Office for National Statistics

Article

Analysis of geographic concentrations of COVID-19 mortality over time, England and Wales: deaths occurring between 22 February and 28 August 2020

Analysis looking at clusters of deaths involving COVID-19 across time and areas in England and Wales.

Contact:
Jordan Parker
health.data@ons.gov.uk
+44 (0)1633 444110

Release date:
11 January 2021

Next release:
To be announced

Table of contents

1 . Main points

- An analysis of geographic clusters of raised COVID-19 mortality suggests that the known risk factors of age, population density, ethnicity and socioeconomic deprivation only partly explain the distribution of deaths across England and Wales.
- Adjusting for these risk factors reveals unexplained clusters of raised COVID-19 mortality in areas such as the South West and East of England, which have generally seen low rates of COVID-19; other explanations such as particular routes of infection, travel patterns, occupations or household types need to be explored.
- Persistently high mortality in some regions such as the North of England and West Midlands may have been driven by a “core” of relatively small areas with the highest mortality, which may have seen the most intense disease transmission.
- A few areas saw COVID-19 mortality more than seven times the expected level compared with the rest of the country.
- Raised COVID-19 mortality was seen in more deprived areas of South East Wales, but consistently high mortality was also seen in some rural areas after accounting for known risk factors.

2 . Defining geographic concentrations of COVID-19 mortality

We have previously reported on the numbers and rates of deaths involving COVID-19 at different levels of geography including local authorities in England and Wales and local health boards in Wales. Numbers of deaths at smaller area level have been published in interactive maps and a downloadable dataset using a statistical geography called Middle Super Output Areas (MSOAs), which are designed so that each MSOA has a similarly sized population of around 10,000 people.

There are around 7,200 MSOAs in England and Wales. Looking at the period 22 February 2020 to 28 August 2020, the largest number of deaths due to COVID-19 in a single MSOA was 67, while 206 MSOAs (2.9%) had no deaths due to COVID-19.

In this bulletin we investigate more closely the geographic concentration of COVID-19 mortality and how it changed over time, using the MSOA as the basic unit and identifying clusters of MSOAs that had higher mortality rates than other areas. We used a statistical method called spatio-temporal scanning, which compares the mortality rate in each area individually to the mortality rate in the rest of the country for each week. An area is identified as a cluster if its mortality rate in at least one week was statistically significantly higher than in other areas. A cluster can vary in size from a single MSOA to a group of many neighbouring MSOAs as big as a local authority or region.

For this analysis we looked at all deaths involving COVID-19 that occurred between 22 February and 28 August 2020 and had been registered by 12 September 2020. Days were aggregated into weeks that correspond with our Deaths registered weekly in England and Wales provisional release. When discussing a certain week, we are accounting for all deaths that occurred from the start of the period up to and including that week.

All deaths involving COVID-19 were counted and the calculations were carried out for England and Wales separately. Not all deaths involving COVID-19 were in one of these clusters, since a MSOA could have some deaths but at a rate equal to, or lower than, the rest of the country at that time. More information on the methods can be found in Measuring the data.

It is already known that the risk of death involving COVID-19 is influenced by geographic and population characteristics such as the proportion of older people and ethnic minorities, population density and socioeconomic deprivation. To see whether these characteristics explain the patterns observed we present two types of results in this bulletin. One, referred to as “unadjusted”, looks at the clusters identified as having raised mortality rates before taking any area characteristics into account. The second, referred to as “adjusted for known risk factors”, looks at clusters identified after taking account of four area characteristics:

- age structure
- rural-urban classification
- English and Welsh Indices of Deprivation
- proportion of the population in ethnic minority groups

Therefore, the clusters seen after adjustment are areas where there was an unexplained raised rate of COVID-19 deaths even after taking known risk factors into account.

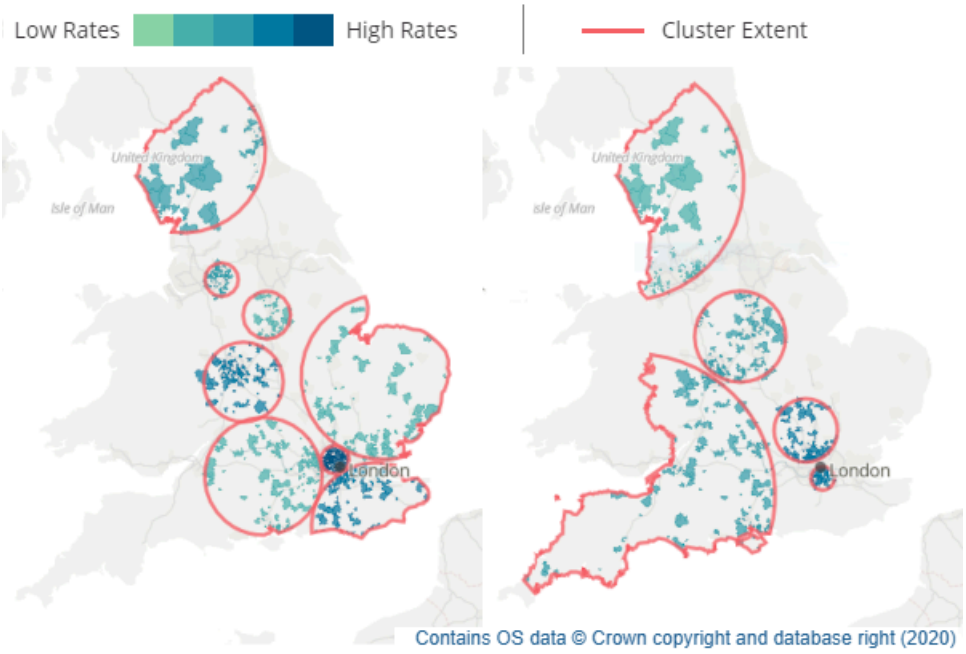
3 . Clusters of higher COVID-19 mortality in England

Figures 1 to 3 show “snapshots” of cluster development in England at different dates over the first wave of the pandemic. Figure 1 shows the areas (Middle Super Output Areas (MSOAs) and groups of MSOAs) identified as having rates of COVID-19 mortality higher than the rest of the country, based on deaths occurring up to 27 March, the week England went into “lockdown”. Figure 1a shows the unadjusted clusters, while Figure 1b shows the clusters identified after adjusting for known risk factors; both maps are shown to illustrate the effect of the adjustment.

