

# Using data linkage to understand the impact of social and ethnic inequalities in health

24<sup>th</sup> Sept 2022

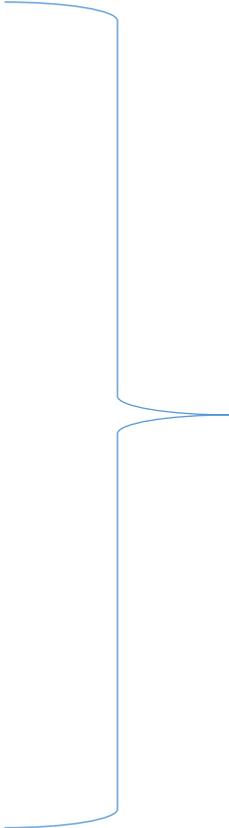
Professor Ronan Lyons

# Secure Anonymised Information Linkage (SAIL) Databank

- Built from the ground up as a safe, legal, social acceptable solution to data linkage and data sharing
- Underpinned by world-leading technical infrastructure (SeRP)
- “No data leaves” model, keeping data under control at all times
- Highly automated, highly efficient, highly secure, completely trustworthy
- Creating unsurpassed research-ready linked data resources for both research community and public sector analytical teams
- SAIL is not limited to Welsh data but supports many UK and international projects

# Privacy by design

- 1) Data sharing agreements and secure transfer
- 2) Reliable matching process
- 3) Anonymisation and encryption
- 4) Disclosure control and risk mitigation
- 5) Data access controls and agreements
- 6) Scrutiny of data use proposals and results
- 7) External verification of compliance with IG



Safe Projects  
Safe People  
Safe Data  
Safe Settings  
Safe Outputs

The 'Five Safes': a framework for planning, designing and evaluating data access solutions (Felix Ritchie)

# Stakeholder & Public Engagement

- The SAIL Consumer Panel
- SAIL Databank has a long standing Consumer Panel, established in 2011
- It actively recruits and is involved in all aspects of SAIL, from bids to IGRP to dissemination and impact



# Total population data linkage

- Tracking 3.2M pop with multiple datasets in privacy protecting environment with public engagement
- Population demographic spine, including **2011 Census** data and health, social care and education workforce
- Infection: antigen and antibody
- Viral subtypes
- Emergency department attendances
- Hospital outpatient and admissions
- Critical care dataset
- General practice
- All laboratory tests
- Deaths
- Longer term outcomes
- Vaccination
- Embedded trials and cohorts

Open access

Protocol

## BMJ Open Understanding and responding to COVID-19 in Wales: protocol for a privacy-protecting data platform for enhanced epidemiology and evaluation of interventions

Jane Lyons <sup>1</sup>, Ashley Akbari <sup>1</sup>, Fatemeh Torabi,<sup>1</sup> Gareth I Davies,<sup>1</sup> Laura North,<sup>1</sup> Rowena Griffiths,<sup>1</sup> Rowena Bailey,<sup>1</sup> Joseph Hollinghurst <sup>1</sup>, Richard Fry <sup>1</sup>, Samantha L Turner,<sup>1</sup> Daniel Thompson,<sup>1</sup> James Rafferty,<sup>1</sup> Amy Mizen,<sup>1</sup> Chris Orton <sup>1</sup>, Simon Thompson,<sup>1</sup> Lee Au-Yeung,<sup>1</sup> Lynsey Cross,<sup>1</sup> Mike B Gravenor,<sup>2</sup> Sinead Brophy,<sup>1</sup> Biagio Lucini,<sup>1</sup> Ann John <sup>1</sup>, Tamas Szakmany,<sup>3,4</sup> Jan Davies,<sup>5</sup> Chris Davies,<sup>5</sup> Daniel Rh Thomas,<sup>6</sup> Christopher Williams,<sup>6</sup> Chris Emmerson,<sup>6</sup> Simon Cottrell,<sup>6</sup> Thomas R Connor,<sup>7</sup> Chris Taylor,<sup>8</sup> Richard J Pugh,<sup>9</sup> Peter Diggle,<sup>10,11</sup> Gareth John,<sup>12</sup> Simon Scourfield,<sup>12</sup> Joe Hunt,<sup>12</sup> Anne M Cunningham,<sup>12</sup> Kathryn Helliwell,<sup>13</sup> Ronan Lyons <sup>1</sup>

# publication exemplars

medRxiv

THE PREPRINT SERVER FOR HEALTH SCIENCES



BMJ Yale

HOME | ABC

Search

[Comment on this paper](#)

## Evaluating the effect of COVID-19 on dispensing patterns: a national cohort analysis

[Fatemeh Torabi](#), [Ashley Akbari](#), [Laura North](#), [Daniel Harris](#), [Gareth Davies](#), [Mike Gravenor](#), [Rowena Griffiths](#), [Jane Lyons](#), [Neil Jenkins](#), [Andrew Morris](#), [Julian Halcox](#), [Ronan A. Lyons](#)

doi: <https://doi.org/10.1101/2021.02.15.21251552>

This article is a preprint and has not been peer-reviewed [what does this mean?]. It reports new medical research that has yet to be evaluated and so should not be used to guide clinical practice.

Abstract

Full Text

Info/History

Metrics

[Preview PDF](#)

## The impact of COVID-19 on adjusted mortality risk in care homes for older adults in Wales, UK: a retrospective population-based cohort study for mortality in 2016–2020

[Joe Hollinghurst](#), [Jane Lyons](#), [Richard Fry](#), [Ashley Akbari](#), [Mike Gravenor](#), [Alan Watkins](#), [Fiona Verity](#), [Ronan A Lyons](#)

*Age and Ageing*, Volume 50, Issue 1, January 2021, Pages 25–31,  
<https://doi.org/10.1093/ageing/afaa207>

Published: 19 September 2020 [Article history](#)

medRxiv

THE PREPRINT SERVER FOR HEALTH SCIENCES



BMJ Yale

HOME | ABC

Search

[Comment on this paper](#)

## Surgical activity in England and Wales during the COVID-19 pandemic: a nationwide observational cohort study

[T D Dobbs](#), [J A G Gibson](#), [A J Fowler](#), [T E Abbott](#), [T Shahid](#), [F Torabi](#), [R Griffiths](#), [R A Lyons](#), [R M Pearse](#), [I S Whitaker](#)

doi: <https://doi.org/10.1101/2021.02.27.21252593>

This article is a preprint and has not been peer-reviewed [what does this mean?]. It reports new medical research that has yet to be evaluated and so should not be used to guide clinical practice.

Abstract

Full Text

Info/History

Metrics

[Preview PDF](#)

Open access

Original research

BMJ Paediatrics Open

## Staff-pupil SARS-CoV-2 infection pathways in schools in Wales: a population-level linked data approach

[Daniel A Thompson](#), [Hoda Abbaszajani](#), [Richard Fry](#), [Emily Marchant](#), [Lucy Griffiths](#), [Ashley Akbari](#), [Joe Hollinghurst](#), [Laura North](#), [Jane Lyons](#), [Fatemeh Torabi](#), [Gareth Davies](#), [Mike B Gravenor](#), [Ronan A Lyons](#)

To cite: Thompson DA, Abbaszajani H, Fry R, et al. Staff-pupil SARS-CoV-2 infection pathways in schools in Wales: a population-level linked data approach. *BMJ Paediatrics Open* 2021;5:e001049. doi:10.1136/bmjpo-2021-001049

### ABSTRACT

**Background** Better understanding of the role that children and school staff play in the transmission of SARS-CoV-2 is essential to guide policy development on controlling infection while minimising disruption to children's education and well-being.

**Methods** Our national e-cohort (n=464531) study used anonymised linked data for pupils, staff and associated households linked via educational settings in Wales. We

### What is known about the subject?

- ▶ Evidence of the role schools play in the transmission of SARS-CoV-2 is limited.
- ▶ Higher positivity rates are observed in school staff compared to pupils.
- ▶ Lack of evidence on transmission pathways transmission into and within schools.



# PHIRI: EU Project Use Case A

---

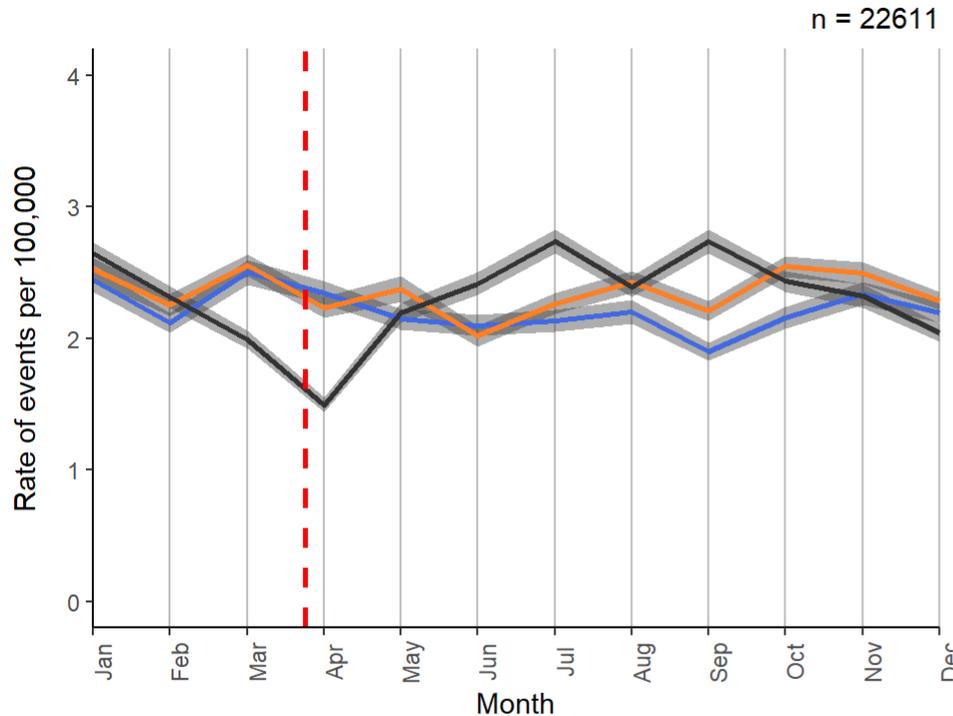
- Has the COVID19 pandemic changed existing patterns of non-COVID-19 health care utilisation for (vulnerable) populations within and between countries?
  - Heart attack and strokes (Cohort 1)
  - Hip and knee replacements (Cohort 2)
  - Serious trauma admissions (Cohort 3)
- Method: Compare age-standardised utilization rates for each month of 2020 (and possibly 2021) compared with pre-existing trends during 2017-2019, supplemented by ecological analyses and comparisons using data on infections and hospitalizations from ECDC.

