attending a mosque in Wales who were offered screening for hepatitis C (n=189), a list of tuberculosis patients notified by doctors in Wales (n=3267) and a list of patients attending an infectious disease clinic in Poland (n=3184).




Determinants of having a positive SARS-CoV-2 PCR test. Adjusted odds ratios (aOR) with 95% confidence intervals are given for male gender, compared to female, older age groups compared to those aged less than 50 years, small area deprivation quintile comparing with least deprived, and Onomap estimated ethnicities, compared to 'White British'. Odds ratios greater than one represent an increased risk; odds ratios less than one represent a decreased risk. 95% confidence intervals not crossing one reflect that the odds ratio is statistically significant.

156x159mm (96 x 96 DPI)




Determinants of: 1. Being admitted to intensive care unit (ICU); and 2. In-hospital mortality in 4,046 individuals hospitalised with Covid-19 in Wales to 31 May 2020, as at 28 June 2020. Adjusted odds ratios (aOR) with 95% confidence intervals are given for male gender, compared to female, older age groups compared to those aged less than 50 years, small area deprivation quintile comparing with least deprived, and Onomap estimated ethnicities, compared to 'White British'. Odds ratios greater than one represent an increased risk; odds ratios less than one represent a decreased risk. 95% confidence intervals not crossing one reflect that the odds ratio is statistically significant.

220x143mm (96 x 96 DPI)

STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

| | Item No | Recommendation | |
|---|---|---|---|
| **Title and abstract** | 1 | (*a*) Indicate the study's design with a commonly used term in the title or the abstract | ✓ |
| | | (*b*) Provide in the abstract an informative and balanced summary of what was done and what was found | ✓ |
| **Introduction** | | | |
| Background/rationale | 2 | Explain the scientific background and rationale for the investigation being reported | ✓ |
| Objectives | 3 | State specific objectives, including any prespecified hypotheses | ✓ |
| **Methods** | | | |
| Study design | 4 | Present key elements of study design early in the paper | ✓ |
| Setting | 5 | Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection | ✓ |
| Participants | 6 | (*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up | ✓ |
| | | (*b*) For matched studies, give matching criteria and number of exposed and unexposed | |
| Variables | 7 | Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable | ✓ |
| Data sources/ measurement | 8* | For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group | ✓ |
| Bias | 9 | Describe any efforts to address potential sources of bias | ✓ |
| Study size | 10 | Explain how the study size was arrived at | No sample taken. Used Welsh population |
| Quantitative variables | 11 | Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why | ✓ |
| Statistical methods | 12 | (*a*) Describe all statistical methods, including those used to control for confounding | ✓ |
| | | (*b*) Describe any methods used to examine subgroups and interactions | ✓ |
| | | (*c*) Explain how missing data were addressed | ✓ |
| | | (*d*) If applicable, explain how loss to follow-up was addressed | Not applicable |
| | | (*e*) Describe any sensitivity analyses | ✓ |
| **Results** | | | |
| Participants | 13* | (a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed | ✓ |
| | | (b) Give reasons for non-participation at each stage | ✓ |

| | | | |
|---|---|---|---|
| | | (c) Consider use of a flow diagram | ✓ |
| Descriptive data | 14* | (a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders | ✓ |
| | | (b) Indicate number of participants with missing data for each variable of interest | ✓ |
| | | (c) Summarise follow-up time (eg, average and total amount) | ✓ |
| Outcome data | 15* | Report numbers of outcome events or summary measures over time | ✓ |
| Main results | 16 | (*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included | ✓ |
| | | (*b*) Report category boundaries when continuous variables were categorized | ✓ |
| | | (*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period | ✓ |
| Other analyses | 17 | Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses | ✓ |
| **Discussion** | | | |
| Key results | 18 | Summarise key results with reference to study objectives | ✓ |
| Limitations | 19 | Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias | ✓ |
| Interpretation | 20 | Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence | ✓ |
| Generalisability | 21 | Discuss the generalisability (external validity) of the study results | ✓ |
| **Other information** | | | |
| Funding | 22 | Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based | ✓ |

*Give information separately for exposed and unexposed groups.

**Note:** An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.