The scale shows the ratio of the observed COVID-19 mortality rate in that cluster of MSOAs to the rate in the rest of the country, so (for example) 2.0 means that the mortality rate was twice as high as would have been expected if the mortality rate in that area was the same as other areas.

Figure 1a (left) and 1b (right): Adjustment for known risk factors explains some of the geographical patterns seen early in the pandemic

Clusters of higher COVID-19 mortality, before (left) and after (right) adjustment for known risk factors, deaths occurring up to 27 March 2020, observed to expected ratios, England



Source: Office for National Statistics – Analysis of geographic concentrations of COVID-19 mortality

Notes:

1. Results are observed to expected mortality ratios, calculated by the SaTScan program based on complete weeks corresponding to the dates of the Deaths registered weekly publication. The ratio is calculated for the cluster as a whole; a cluster may contain within it some MSOAs with lower mortality rates, or even no COVID-19 related deaths.

2. Areas that are not in a cluster had equal or lower mortality compared with the rest of the country.

3. Based on date of death, registered up to 12 September 2020.

4. The International Classification of Diseases, 10th Edition (ICD-10) definitions are as follows; coronavirus (COVID-19) (U07.1 and U07.2).

5. Figures exclude deaths of non-residents.

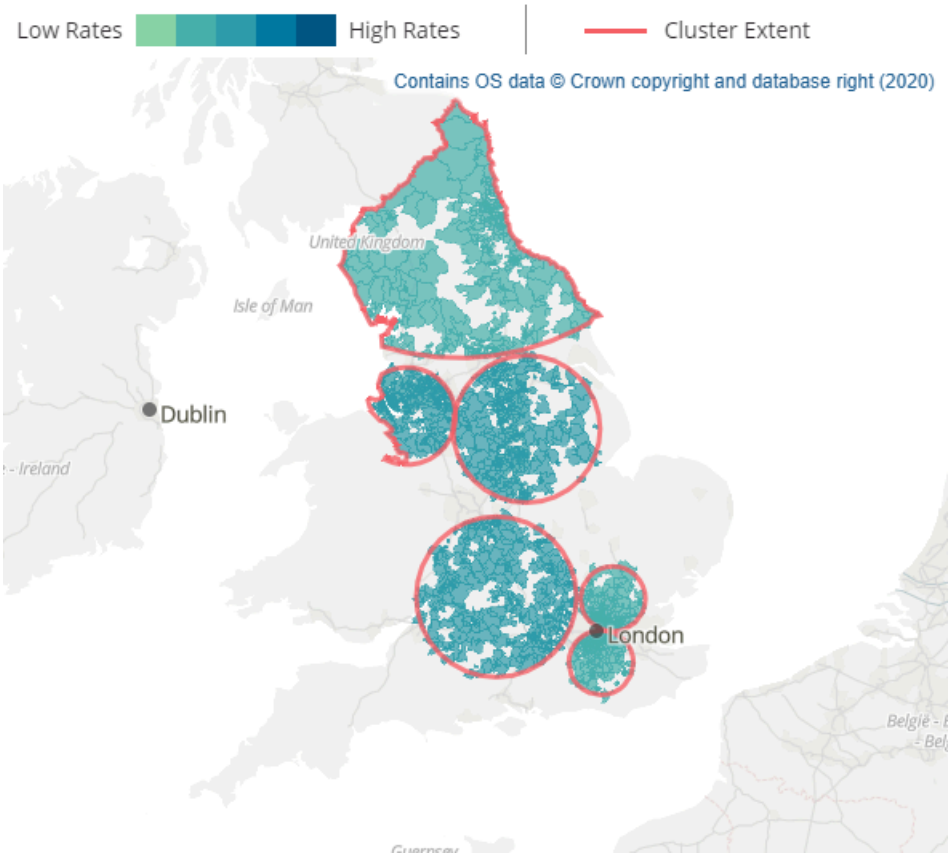
6. All figures for 2020 are provisional.

Looking at the unadjusted map (Figure 1a) we see that there was already a lot of variation in the rate of COVID-19 mortality across England, with some concentration in London and the South East but no one region predominating. The highest ratios relative to the country as a whole were in London (more than five times the mortality rate expected) and the West Midlands. In Figure 1b some clusters are no longer visible after adjusting for the known risk factors of age, population density, ethnicity and deprivation: this suggests that the early clusters of high mortality seen in the North and East of England can be explained by those known factors. The highest mortality is still in London, but unexpectedly high rates are apparent in the East Midlands and small areas of the South West.

Figure 2 shows the clusters having higher rates of COVID-19 mortality than the rest of the country based on deaths occurring up to 17 April, the week containing the “peak” number of deaths in England. This map shows the clusters after adjusting for the known risk factors; that is the areas of higher mortality that are not explained by those factors already identified as important.

Figure 2: By the peak of the first wave of the pandemic, the areas with consistently higher mortality were well-established and less scattered

Clusters of higher COVID-19 mortality after adjustment for known risk factors, deaths occurring up to 17 April 2020, observed to expected ratios, England



Source: Office for National Statistics – Analysis of geographic concentrations of COVID-19 mortality

Notes:

1. Results are observed to expected mortality ratios, calculated by the SaTScan program based on complete weeks corresponding to the dates of the Deaths registered weekly publication. The ratio is calculated for the cluster as a whole; a cluster may contain within it some MSOAs with lower mortality rates, or even no COVID-related deaths.

2. Areas that are not in a cluster had equal or lower mortality compared with the rest of the country.

3. Based on date of death, registered up to 12 September 2020.

4. The International Classification of Diseases, 10th Edition (ICD-10) definitions are as follows; coronavirus (COVID-19) (U07.1 and U07.2).