## Objectives of the study:

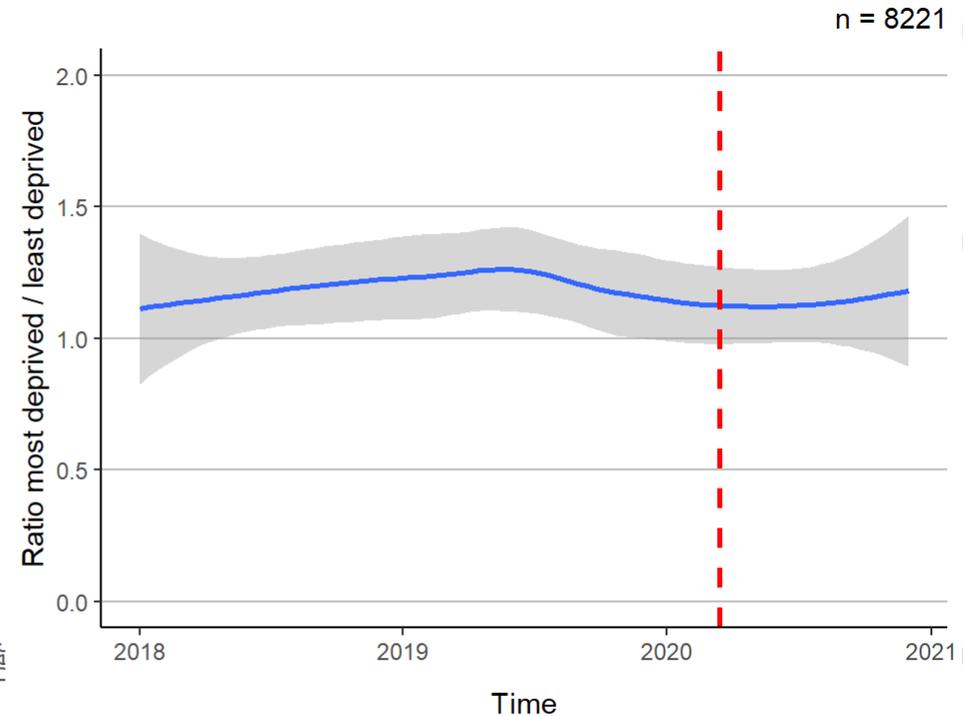
- ✓ Demonstrate how a broad variety of secondary data (e.g. administrative and survey data) can be pooled and/or reused in a distributed way across Europe
- ✓ Gain insights into the situation of socially (and potentially clinically) vulnerable groups during the COVID-19 pandemic
- ✓ Understand gaps in health system performance during crisis
- ✓ Develop learnings on system resilience and inclusive pandemic preparedness

# EASR for myocardial infarction and inequality ratio in Wales, UK

**C** Rate of heart attack in Wales



**D** Ratio of most to least deprived



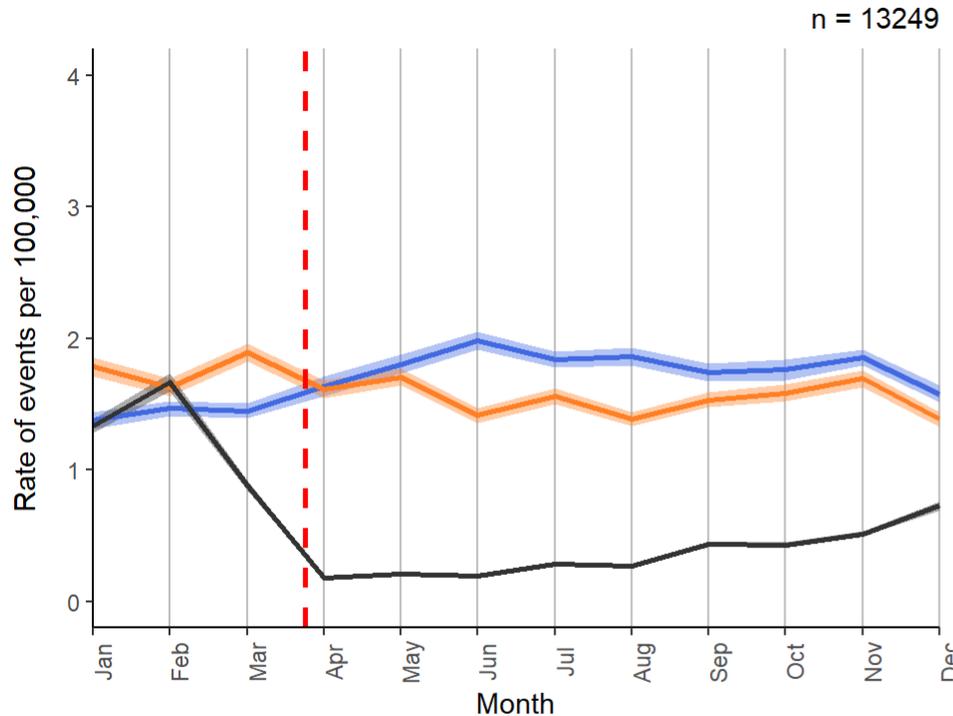
YEAR ■ 2018 ■ 2019 ■ 2020

The European age standardised rate of strokes (A) and heart attacks (C) occurring in Wales with the corresponding ratio of individuals with the highest deprivation score to those with the lowest (B and D)

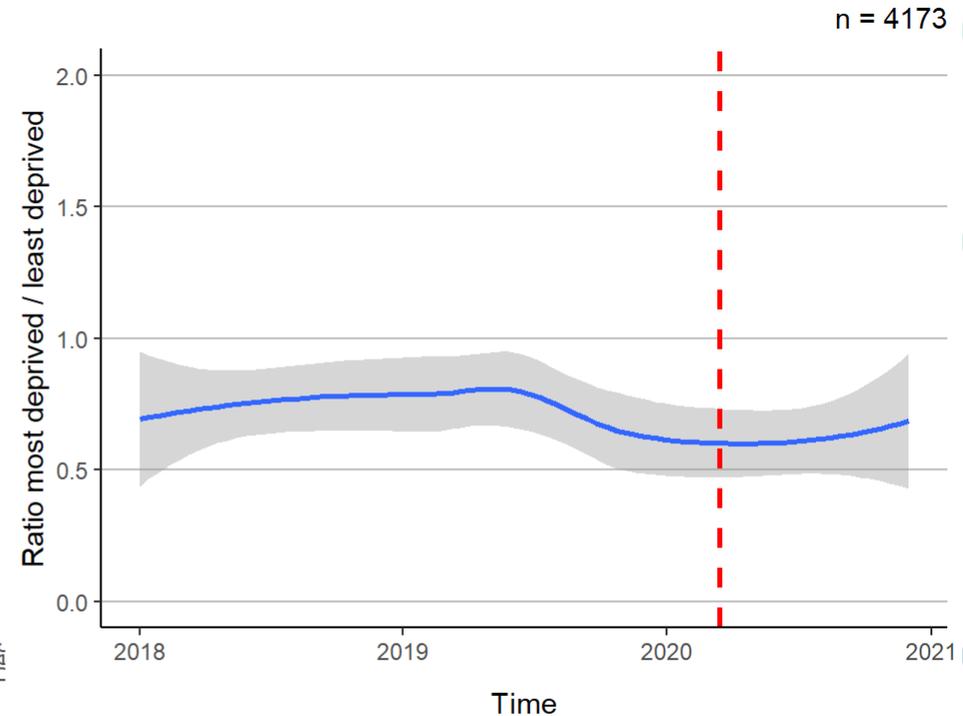


# EASR for hip replacements and inequality ratio in Wales, UK

**C** Rate of hip replacement surgeries in Wales



**D** Ratio of most to least deprived



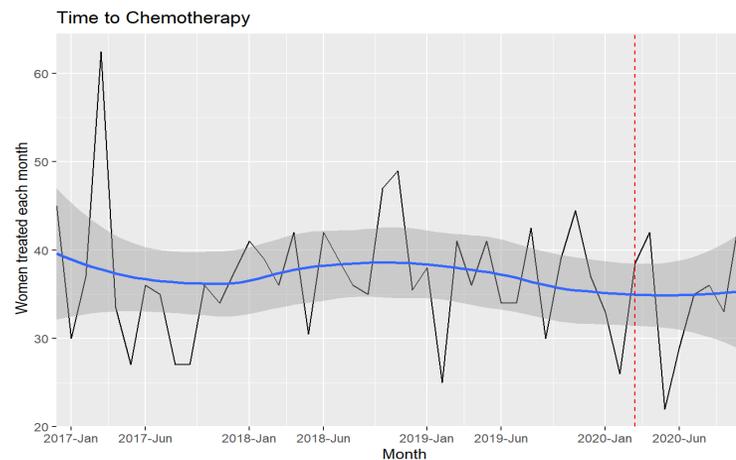
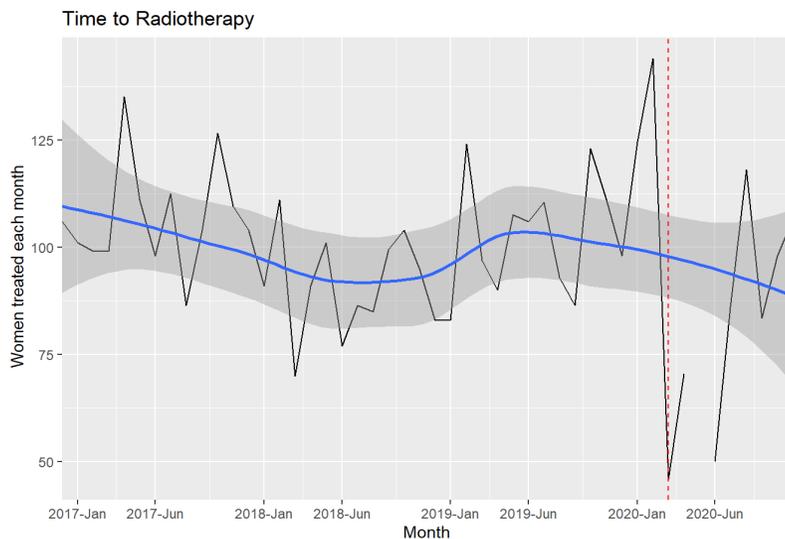
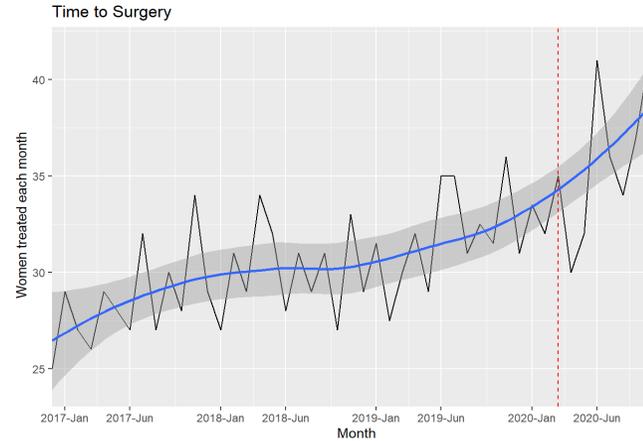
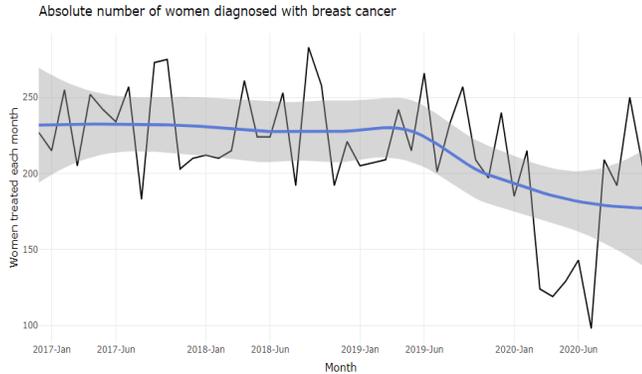
YEAR 2018 2019 2020

The European age standardised rate of elective knee replacement surgeries (A) and elective hip replacement surgeries (C) occurring in Wales with the corresponding ratio of individuals with the highest deprivation score to those with the lowest (B and D)



# Use Case B: Breast cancer during COVID-19: detection and treatment

Courtesy of Sarah Aldridge – PHIRI project [www.phiri.eu](http://www.phiri.eu)



03:25 📶 🔋

← Top Stories 📌

The Guardian

Black and Asian people wait longer for a cancer diagnosis than white people - with some waiting an extra six weeks, The Guardian reports.