5. Figures exclude deaths of non-residents.

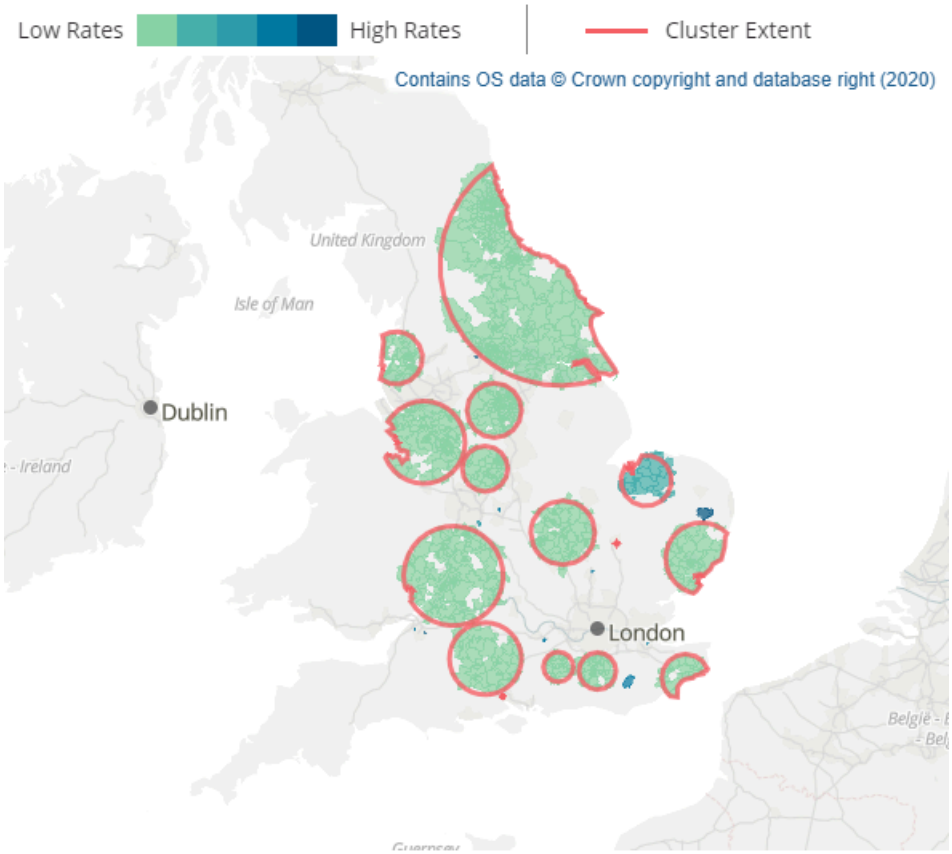
6. All figures for 2020 are provisional.

By 17 April, the clusters are more concentrated and show six distinct geographical groupings. Much of the North of England is now included in clusters, but the highest ratio of unexplained deaths (almost three times that expected) up to this date is in a large area that is centred roughly between Birmingham and London. Despite an increase in the number of MSOAs in clusters, the East of England and South West no longer have any clusters identified. This change demonstrates how previous clusters can “disappear” as their ratio of observed to expected deaths becomes insignificant in comparison with higher-rate clusters. The relative absence of single-MSOA clusters that were widespread and included some rural areas in earlier weeks may reflect higher transmission in urban populations.

Figure 3 shows the clusters having higher rates of COVID-19 mortality than the rest of the country based on deaths occurring up to 5 June, towards the end of the first wave of the pandemic. This map shows the clusters after adjusting for the known risk factors; that is the areas of higher mortality that are not explained by those factors already identified as important.

Figure 3: Towards the end of the first wave, clusters of unexplained higher mortality became mostly smaller and of lower relative intensity

Clusters of higher COVID-19 mortality after adjustment for known risk factors, deaths occurring up to 5 June 2020, observed to expected ratios, England



Source: Office for National Statistics – Analysis of geographic concentrations of COVID-19 mortality

Notes:

1. Results are observed to expected mortality ratios, calculated by the SaTScan program based on complete weeks corresponding to the dates of the Deaths registered weekly publication. The ratio is calculated for the cluster as a whole; a cluster may contain within it some MSOAs with lower mortality rates, or even no COVID-19 related deaths.

2. Areas that are not in a cluster had equal or lower mortality compared with the rest of the country.

3. Based on date of death, registered up to 12 September 2020.

4. The International Classification of Diseases, 10th Edition (ICD-10) definitions are as follows; coronavirus (COVID-19) (U07.1 and U07.2).

5. Figures exclude deaths of non-residents.

6. All figures for 2020 are provisional.

Figure 3 shows how in week ending 5 June, towards the end of the first wave of COVID-19 deaths, clusters of unexplained higher mortality continue to be widespread across England. Individual clusters are now concentrated in smaller areas than during the peak period, except for a large area covering much of the North East, and continue to be present mostly in urban areas. This may reflect the wider dissemination of infections over time. Clusters are apparent in rural areas of the East of England, and may represent specific instances of infection spread, but none are in the South West. Beyond this date, clusters start to be represented as single MSOAs because of the much lower total number of deaths in the summer months.

4 . How clusters of higher COVID-19 mortality changed over time in England

Figure 4 shows the development and change of the clusters of Middle Super Output Areas (MSOAs) where rates of COVID-19 mortality were statistically significantly higher than the rest of the country, week by week, up to 28 August 2020. This map shows the clusters after adjusting for the known risk factors; that is the areas of higher mortality that are not explained by those factors already identified as important. An individual MSOA can be identified by searching for the postcode.

Figure 4:Change over time in clusters of unexplained higher COVID-19 mortality

Clusters of higher COVID-19 mortality after adjustment for known risk factors, by week from 22 February to 28 August 2020, observed to expected ratios, England

Footnotes:

1. Results are observed to expected mortality ratios, calculated by the SaTScan program based on complete weeks corresponding to the dates of the Deaths registered weekly publication. The ratio is calculated for the cluster as a whole; a cluster may contain within it some MSOAs with lower mortality rates, or even no COVID-19 related deaths.