Business (1) Politics (1) Science (1) Health (1) Environment (1) Sport (1) News (1) World (1)

**Revealed: black and Asian cancer patients wait longer for diagnosis**

**Pakistan's fatal floods 'a climate dystopia'**

**Will VAT cut help people pay bills?**

**Calling foul over water pollution**

**Crashland times: Any returns to 'Nothing Hill' viewers after Covid absence**

f t e

4h ago 22:47

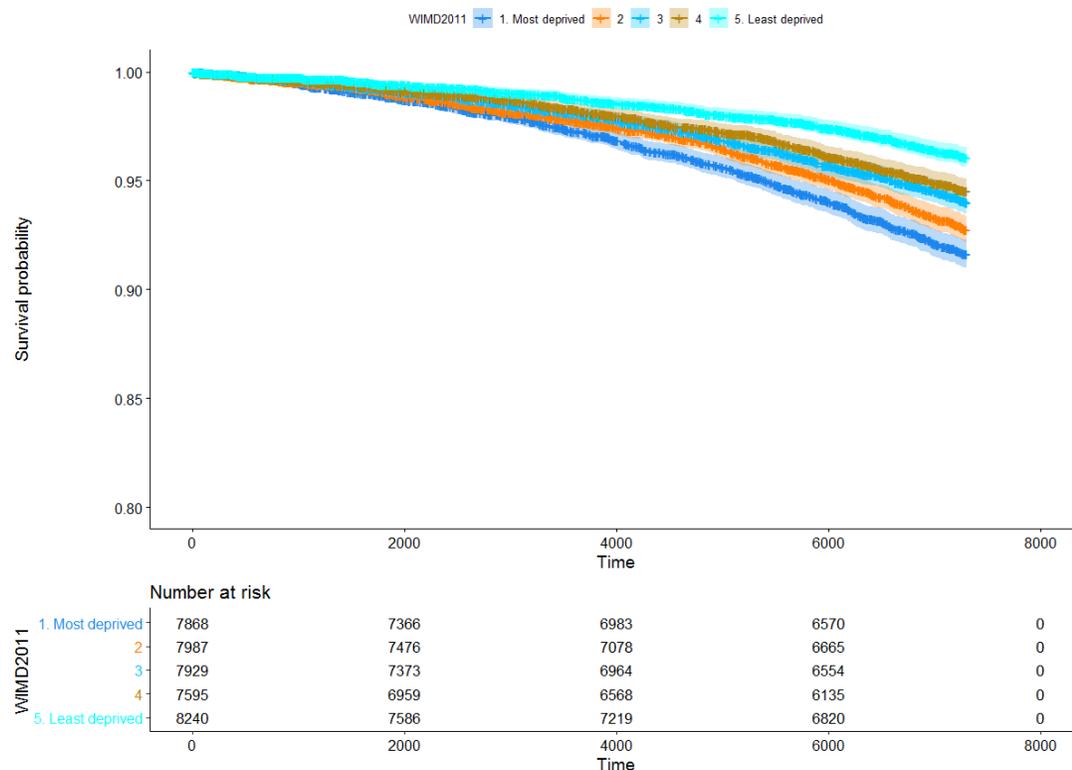
Daily Mail

More than 13,500 NHS beds are occupied by people who are well enough to go home - triple the pre-pandemic average, according to the Daily Mail

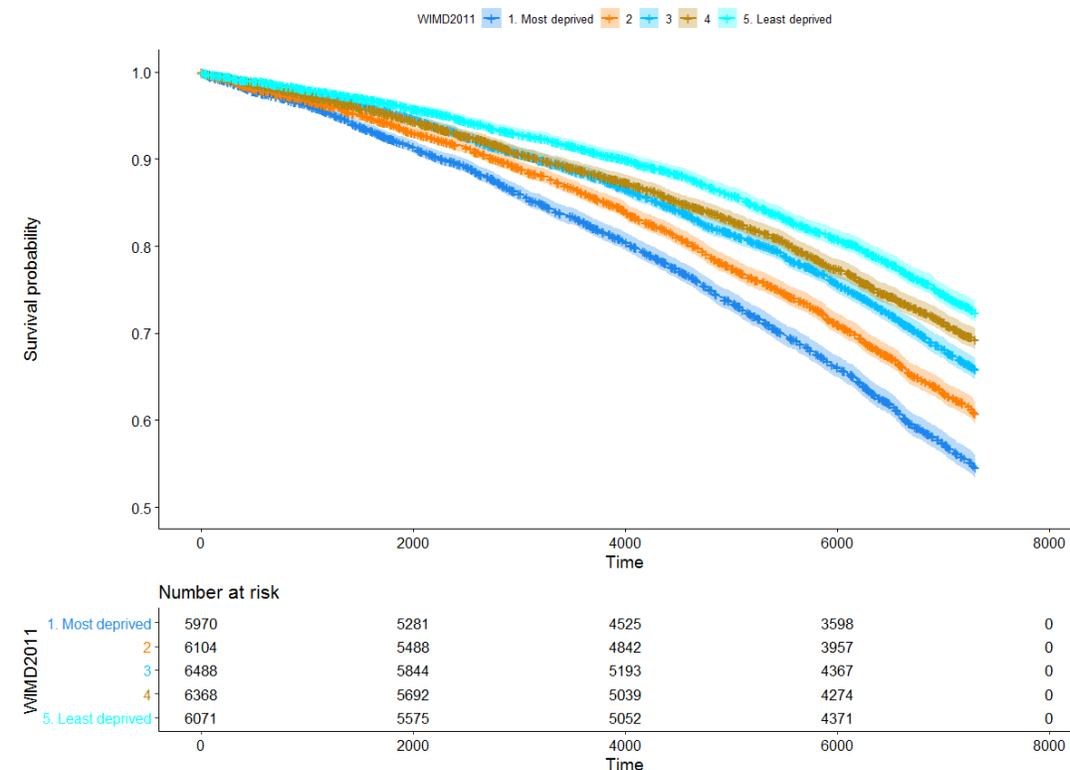
# Wales Multimorbidity Cohort: Survival by Deprivation