2. Areas that are not in a cluster had equal or lower mortality compared with the rest of the country.

3. Based on date of death, registered up to 12 September 2020.

4. The International Classification of Diseases, 10th Edition (ICD-10) definitions are as follows; coronavirus (COVID-19) (U07.1 and U07.2).

5. Figures exclude deaths of non-residents.

6. All figures for 2020 are provisional.

5 . Duration and intensity of clusters in England

To summarise some of the patterns shown week by week, we can classify Middle Super Output Areas (MSOAs) by how often each one appeared in a cluster of higher COVID-19 mortality, and how high its mortality rates were relative to the rest of the country, after adjusting for known risk factors. Figure 5 shows areas in five categories:

High persistence and rates: the MSOA has a high average ratio of observed to expected COVID-19 deaths and appears frequently in clusters across the study period.

Low persistence, high rates: the MSOA has a high average ratio of observed to expected COVID-19 deaths but appears infrequently in clusters across the study period.

High persistence, low rates: the MSOA has a low average ratio of observed to expected COVID-19 deaths but appears frequently in clusters across the study period.

Low persistence and rates: the MSOA has a low average ratio of observed to expected COVID-19 deaths and appears infrequently in clusters across the study period.

Not in a cluster: the MSOA did not appear in any clusters across the study period.

Those areas in the high persistence and rates category had the highest, most persistent raised mortality that is not explained by the known risk factors.

An individual MSOA can be identified by searching for the postcode.

Figure 5: Most parts of England were in a cluster of higher mortality at least once, the main exceptions being in the South West and parts of Yorkshire

Classification of MSOAs by persistence in a cluster of higher COVID-19 mortality and local COVID-19 mortality rates, after adjustment for known risk factors, 22 February to 28 August 2020, England

Footnotes:

1. Results are obtained from a bivariate combination of persistence in a cluster of raised mortality and raised mortality in the individual MSOA. Periods are based on complete weeks corresponding to the dates of the Deaths registered weekly publication. More information on the methods is contained in Measuring the data.
2. Areas that are not in a cluster had equal or lower mortality compared with the rest of the country.
3. Based on date of death, registered up to 12 September 2020.
4. The International Classification of Diseases, 10th Edition (ICD-10) definitions are as follows; coronavirus (COVID-19) (U07.1 and U07.2).
5. Figures exclude deaths of non-residents.
6. All figures for 2020 are provisional.

The general pattern which emerges across England shows that areas in the North have frequently been in a cluster of higher mortality, with the majority of MSOAs appearing in the high persistence and rates and high persistence and low rates categories. This high-persistence pattern would be expected in the North and areas around London because of their high rates of infection. We also see an area stretching from the West Midlands to the South East that has frequently been included in clusters of higher mortality. The high persistence and low rates areas are more widely distributed around the higher-rate areas, suggesting that the wider regional patterns are be driven by a core of relatively small areas with the highest mortality, which may have seen the most intense disease transmission.

6 . Clusters of higher COVID-19 mortality in Wales

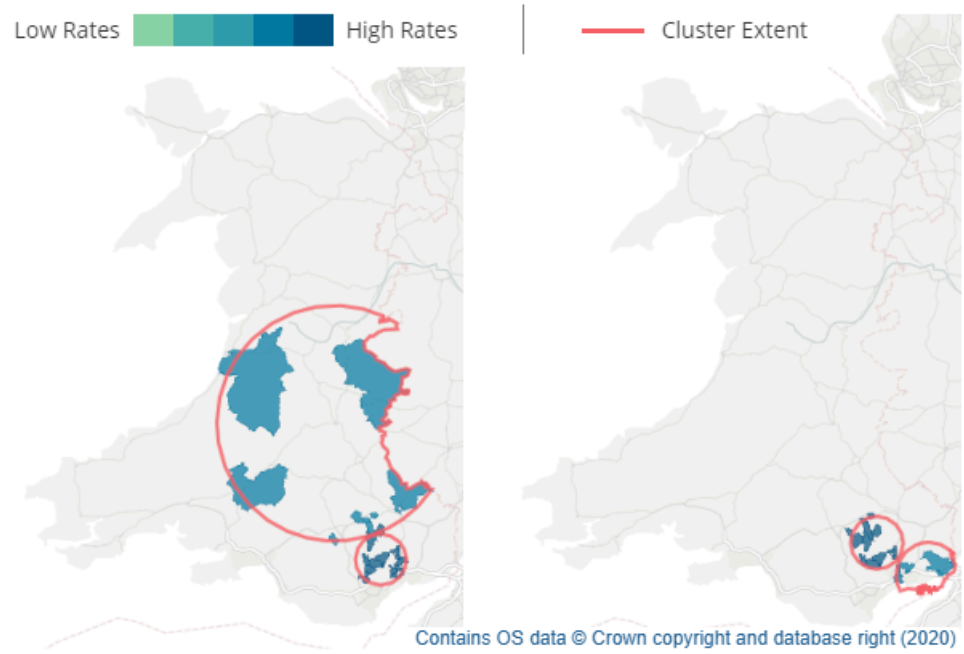
Figures 6 to 8 show snapshots of cluster development in Wales at different dates over the first wave of the pandemic. The same dates have been used as for England. The methods used for the analysis of Middle Super Output Areas (MSOAs) in Wales were identical to the methods used for England, except that the Welsh Index of Multiple Deprivation was used instead of the English Indices of Deprivation. The results for England and Wales are therefore not completely comparable.

Figure 6 shows the areas (MSOAs and groups of MSOAs) identified as having rates of COVID-19 mortality higher than the rest of the country, based on deaths occurring up to 27 March. Figure 6a shows the unadjusted clusters, while Figure 6b shows the clusters identified after adjusting for known risk factors; both maps are shown to illustrate the effect of the adjustment.

The scale shows the ratio of the observed COVID-19 mortality rate in that cluster of MSOAs to the rate in the rest of the country, so (for example) 4.0 means that the mortality rate was four times what would have been expected if the mortality rate in that area was the same as other areas.