## Fifth in those in their 40s or 60s at inception

Kaplan-Meier Curve for WMC survival by WIMD2011 and aged 40 at inception



Kaplan-Meier Curve for WMC survival by WIMD2011 and aged 60 at inception

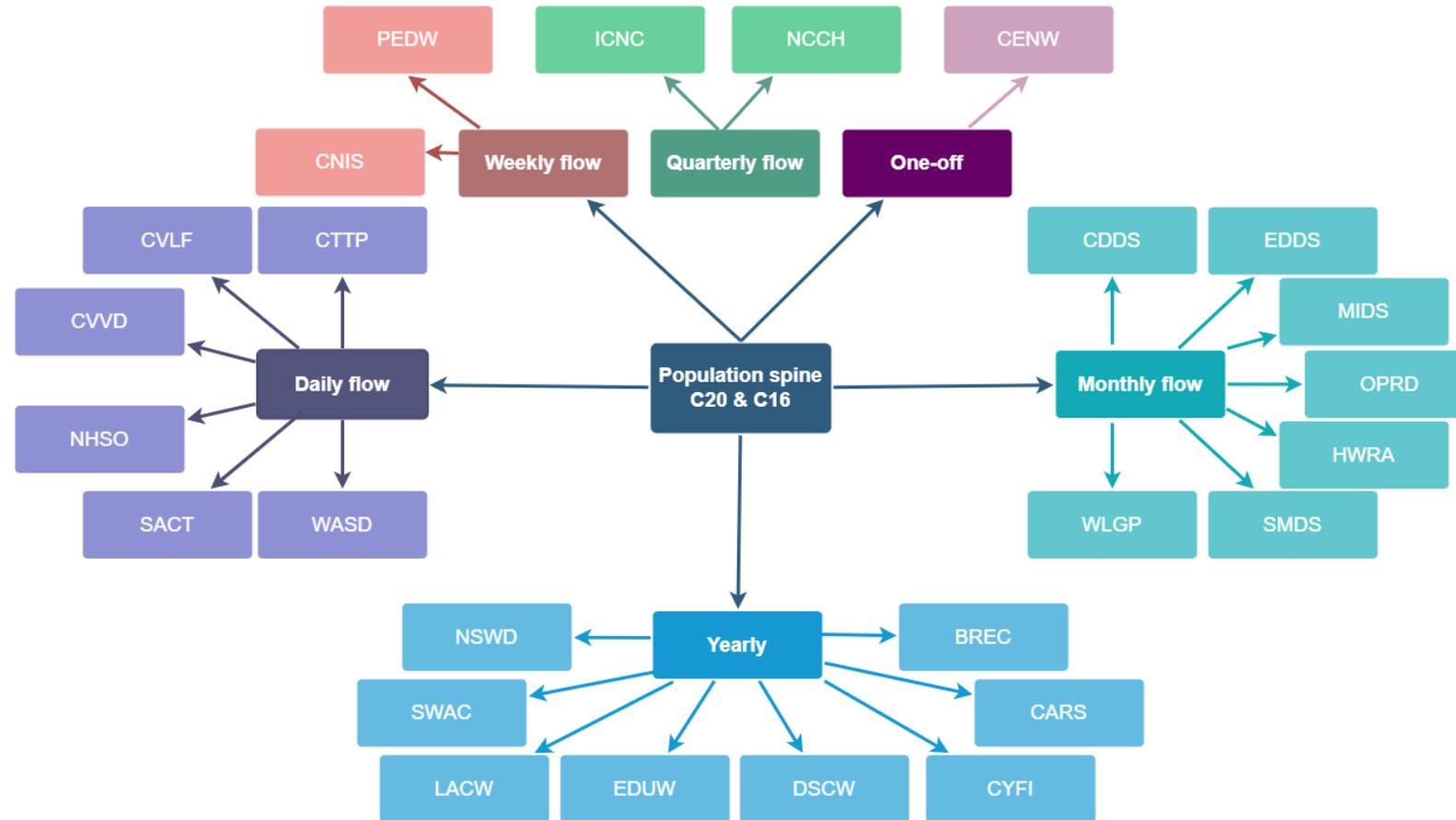


## Developing and comparing a population-scale linked data ethnicity-spine in Wales to assess COVID-19 outcomes

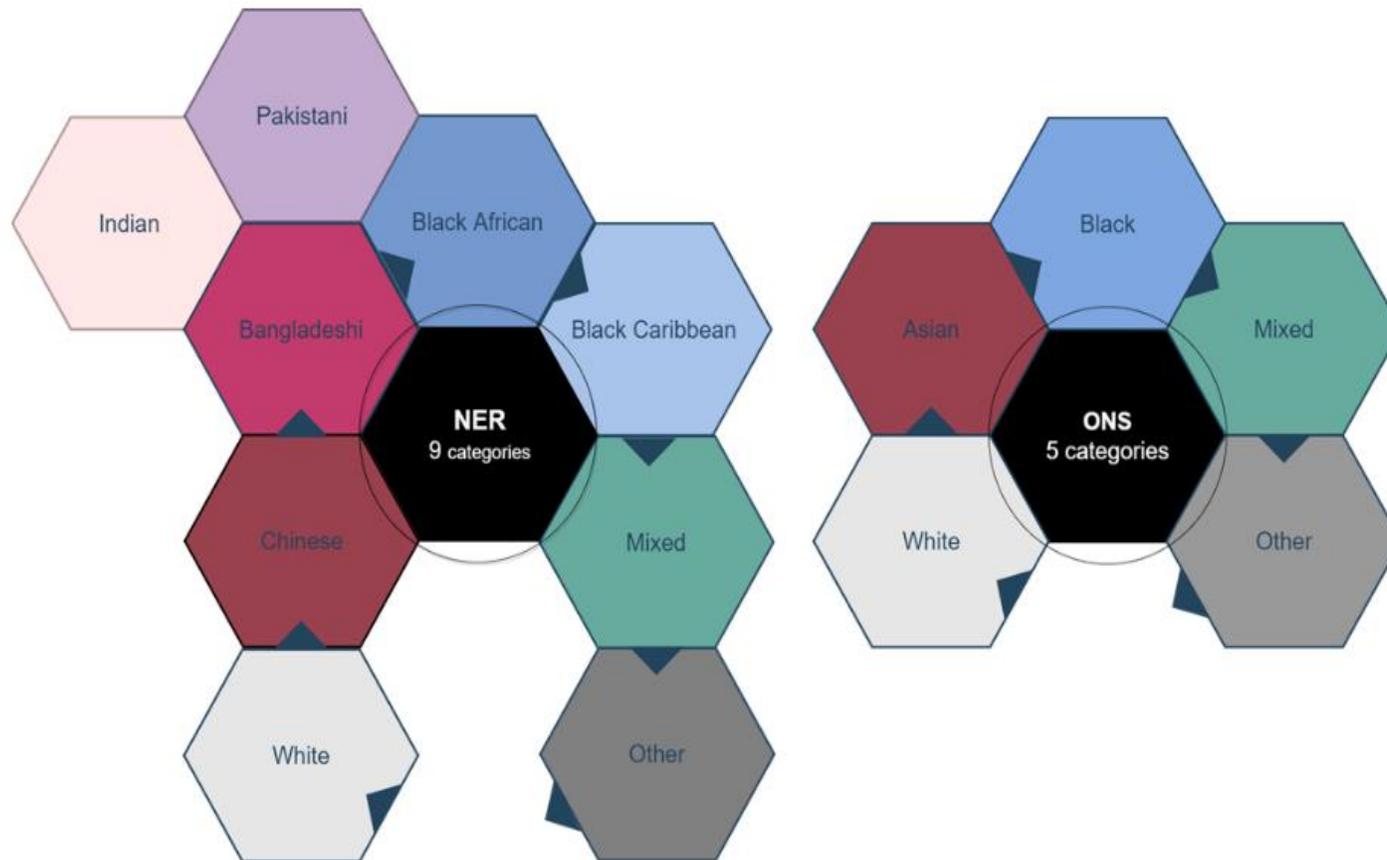
Ashley Akbari, Fatemeh Torabi, Stuart Bedston, Emily Lowthian , Hoda Abbasizanjani , Rich Fry , Jane Lyons, Rhiannon Owen, Kamlesh Khunti, Ronan A. Lyons

- 26 data sources within SAIL
- Derive a set of harmonised ethnic group spine for the population of Wales.
- Longitudinal, varied ethnicity values into a harmonised, de-duplicated lean research ready data asset (RRDA).
- Four different approaches: mode, **most recent (similar to population estimates)**, weighted mode, composite (if >1 then code to mixed)
- Compared the coverage and frequency of records for the entire population of Wales using an existing national-scale e-cohort (C20) across sex, age groups and major COVID-19 outcomes: infection, vaccination, hospitalisation and death.

# Datasets used



# Classifications of ethnicity – from COVID19 NERVTAG and ONS



# Ethnicity distribution – by NERVTAG groupings and method

- 46 million ethnicity codes found between 2000 and 2021
- across 26 data sources for 4,297,694 individuals
- 3,345,694 individuals who were alive and residing in Wales as identified in the C20 national-scale e-cohort for Wales on the 1st January 2020
- 4.4% had not had any associated ethnicity records in their health records

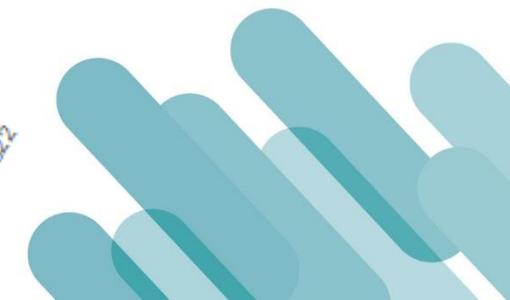
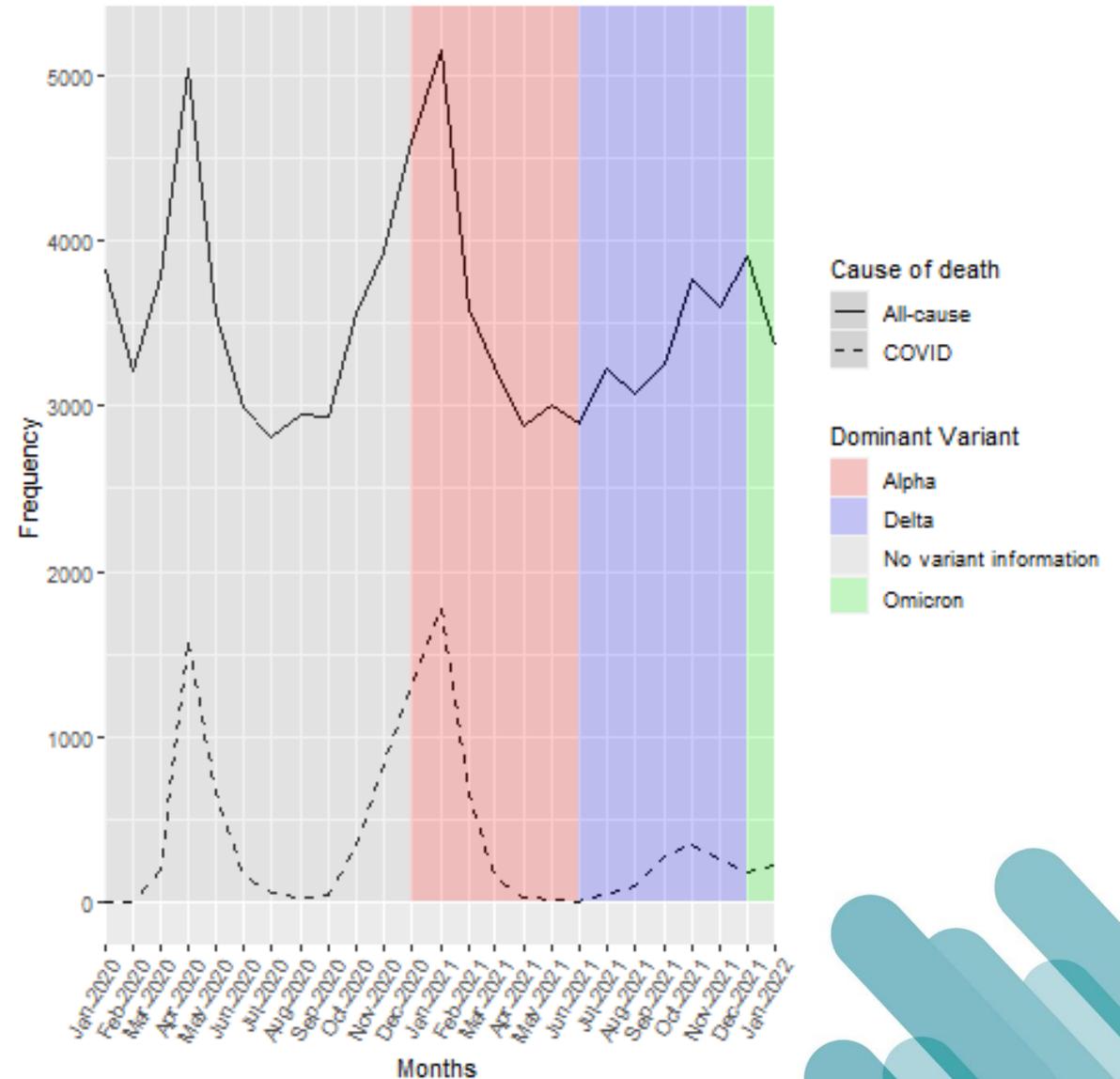
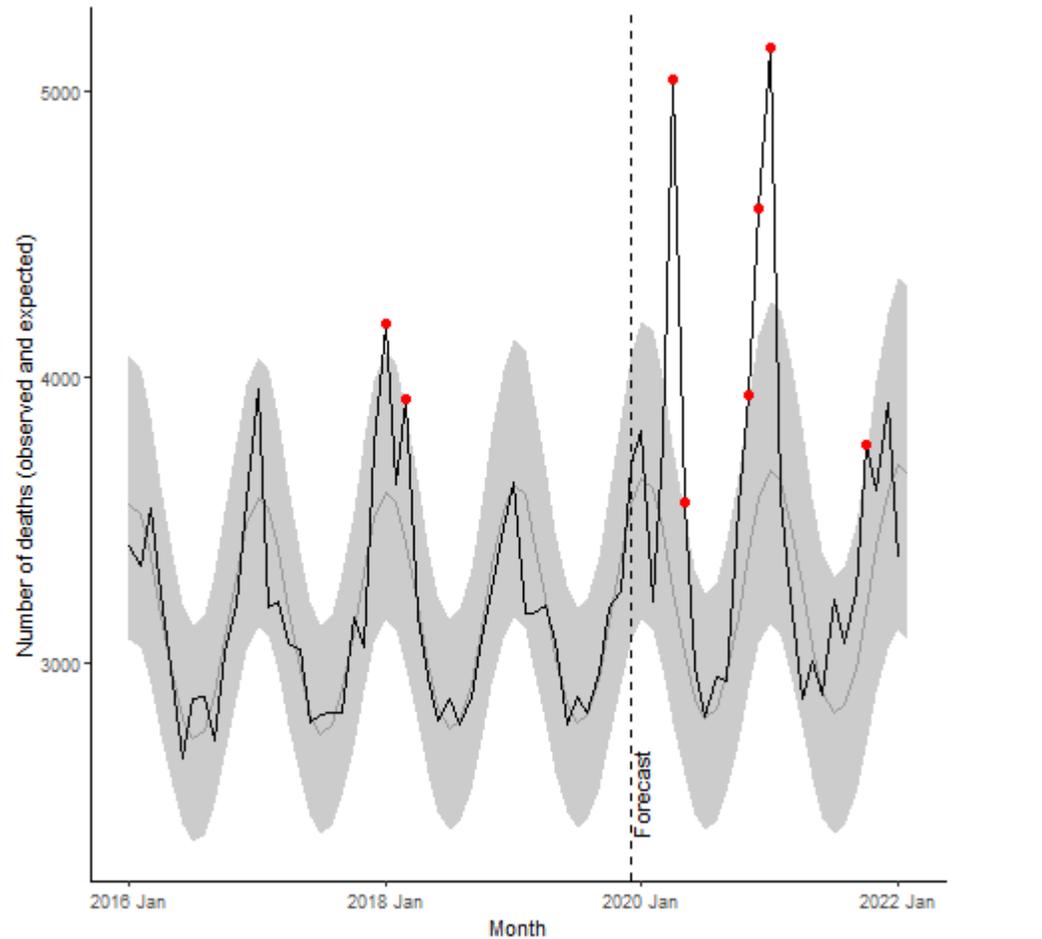
Characteristics	Latest		Mode		Weighted Mode		Composite	
	Ethnic group N = 3,457,694		Ethnic group N = 3,457,694		Ethnic group N = 3,457,694		Ethnic group N = 3,457,694	
	N(%)		N(%)		N(%)		N(%)	
<b>Ethnic Group (NER)</b>								
<b>White</b>	3,080,277	(89.0%)	3,093,817	(89.0%)	2,987,420	(86.0%)	2,987,245	(86.0%)
<b>Bangladeshi</b>	25,780	(0.7%)	16,243	(0.5%)	55,675	(1.6%)	9,204	(0.3%)
<b>Chinese</b>	17,970	(0.5%)	17,851	(0.5%)	20,302	(0.6%)	13,703	(0.4%)
<b>Indian</b>	25,058	(0.7%)	25,190	(0.7%)	27,078	(0.8%)	17,075	(0.5%)
<b>Pakistani</b>	17,035	(0.5%)	17,043	(0.5%)	20,287	(0.6%)	11,030	(0.3%)
<b>Black African</b>	21,578	(0.6%)	22,003	(0.6%)	32,671	(0.9%)	9,904	(0.3%)
<b>Black Caribbean</b>	4,730	(0.1%)	4,083	(0.1%)	7,145	(0.2%)	1,135	(<0.1%)
<b>Mixed</b>	43,672	(1.3%)	43,572	(1.3%)	66,895	(1.9%)	217,059	(6.3%)
<b>Other</b>	68,262	(2.0%)	64,560	(1.9%)	86,889	(2.5%)	38,007	(1.1%)
<b>Unknown</b>	153,332	(4.4%)	153,332	(4.4%)	153,332	(4.4%)	153,332	(4.4%)