Figure 6a (left) and 6b (right): High mortality was concentrated in only a few areas of South East Wales early in the pandemic

Clusters of higher COVID-19 mortality, before and after adjustment for known risk factors, deaths occurring up to 27 March 2020, observed to expected ratios, Wales



Source: Office for National Statistics – Analysis of geographic concentrations of COVID-19 mortality

Notes:

1. Results are observed to expected mortality ratios, calculated by the SaTScan program based on complete weeks corresponding to the dates of the Deaths registered weekly publication. The ratio is calculated for the cluster as a whole; a cluster may contain within it some MSOAs with lower mortality rates, or even no COVID-19 related deaths.

2. Areas that are not in a cluster had equal or lower mortality compared with the rest of the country.

3. Based on date of death, registered up to 12 September 2020.

4. The International Classification of Diseases, 10th Edition (ICD-10) definitions are as follows; coronavirus (COVID-19) (U07.1 and U07.2).

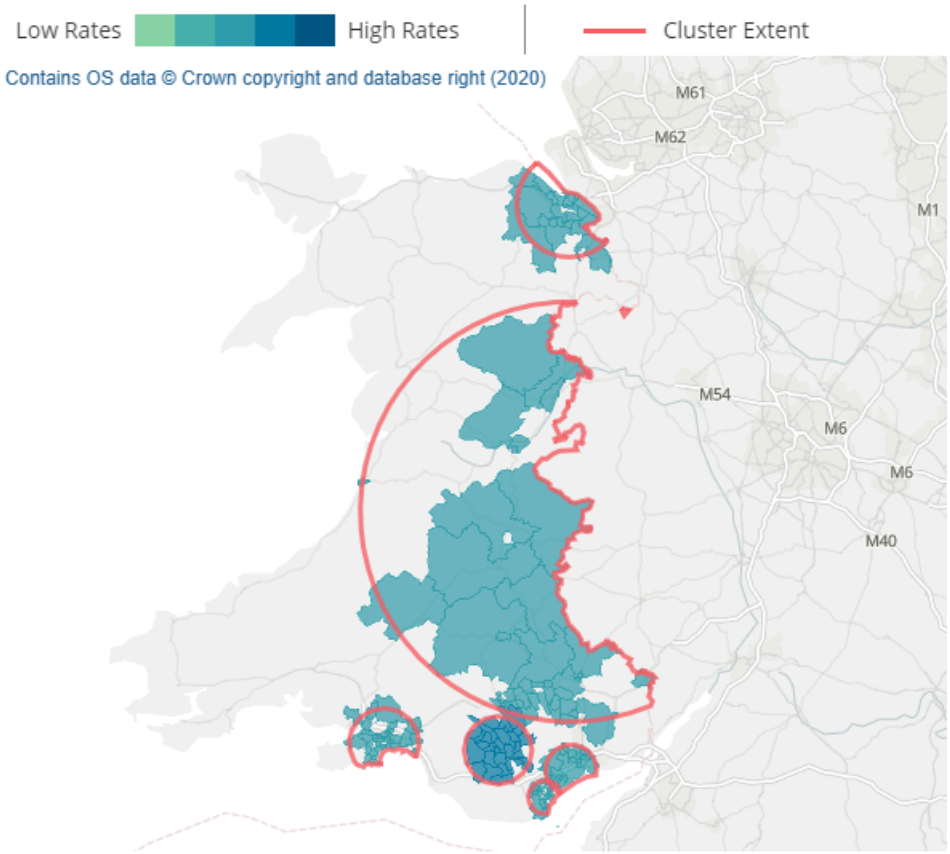
5. Figures exclude deaths of non-residents.

6. All figures for 2020 are provisional.

Lockdown started in Wales on 23 March 2020. Figure 6 shows that in the week ending 27 March, there were only a small number of clusters of raised COVID-19 mortality in Wales, which were around Newport, Torfaen, and Caerphilly. The highest rate was in the area of Caerphilly and the Valleys to the north and was more than seven times what would be expected compared with the whole of Wales. The two clusters identified in Figure 6a are no longer visible in Figure 6b, meaning they have been explained by adjusting for known risk factors. However, two clusters have appeared in the south east of Wales when adjusting for known risk factors.

Figure 7: By the peak of the first wave of the pandemic, areas of Wales with unexplained higher COVID-19 mortality were more widespread

Clusters of higher COVID-19 mortality after adjustment for known risk factors, deaths occurring up to 17 April 2020, observed to expected ratios, Wales



Source: Office for National Statistics – Analysis of geographic concentrations of COVID-19 mortality

Notes:

1. Results are observed to expected mortality ratios, calculated by the SaTScan program based on complete weeks corresponding to the dates of the Deaths registered weekly publication. The ratio is calculated for the cluster as a whole; a cluster may contain within it some MSOAs with lower mortality rates, or even no COVID-19 related deaths.

2. Areas that are not in a cluster had equal or lower mortality compared with the rest of the country.

3. Based on date of death, registered up to 12 September 2020.

4. The International Classification of Diseases, 10th Edition (ICD-10) definitions are as follows; coronavirus (COVID-19) (U07.1 and U07.2).