# Statistical Methods for Covid-19: Mortality Statistics - Wales

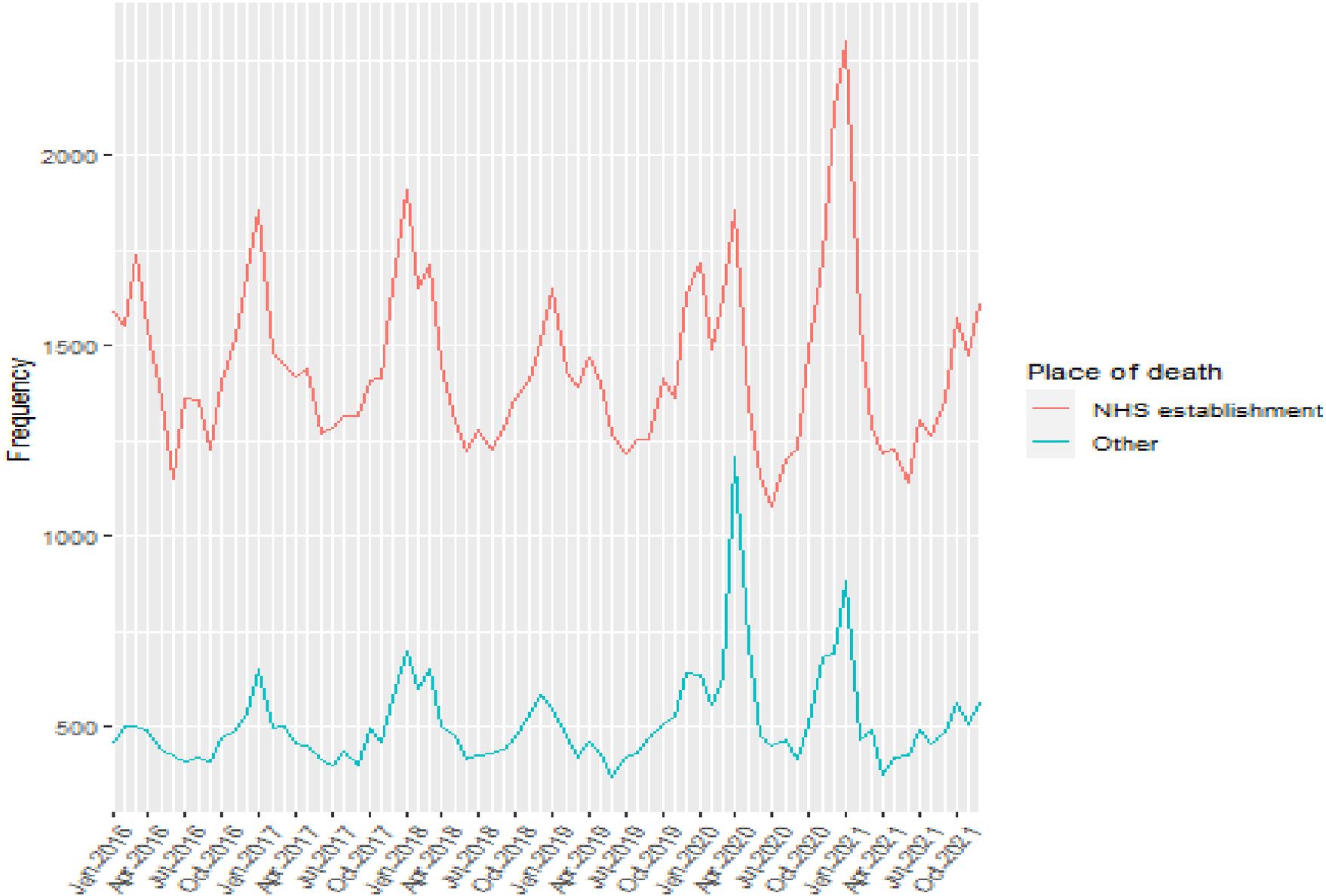
Ronan Lyons and Rhiannon Owen,  
Swansea University Medical School  
on behalf of the One Wales Team



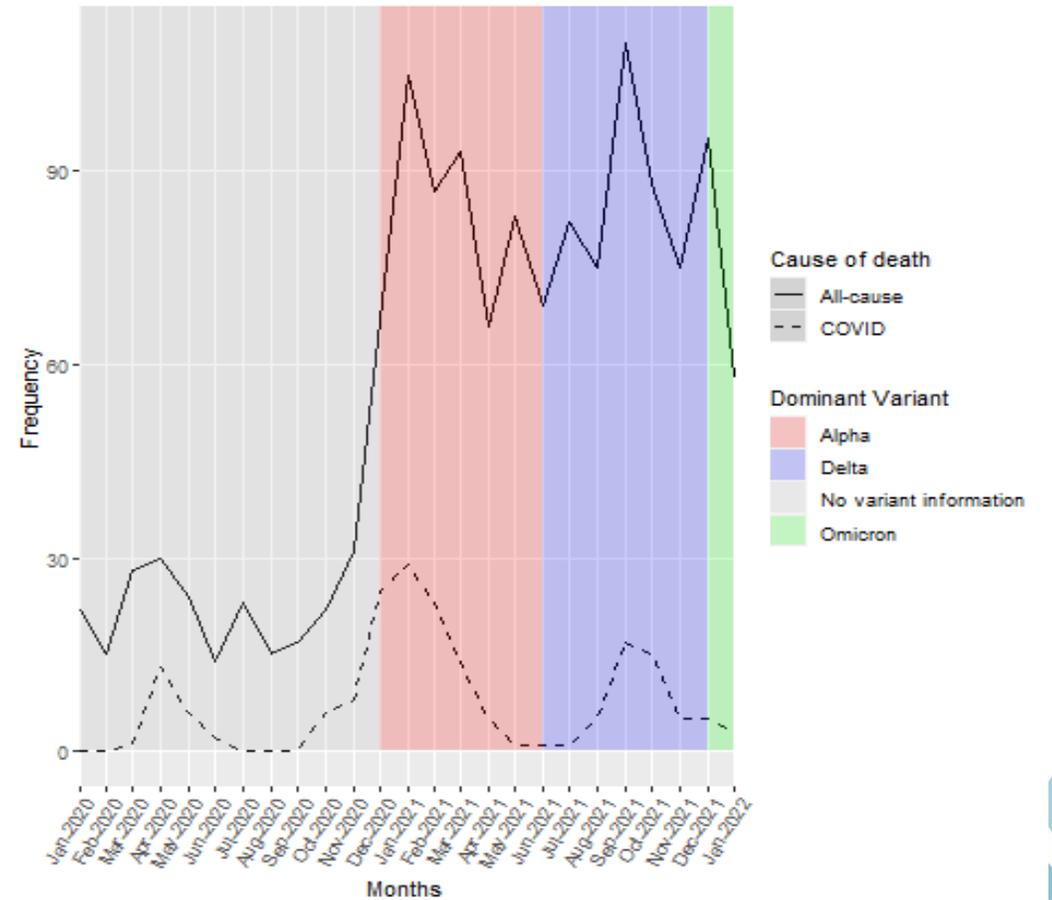
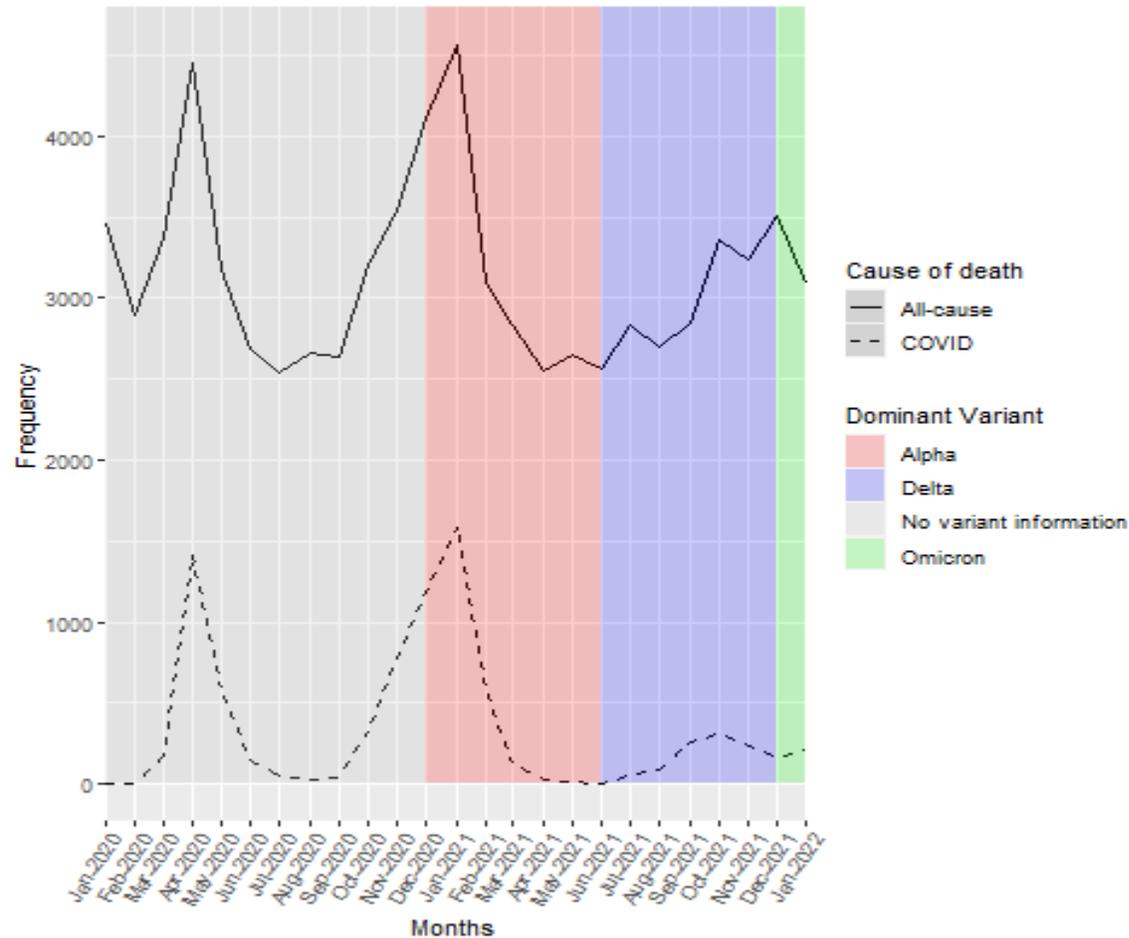
# All-cause mortality Wales: 2016-Jan 2022



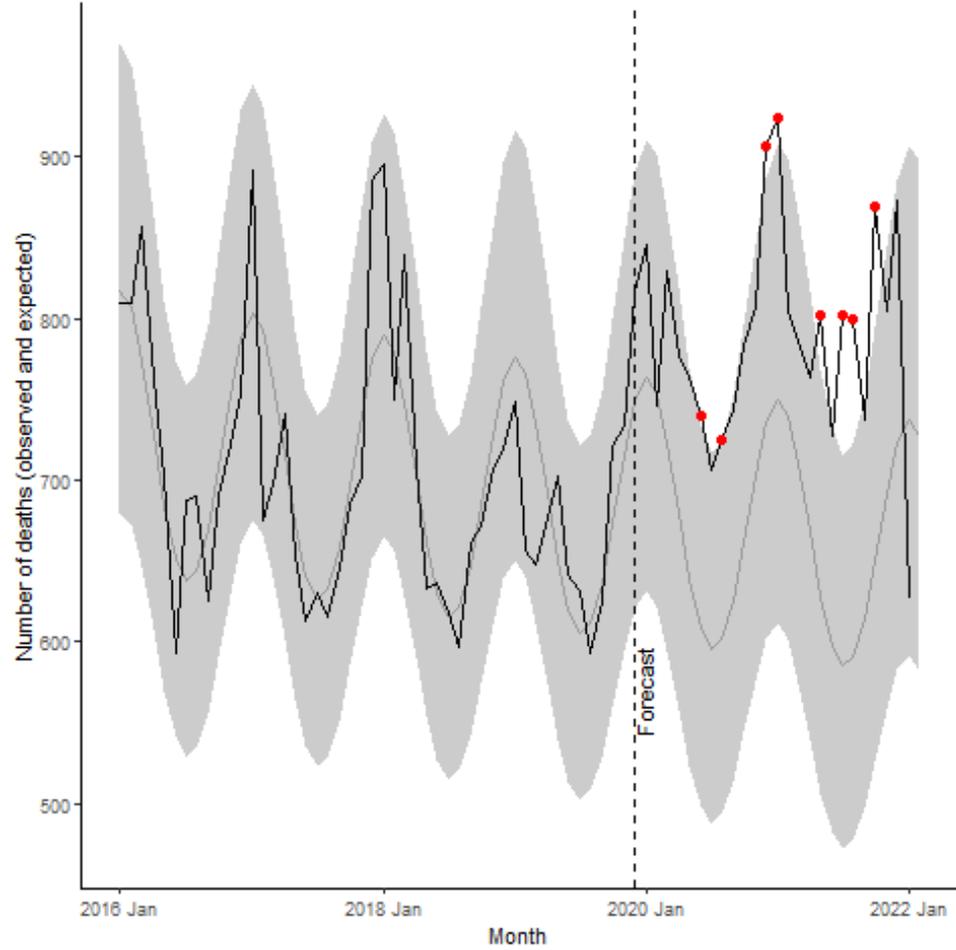
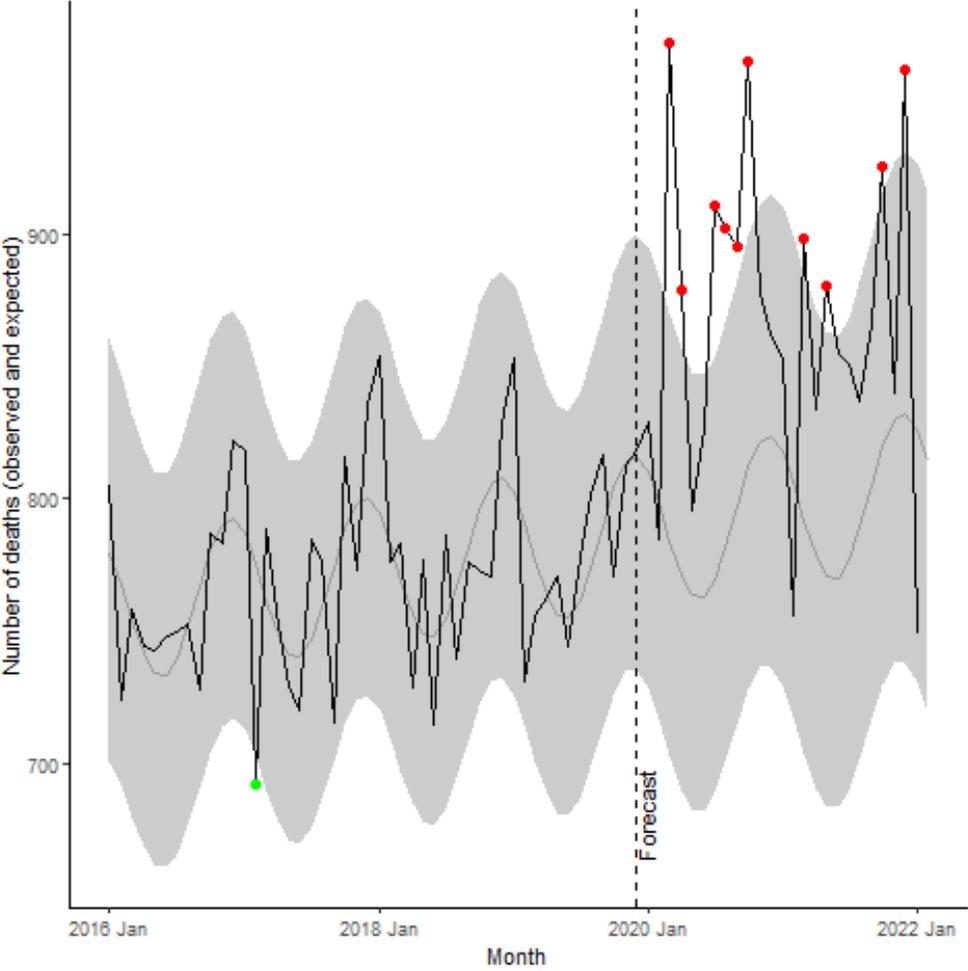
# Place of Death



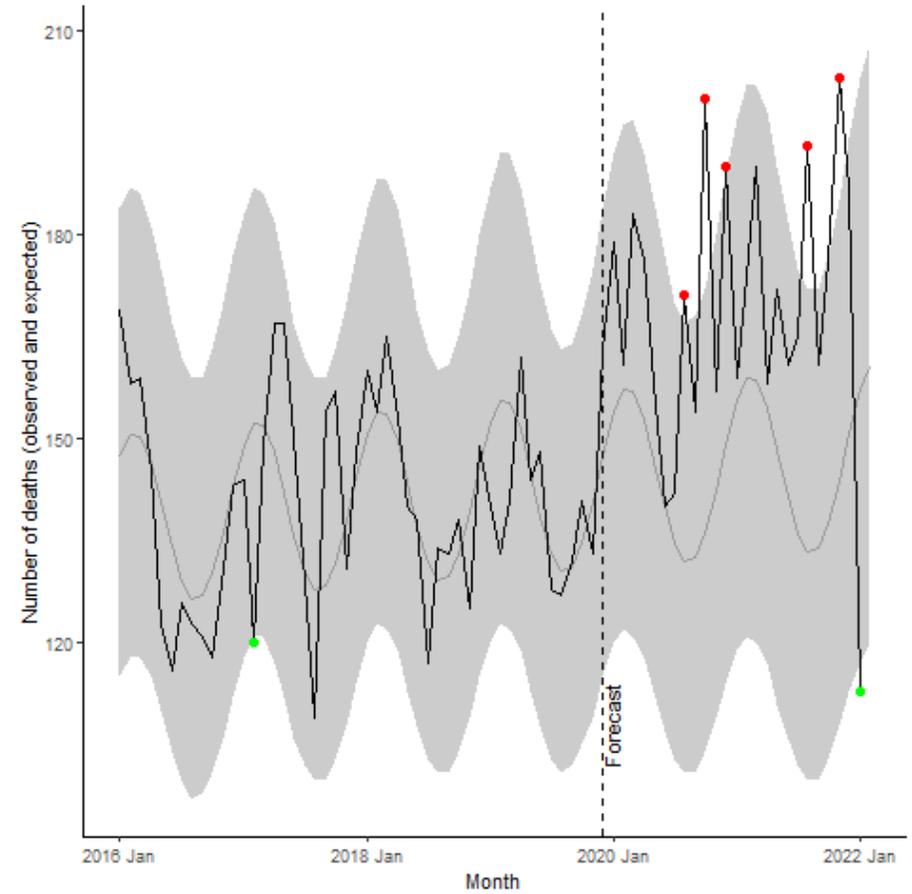
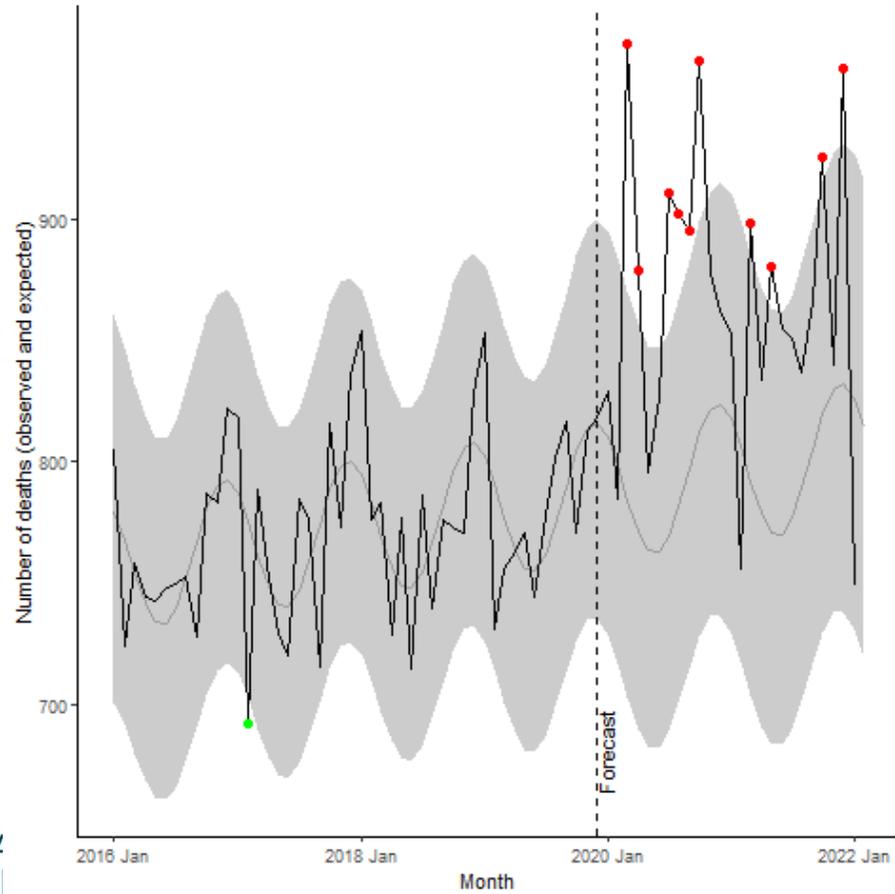
# All-cause and COVID mortality by ethnicity (a) White population (b) Asian population, by dominant variant