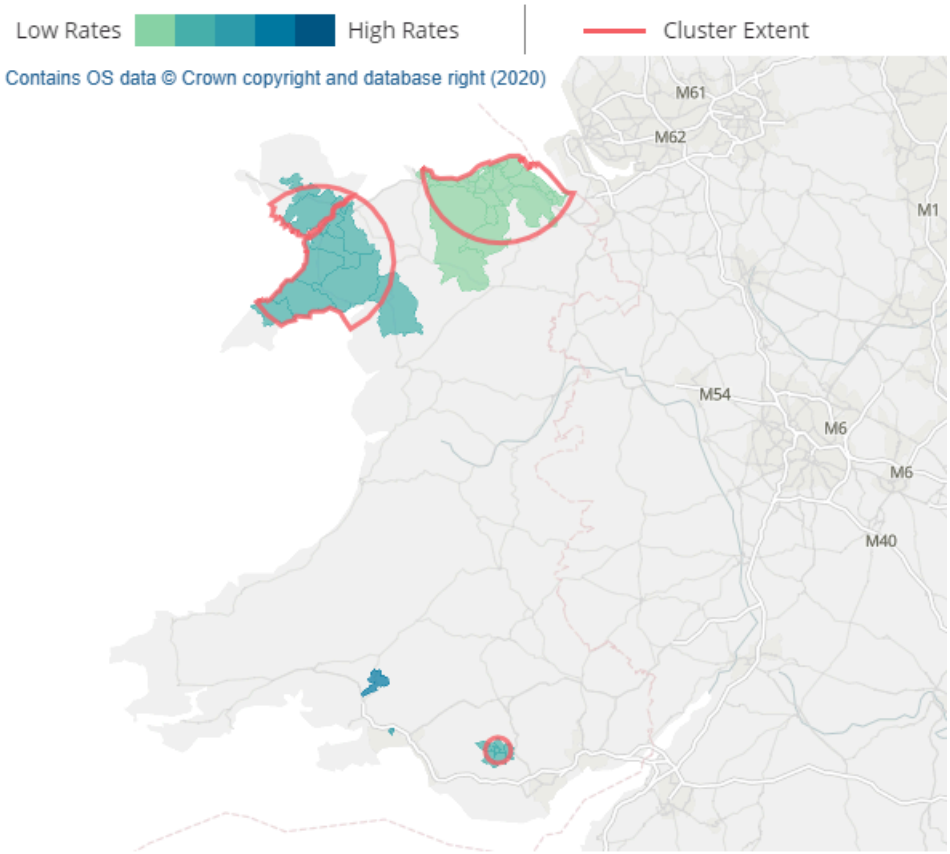
5. Figures exclude deaths of non-residents.

6. All figures for 2020 are provisional.

Looking at clusters up to 17 April 2020, around the peak of the pandemic, the extent of clusters of unexplained high rates of COVID-19 mortality in Wales expanded, now covering much of the border region as well as parts of Cardiff, Newport, and Swansea. This shows evidence of the spread of infections even into more rural areas. Newport and the Valleys no longer had the highest mortality compared with the level expected after adjustment for known risk factors.

Figure 8: Towards the end of the first wave, clusters of unexplained higher mortality in Wales were small and scattered

Clusters of higher COVID-19 mortality after adjustment for known risk factors, deaths occurring up to 5 June 2020, observed to expected ratios, Wales



Source: Office for National Statistics – Analysis of geographic concentrations of COVID-19 mortality

Notes:

1. Results are observed to expected mortality ratios, calculated by the SaTScan program based on complete weeks corresponding to the dates of the Deaths registered weekly publication. The ratio is calculated for the cluster as a whole; a cluster may contain within it some MSOAs with lower mortality rates, or even no COVID-19 related deaths.

2. Areas that are not in a cluster had equal or lower mortality compared with the rest of the country.

3. Based on date of death, registered up to 12 September 2020.

4. The International Classification of Diseases, 10th Edition (ICD-10) definitions are as follows; coronavirus (COVID-19) (U07.1 and U07.2).

5. Figures exclude deaths of non-residents.

6. All figures for 2020 are provisional.

Looking at clusters identified up to 5 June, towards the end of the first wave, many of Wales' previous clusters of higher mortality are no longer present, with only a few affected MSOAs scattered around the country. Although previously large clusters were seen in the south and east, the clusters identified late in the pandemic included areas of North Wales. No statistically significant clusters were detected in Wales after 10 July 2020, this means that looking at the period as a whole, there were no significant clusters in Wales when adjusting for known risk factors.

7 . How clusters of higher COVID-19 mortality changed over time in Wales – interactive

Figure 9 shows the development and change of the clusters of Middle Super Output Areas (MSOAs) where rates of COVID-19 mortality were statistically significantly higher than the rest of the country, week by week, up to 28 August 2020. This map shows the clusters after adjusting for the known risk factors; that is the areas of higher mortality that are not explained by those factors already identified as important. An individual MSOA can be identified by searching for the postcode.

Figure 9: Change over time in clusters of unexplained higher COVID-19 mortality

Clusters of higher COVID-19 mortality after adjustment for known risk factors, by week from 22 February to 28 August 2020, observed to expected ratios, Wales

Footnotes:

1. Results are observed to expected mortality ratios, calculated by the SaTScan program based on complete weeks corresponding to the dates of the Deaths registered weekly publication. The ratio is calculated for the cluster as a whole; a cluster may contain within it some MSOAs with lower mortality rates, or even no COVID-19 related deaths.
2. Areas that are not in a cluster had equal or lower mortality compared with the rest of the country.
3. Based on date of death, registered up to 12 September 2020.
4. The International Classification of Diseases, 10th Edition (ICD-10) definitions are as follows; coronavirus (COVID-19) (U07.1 and U07.2).
5. Figures exclude deaths of non-residents.
6. All figures for 2020 are provisional.

8 . Duration and intensity of clusters in Wales

To summarise some of the patterns shown week by week above, we can classify Middle Super Output Areas (MSOAs) by how often each one appeared in a cluster of higher COVID-19 mortality, and how high its mortality rates were relative to the rest of the country, after adjusting for known risk factors. Figure 10 shows areas in five categories:

High persistence and rates: the MSOA was included in one or more clusters of higher rates for most of the time period and had repeatedly high COVID-19 mortality rates itself.

Low persistence, high rates: the MSOA had high COVID-19 mortality rates but was only included in a cluster of higher rates once or a few times.

High persistence, low rates: the MSOA was included in one or more clusters of higher rates for most of the time period but did not have notably high COVID-19 mortality rates itself for most of the time.

Low persistence and rates: the MSOA was only included in a cluster of higher rates once or a few times and did not have notably high COVID-19 mortality rates itself.

Not in a cluster: the MSOA had COVID-19 mortality rates equal to or lower than the rest of the country and was never included in a cluster of higher rates.

Those areas in the high persistence and rates category had the highest, most persistent raised mortality that is not explained by the known risk factors.

An individual MSOA can be identified by searching for the postcode.