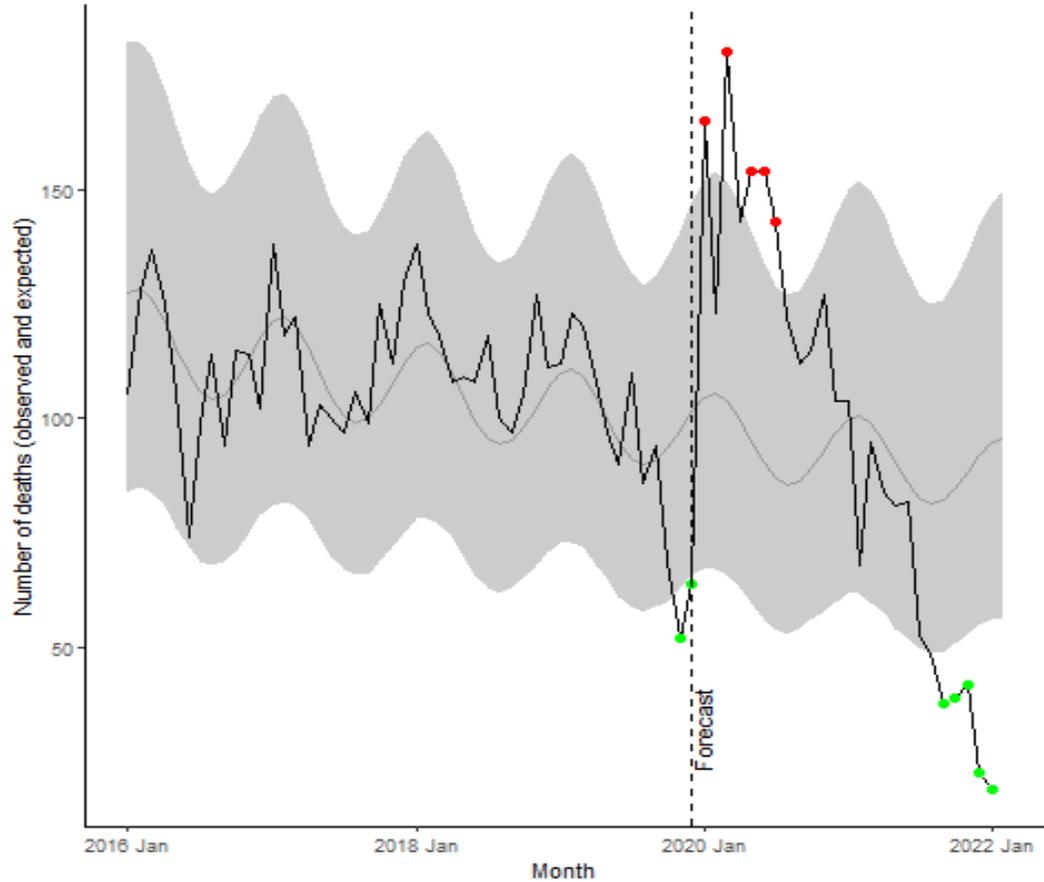
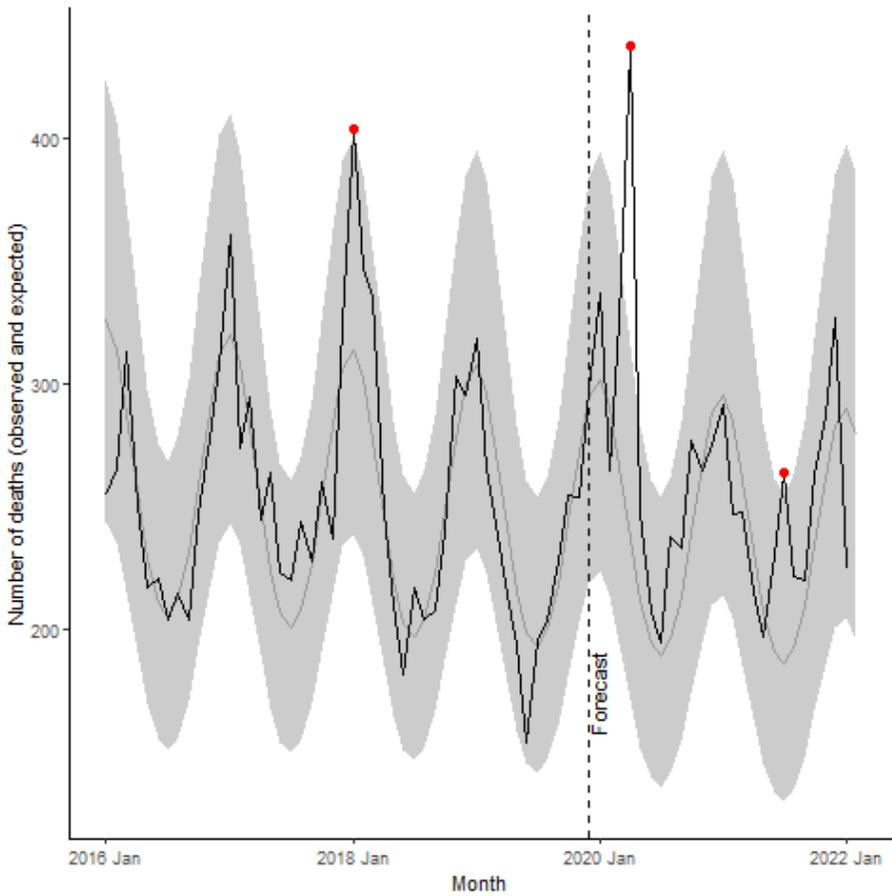
# Mortality by ICD10 chapter: (a) respiratory (b) circulatory Page 20



# Mortality by ICD10 chapter: (a) neoplasms, (b) digestive disorders



# Mortality by ICD10 chapter: (a) mental health, (b) trauma Page 22

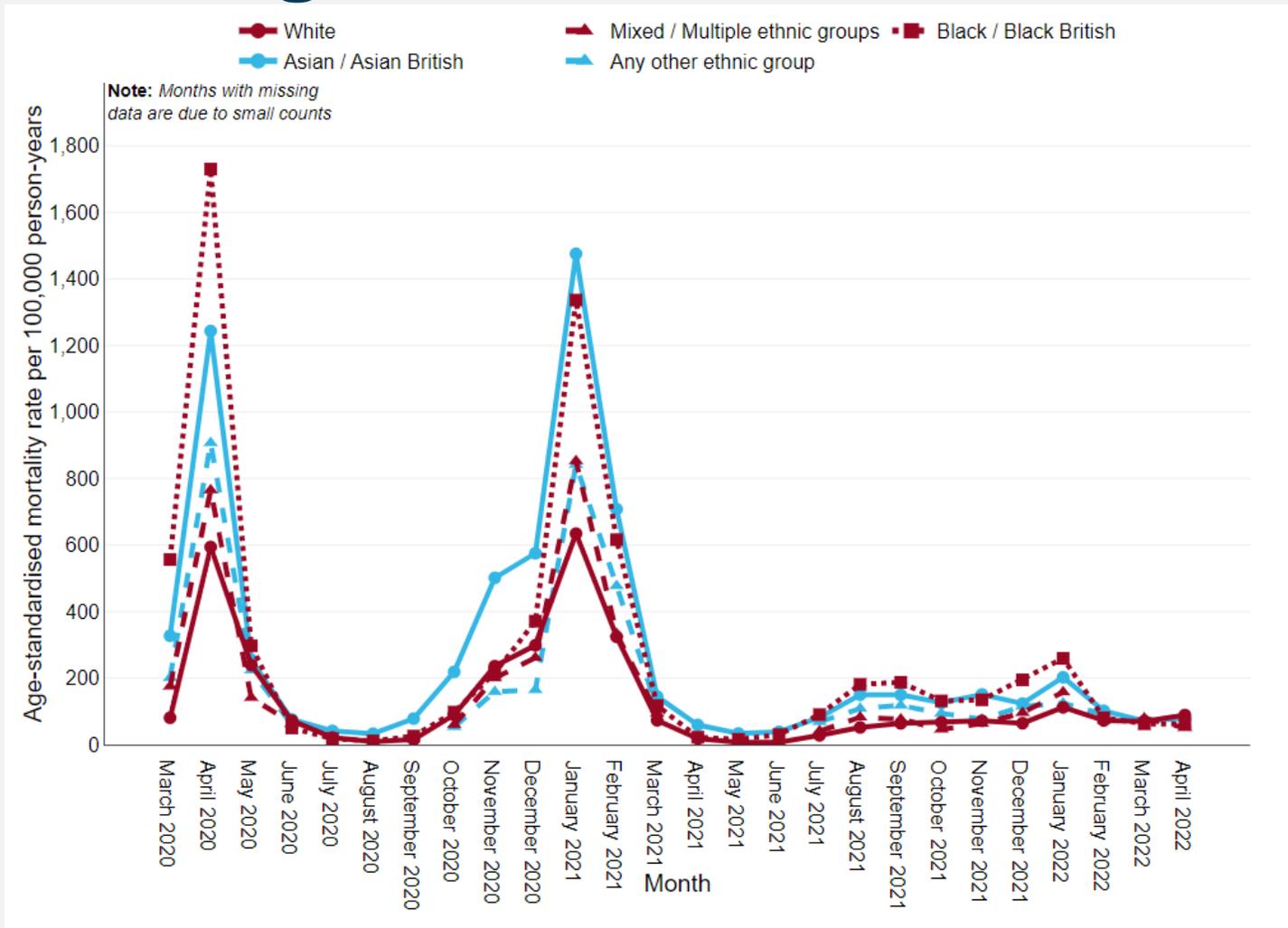


# **Ethnic differences in COVID-19 mortality during the vaccine roll-out in England**

**Matthew Bosworth, Vahé Nafilyan**

Health Analysis and Life Events Division  
Office for National Statistics

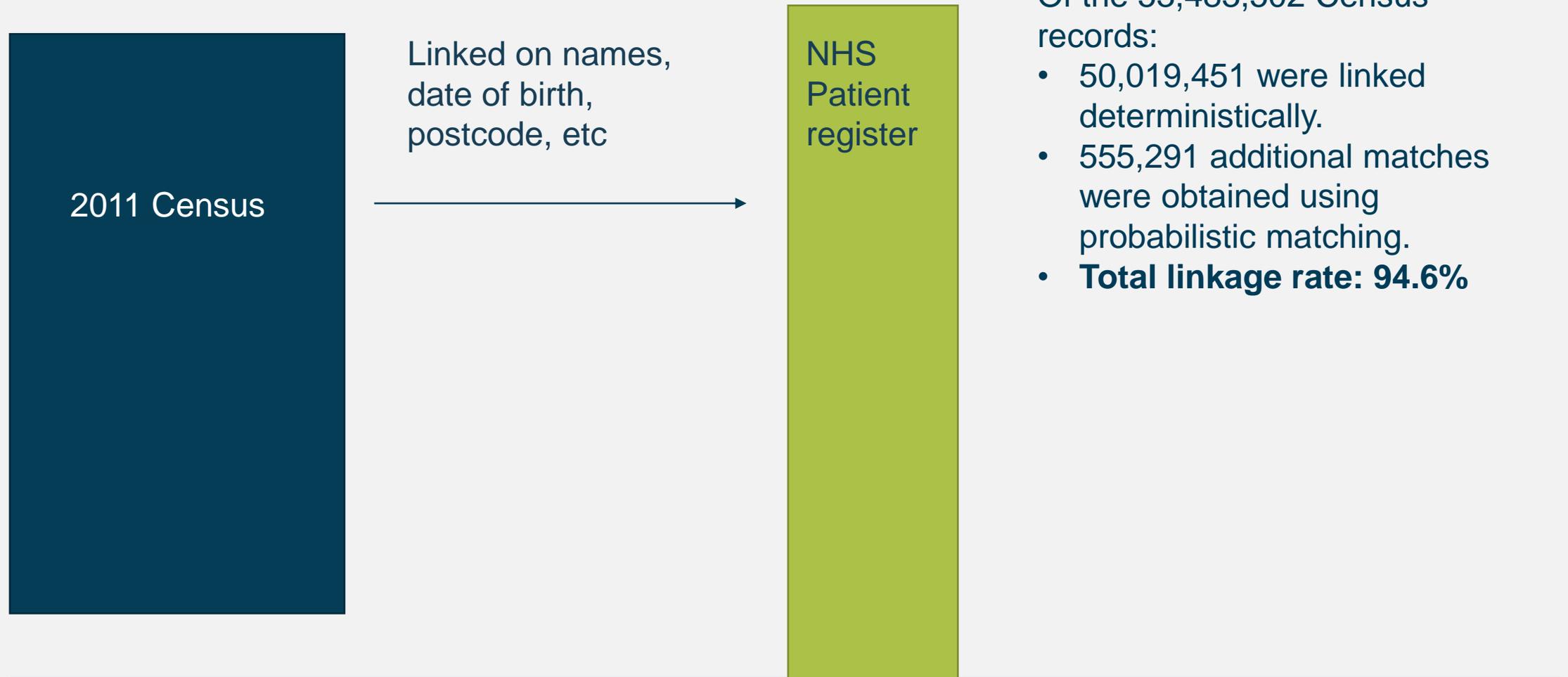
# Background



- Ethnic differences in Covid-19 mortality well documented
- What are the factors explaining these differences?
- What is the role of differences in COVID-19 vaccination uptake?
- Use new linked data based on Census

Source: Public Health England [CHIME tool](#)

# Obtaining NHS numbers for Census respondents



# Linkage failure by ethnic group

	N at 2011 Census	N linked to Patient Register	Rate of linkage failure (%)	Unadjusted OR (95% CI)	OR (95% CI) adjusted for age and sex	OR (95% CI) adjusted for age, sex, ethnicity, region and IMD decile
<b>Bangladeshi</b>	245,164	226,755	7.5	1.86 (1.83-1.89)	1.28 (1.26-1.30)	0.92 (0.91-0.94)
<b>Black African</b>	599,118	527,772	11.9	3.10 (3.07-3.12)	2.36 (2.34-2.38)	1.76 (1.74-1.77)
<b>Black Caribbean</b>	445,503	413,411	7.2	1.78 (1.76-1.80)	1.72 (1.70-1.74)	1.30 (1.28-1.31)
<b>Chinese</b>	286,419	242,791	15.2	4.11 (4.07-4.16)	3.00 (2.97-3.03)	2.69 (2.66-2.72)
<b>Indian</b>	1,036,678	979,485	5.5	1.34 (1.33-1.35)	1.05 (1.04-1.06)	0.92 (0.92-0.93)
<b>Mixed</b>	524,235	439,168	16.2	4.44 (4.40-4.47)	3.20 (3.18-3.23)	2.76 (2.73-2.78)
<b>Pakistani</b>	650,738	610,439	6.2	1.51 (1.50-1.53)	1.09 (1.08-1.10)	0.84 (0.83-0.85)
<b>White British</b>	31,923,883	30,587,948	4.2	Ref	Ref	Ref
<b>White other</b>	2,440,994	2,220,570	9.0	2.27 (2.26-2.28)	1.78 (1.77-1.79)	1.56 (1.55-1.57)
<b>Other</b>	1,095,623	873,949	20.2	5.81 (5.78-5.84)	4.54 (4.51-4.56)	3.72 (3.70-3.74)

# Data on 40.0 million people aged 9+, alive on 24<sup>th</sup> Jan 2020

## Demographics:

sex, age, ethnicity, religion  
Census

## Socio-economics:

Household deprivation,  
education, SES, tenure,  
housing, household  
composition  
Census

## Pre-existing conditions:

As per QCOVID risk model:  
BMI, learning disability, cancer and  
immunosuppression, chronic  
kidney disease, diabetes, COPD,  
other pulmonary diseases,  
coronary heart disease , stroke ,  
dementia , severe mental illness ,  
etc  
GPES/HES

## Geographical factors:

Place of residence, Care  
home/private household,  
Population density,  
rural/urban, area  
deprivation  
2019 Patient register

## Occupational exposure:

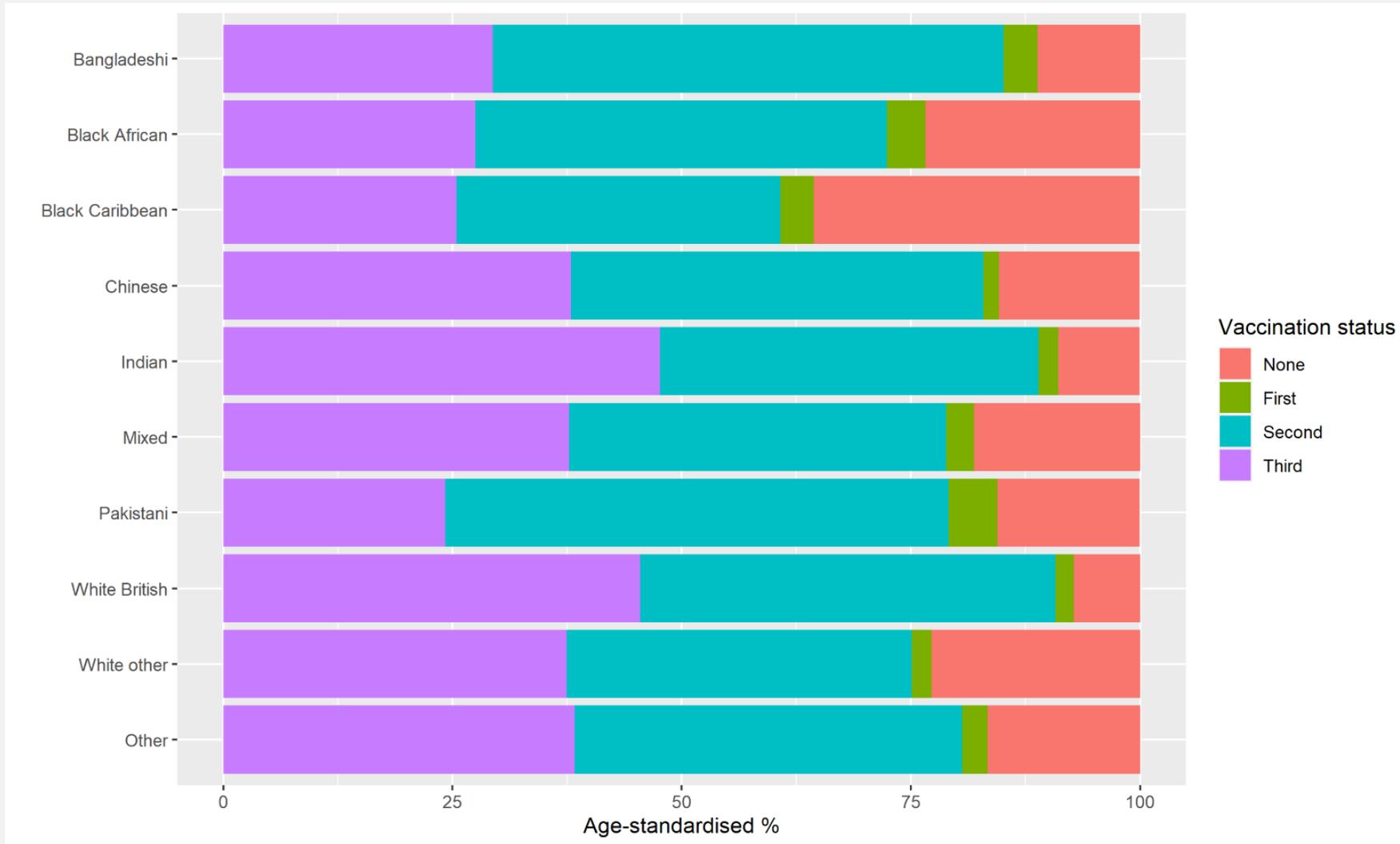
Occupation; O\*NET  
exposure measures  
Census

## Outcomes:

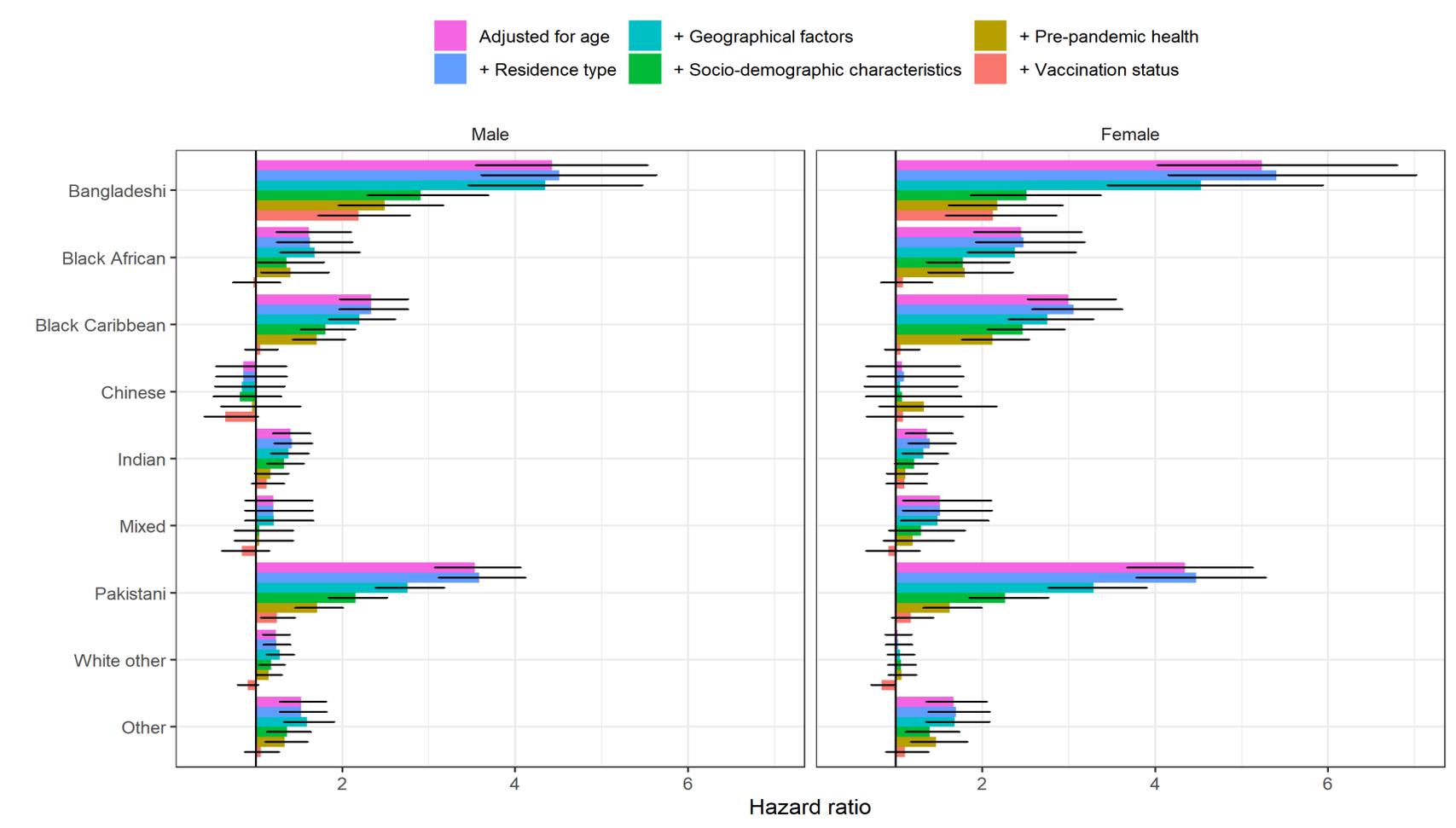
COVID-19 mortality & hospitalisation  
Hospitalisation & diagnosis in primary care

Age	Coverage (% of total pop)
9+	78.6%
40+	85.3%
65+	89.1%

# Vaccination coverage by ethnic groups (1 dec 2021)



# HRs for death involving COVID-19 by ethnic group during the third wave of the pandemic (13 June 2021 to 1 December 2021)



# Summary

- SAIL has a lot to offer for future research on inequalities
- Updating the ethnicity spine with 2021 Census will enable much better analyses to be undertaken
- Funds needed for data analysts to answer questions
- Social Justice theme of the Administrative Data Research Wales plan analyses on many marginalised groups, including ethnic minority groups
- Need for studies of interventions to reduce inequalities