Figure 10: All parts of Wales were in a cluster of higher mortality at least once

Classification of MSOAs by persistence in a cluster of higher COVID-19 mortality and local COVID-19 mortality rates, after adjustment for known risk factors, 22 February to 28 August 2020, Wales

Footnotes:

1. Results are obtained from a bivariate combination of persistence in a cluster of raised mortality and raised mortality in the individual MSOA. Periods are based on complete weeks corresponding to the dates of the Deaths registered weekly publication. More information on the methods is contained in Measuring the data.
2. Areas that are not in a cluster had equal or lower mortality compared with the rest of the country.
3. Based on date of death, registered up to 12 September 2020.
4. The International Classification of Diseases, 10th Edition (ICD-10) definitions are as follows; coronavirus (COVID-19) (U07.1 and U07.2).
5. Figures exclude deaths of non-residents.
6. All figures for 2020 are provisional.

Unlike England, every MSOA in Wales appeared in a cluster at some point through the first wave of the pandemic. The majority of areas appeared in a cluster over a long period of time (the high persistence and rates and high persistence, low rates categories). The inclusion of rural areas in the higher-rate categories, which seems initially surprising, shows that the known risk factors for which we adjusted do not account for the wider patterns seen across Wales.

9 . Interpreting the adjusted cluster patterns

These findings based on a spatio-temporal clustering method add to our previous publications on the Deaths involving COVID-19 by local area and socioeconomic deprivation involving COVID-19 by showing which areas had higher than expected COVID-19 mortality rates compared with the rest of the country. The unadjusted clustering patterns largely reflect known risk factors such as age, population density, socioeconomic deprivation, and ethnic minority population, while the adjusted clustering patterns show the areas where mortality was higher than expected even after taking these factors into account.

The results suggest that while some geographical variation in COVID-19 mortality is explained by the risk factors mentioned, substantial variation remains unexplained. In fact, adjustment for these risk factors also reveals higher than expected mortality in some areas, such as parts of the South West of England, where there are no clusters in the unadjusted rates.

This means that it is necessary to look beyond the major risk factors already identified to understand the spatio-temporal patterns of the pandemic. Two particular areas need further research: firstly the routes of infection into and around the country, which must partly determine the spread of COVID-19 leading to the observed deaths; and secondly more nuanced aspects of socioeconomic and demographic difference that commonly-used measures such as the Index of Multiple Deprivation may not detect, such as specific groups of occupations or types of household associated with high infection risk.

The temporal trends observed also highlight the varying ways in which some areas experienced waves of COVID-19 deaths. While some local areas experienced infrequent, low numbers of COVID-19 deaths, other areas throughout the pandemic have experienced consistently elevated rates. This seems surprising in some areas such as the area from Birmingham to Southampton, which is a continuously changing mix of rural and urban populations. A continuously high rate of COVID-19 deaths also implies a high rate of infections, which allow for more deaths, and that these infections may have found their way into populations not seen as vulnerable. Therefore, the dynamic nature of how infections spread through a given time period, and the connections between rural and urban communities (for example, through travel for work, shopping or social life), is also worthy of investigating to help explain these trends.

10 . Analysis of geographic concentrations of COVID-19 mortality over time, England and Wales data

Analysis of geographic concentrations of COVID-19 mortality over time, England and Wales
Dataset | Released 11 January 2021
Analysis looking at clusters of deaths involving COVID-19 across time and areas in England and Wales

11 . Glossary

Coronaviruses

The World Health Organization (WHO) defines coronaviruses as "a large family of viruses that are known to cause illness ranging from the common cold to more severe diseases such as Middle East Respiratory Syndrome (MERS) and Severe Acute Respiratory Syndrome (SARS)". Between 2001 and 2018, there were 12 deaths in England and Wales due to a coronavirus infection, with a further 13 deaths mentioning the virus as a contributory factor on the death certificate.

Coronavirus (COVID-19)

COVID-19 refers to the "coronavirus disease 2019" and is a disease that can affect the lungs and airways. It is caused by a type of coronavirus. Further information is available from the WHO.

Registration delay

Mortality statistics are compiled from information supplied when deaths are certified and registered as part of civil registration, a legal requirement. According to the Births and Deaths Registration Act 1953, a death should be registered within five days unless it is referred to a coroner for investigation. Mortality statistics for a given time period can be based on occurrence (death date) or registration (registration date); registration delay is the difference between date of occurrence and date of registration.

Statistical significance

The term "significant" refers to statistically significant changes or differences. Significance has been determined using the 95% confidence intervals, where instances of non-overlapping confidence intervals between estimates indicate the difference is unlikely to have arisen from random fluctuation. In some circumstances, significance has also been tested using z scores. More information about this z test is available in Appendix 1 of the Sullivan guide.

12 . Measuring the data

Data used

The figures represent the number of deaths involving COVID-19 that occurred between 22 February and 28 August 2020 and were registered by 12 September 2020; a proportion of deaths occurring in this year will not be registered until subsequent years (more information can be found in our Impact of Registration Delays release).

Identifying clusters

We used SaTScan v9.6.1, a publicly available software program which can assist in automatically detecting clusters of events (such as deaths) which have both spatial and temporal dimensions. SaTScan has been widely used in epidemiology.

The method used to complete the analysis for MSOA-level clusters was based primarily on work done in Hohl et al, 2020. This method allows for a time series to be created through the use of a Poisson-distributed, prospective analysis in SaTScan when supplying a varying time window. This was found to be more useful for tracking the course of death clusters than a retrospective analysis as it effectively simulates a live view of cluster distribution at the end of each weekly time-slice.

The weeks used to create these time slices follow the sequence:

‘Week 1, Week 1 to Week 2, Week 1 to Week 3, … , Week 1 to Week N’

Input deaths data is split according to this sequence for the full range of weeks ending 28 February 2020 to 28 August 2020.

Three types of data were used as inputs.

Case files:

- individual deaths with registered date of death and age of the deceased by MSOA
- ages of the deceased grouped to “under 65 years” and “65 years and over”
- all combinations of MSOAs, dates and age groups were created such that all possible data points are included –this includes MSOAs on dates with no deaths in either age group

Population file:

- mid-year 2018 MSOA population estimates by age
- ages of the population grouped to “under 65 years” and “65 years and over”
- each MSOA included once per age group

Coordinates file:

- 2011 MSOA population-weighted centroids
- centroid coordinates represented as latitude (y) and longitude (x)

Data sources:

- age: Lower layer Super Output Area population estimates (Mid-2018)
- Rural Urban Classification: Rural Urban Classification (2011) of Middle Layer Super Output Areas in England and Wales
- ethnicity: CDRC Modelled Ethnicity Proportions (LSOA Geography) 2016
- Index of Multiple Deprivation: Index of Multiple Deprivation (December 2019) Lookup in England, Welsh Index of Multiple Deprivation

SaTScan was run separately across the time sequence for five covariate combinations, with new covariates included additively:

- no adjustments
- age groups
- age groups and RUC2011 (2011 Rural-Urban Classification)
- age groups, RUC2011 and BAME DECILE (MSOA-aggregated, decile of mean LSOA-level proportion of population as BAME)
- age group, RUC2011, BAME Decile and IMD Decile (median of MSOA-aggregated LSOA-level IMD)

Calculations for England and Wales were run separately as they have different Indices of Deprivation. Inspecting these runs showed a clear trend in that adding known risk factors could “explain” some clusters, and as such only outputs with no covariates and all covariates have been included in the analysis.

To ensure that clusters are limited in both spatial and temporal size, and can be found at a resolution appropriate for the scale of the analysis, the following parameters were used in SaTScan:

- minimum cases (deaths) per cluster = 10
- high-rate clusters minimum relative risk = 1
- time aggregation = 7 days
- maximum spatial cluster size = 10% of the population at risk
- maximum temporal cluster size = 50% of the study period
- minimum temporal cluster size = 7 days

Classifying areas according to cluster persistence and mortality rates

The time series maps rates are classified using Jenks breaks on the full dataset individually for England and Wales. The bivariate map categories are based on the median value for rates and persistence individually for England and Wales. Therefore, a high persistence or rate has a value higher than half of all MSOAs in the country, while low persistence or rates have a value lower than half of all MSOAs. Cut-off rates are:

- England: high persistence 6 weeks, high rate 1.19
- Wales: high persistence 5 Weeks, high rate 1.21

Acknowledgements

We would like to thank Naru Shiode, Shino Shiode and Richard Elson for their help in the understanding and use of SaTScan and providing quality assurance on the methods used and technical decisions made.

13 . Strengths and limitations

Strengths

Coding for cause of death is carried out according to the World Health Organization (WHO) ICD-10 and internationally agreed rules.

Information is supplied when a death is registered, which gives complete population coverage and ensures the estimates are of high precision, and representative of the underlying population at risk.

The location of deaths (based on the usual residence of the deceased) is accurately captured and mapped to MSOA boundaries.

Limitations

Deaths used in this analysis are based on registrations up to 12 September 2020. Some deaths that occurred in the period of interest will not have been registered by that cut-off date and will not be included in this analysis. More information can be found in our Impact of Registration Delays release.

Deaths of non-residents of England and Wales are not included as they could not be assigned to a MSOA of residence; these make up only a very small proportion of total deaths.

Deaths were aggregated for the SaTScan calculations into specific complete weeks (Saturday to Friday) rather than rolling weeks; this could potentially cause artefactual changes to the results if the distribution of deaths by day is uneven.

Numbers of infections or COVID-19 cases per area are not included as a covariate. Because of issues with data quality from limited testing capacity during the first wave of the pandemic, it would not be possible to ensure that cases recorded in the first weeks of the analysis are truly representative of the distribution of the virus. With increasing capacity since the first wave, the number of active weekly cases could be used in future analysis.

Sex is not included as a covariate, however differences in the sex composition of areas in England and Wales are not thought to be sufficient to affect the patterns seen.

The number of care homes in an area were not included as a covariate as we currently do not hold information on the number of care homes or care home residents in a specific area.

The prospective space-time analysis used in this approach is a novel method. While the spatial distribution of clusters was validated using a separate retrospective analysis with the same inputs, the temporal size of clusters may not be consistent with that found from a retrospective analysis.

While there have been a large number of COVID-19 related deaths, and these have resulted in higher than average death rates across the UK, the absolute number of deaths is small compared with the overall population of England and Wales. As such, clusters, particularly when adjusted for risk factors, may be very sensitive to small changes in input cases and changes in covariates. The distribution of risk-adjusted death clusters is therefore based solely on the outputs from SaTScan and does not necessarily represent an absolute view of how COVID-19 related death rates have varied across the country.

Despite these limitations we have confidence that, given the data used and validation of the outputs, the overall trends discussed are accurate and representative of how risk-adjusted clusters across England and Wales have varied both spatially and temporally over the course of the study period.

14 . Related links

Deaths registered in England and Wales: 2019
Bulletin | Released 1 July 2020
Registered deaths by age, sex, selected underlying causes of death and the leading causes of death. Contains death rates and death registrations by area of residence and single year of age.

Coronavirus (COVID-19) latest data and analysis
Web page | Updated as and when new data become available
Brings together the latest data and analysis on the coronavirus (COVID-19) pandemic in the UK and its effect on the economy and society.

Coronavirus (COVID-19) roundup
Blog | Updated as and when new data become available
Catch up on the latest data and analysis related to the coronavirus pandemic and its impact on our economy and society.

Coronavirus and the latest indicators for the UK economy and society
Bulletin | Released 3 December 2020
Early experimental data on the impact of the coronavirus on the UK economy and society. These faster indicators are created using rapid response surveys, novel data sources and experimental methods.

Monthly mortality analysis, England and Wales: October 2020
Bulletin | Released 19 November 2020
Provisional death registration data for England and Wales, broken down by sex, age and country. Includes deaths due to COVID-19 and leading causes of death.

Deaths involving COVID-19 by local area and socioeconomic deprivation: deaths occurring between 1 March and 31 July 2020
Bulletin | Released 28 August 2020
Provisional counts of the number of deaths and age-standardised mortality rates involving COVID-19 between 1 March and 31 July 2020 in England and Wales. Figures are provided by age, sex, geographies down to local authority level, and deprivation indices.