Royal Statistical Society
Official Statistics Section
Health Statistics User Group

Collecting and reporting mortality statistics for Covid-19

Wednesday May 20, 12.00-13.30
Collecting and reporting mortality Statistics

Myer Glickman
Head of Quality, Methods and International

20 May 2020
Codes and causes

• U07.1 Confirmed COVID-19 (95%)

• U07.2 Suspected COVID-19 (5%)
  • Suspected COVID-19, maybe COVID-19, possible COVID-19, likely to be COVID-19, pneumonia in COVID-19 outbreak, heart disease during COVID-19 pandemic...

• Underlying cause of death rules similar to influenza: 95% UCOD
• Deaths ‘due to’ is UCOD v. ‘deaths involving’ is any mention
Data and definitions

• DHSC/NHS England: deaths with positive test in hospital
• DHSC/PHE: deaths with positive test anywhere
• PHW, PHS, PHA: deaths with positive test anywhere
• CQC, CIW: deaths with reported COVID in care services
• ONS, NRS, NISRA: deaths with recorded COD anywhere
• + excess deaths – current v. average (usually 5yr)
Figure 1: The cumulative number of deaths involving COVID-19 in England using different data sources, up to 8 May 2020

Source: Office for National Statistics
Death registration data flows
Analysis plans

• Weekly update and comparisons ONGOING
• Monthly occurrences analysis ONGOING
• Geography and deprivation DONE – TO REPEAT
• Ethnic group DONE – TO DEVELOP
• Care home residents - DONE
• Excess deaths – IN PROGRESS
• Religion – PLANNED
• Disability – PLANNED
• ... WHAT ELSE?
COVID-19 and all-cause mortality

Dr Nick Andrews, PHE, Statistics Unit, Colindale
On behalf of

PHE COVID-19 Epi-Cell
PHE COVID-19 Surveillance Cell

May 2020
All Cause Mortality – published in the weekly PHE Surveillance reports and Euro momo

- https://www.euromomo.eu/

- Original Data source: General Register Office, Southport
- Data on all registrations to midnight provided to PHE the following day as aggregated data by date of death, date of registration, age in years, sex, region where death registered.
- Agreements to publish Z-score analysis (Euromomo) and to publish figures in the weekly mortality report as has been done with ONS weekly data since 2009.

- Results reported every Thursday for deaths to week ending previous Sunday.
Methods

Baseline is calculated using 5 years historic data with a Serfling (sine) wave fitted to spring/autumn data and a trend. Model is Poisson with rescaling for over dispersion. Thresholds and Z scores are calculated from this model.

Observed data are corrected based on the historic proportions of deaths registered with delays of 0,1,2,3,…,50 weeks. Each proportion P0 to P50 is modelled with a trend in this proportion over time and based on the number of days registration offices are closed within that interval. This can then be used to adjust recent counts for registration delays.

Most deaths are registered within 3-4 weeks but delays can be long (many months) if deaths go to a Coroner - common in younger ages (deaths where individuals have not been seen recently by a doctor).
Output in latest PHE report

Figure 24: weekly observed and expected number of all-cause deaths in all ages, with the dominant circulating influenza type(s), England, 2015 to week 19 2020
Euromomo output (Z scores)

All ages

UK (England)

0 to 4

UK (England)
15 to 64

UK (England)

65+

UK (England)
Comments

• Correction for most recent weeks is likely affected by a number of factors – e.g. very large # deaths in a week -> longer delay than in the past, but as numbers drop shorter delay than the past.

• So initial under estimation more recently over estimation for most recent week.

• Coroner referral probably does not generally apply to COVID deaths – so they come through faster -> over correction in younger ages.

• Additional analyses at PHE include alternative baseline modelling, looking at daily all-cause excess and regional analysis. Also ONS coded data are assessed at PHE using daily data to look at excess with data on cause, place and other demographics on the death certificate and through linkage.
Hospital deaths data (not PHE)

Provided by NHS England - data are shown by trust and age group, gender, ethnicity, pre existing condition

All deaths occurring in any setting among people who have had a COVID-19 positive test result.

**Data are combined from three different sources:**
- Deaths occurring in hospitals, notified to NHS England by NHS trusts
- Deaths notified to local PHE Health Protection Teams in the course of outbreak management
- Laboratory reports where a person has had a laboratory confirmed COVID-19 test linked to death reports from electronic hospital records

Data from each source are validated and merged into single dataset, removing duplicate records.
Data / reports available

- Numbers are published (just overall total)
  https://coronavirus.data.gov.uk/about#covid-19-associated-deaths
- Full details on the data:
- PHE Surveillance report gives more details including ethnicity, age and sex. PHE are currently looking to update dashboard with actual numbers.
- More analysis is paper submitted for publication in Eurosurveillance
Deaths in laboratory confirmed COVID-19 by date of death: England, 2 March - 12 May 2020
Age/sex pyramid of COVID-19 deaths in England, 2 March to 12 May 2020
Daily deaths and proportion that occur in care homes, hospitals, and other locations, England

65+ year olds

Number of daily deaths:
- Total
- Care home
- Hospital
- Other

% of deaths in past 7d:
- 0
- 100

Dates:
- 02/03/2020
- 09/03/2020
- 16/03/2020
- 23/03/2020
- 30/03/2020
- 06/04/2020
- 13/04/2020
- 20/04/2020
- 27/04/2020
- 04/05/2020
- 11/05/2020
Daily deaths and proportion that occur in care homes, hospitals, and other locations, England

20-64 year olds

Number of daily deaths

% of deaths in past 7d

Total  Care home  Hospital  Other
Strengths and limitations

Strengths
- Mortality system developed rapidly and provides daily outputs
- Multiple source ascertainment minimises underreporting and reporting delay
- Capable of capturing deaths in all settings
- Deaths linkable to ethnicity and locally assigned information on residential deprivation
- Clear, consistent definition allows monitoring of trends

Limitations
- Excludes people who have died of COVID-19 who were not tested
- Doesn’t explain all of the impact of COVID-19 on excess mortality
- Between 13 March and 26 April UK testing policy limited to testing mainly people in hospital settings. This means it is possible that people who died in community or residential settings may have been underestimated.
Questions

Also thanks to Mary Sinnathamby and Alison Brown for help with slides
Reporting of COVID-19 deaths in Scotland

Scott Heald, Public Health Scotland
Julie Ramsay, National Records of Scotland
COVID-19 Statistics in Scotland - 1

3 main producer bodies in Scotland:
  • National Records of Scotland (NRS)
  • Scottish Government
  • Public Health Scotland (PHS)

Public Health Scotland was formed on 1\textsuperscript{st} April 2020, comprises:
  • Information Services Division (ISD)
  • Health Protection Scotland (HPS)
  • NHS Health Scotland

Across Scotland, statistics producers have worked together to ensure COVID-19 statistics are aligned and coherent:
  • Scottish Government publish daily statistics
  • NRS and PHS produce weekly reports
COVID-19 Statistics in Scotland - 2

Daily figures published by Scottish Government at 2pm, including statistics on:

- COVID-19 test numbers (and positive results)
- Deaths from COVID-19 (see next slide)
- Patients in hospital and ICU with COVID-19
- Calls to NHS24, COVID-19 helpline and Scottish Ambulance Service
- Care Homes

Weekly reports published by NRS and PHS every Wednesday at 12 noon:

- NRS – weekly overview of death statistics
- PHS – weekly overview of a range of statistics (including special topics – this week ethnicity)
Daily reporting of COVID-19 deaths

Figures are produced by PHS and published by Scottish Government

**Definition**
People with a lab-confirmed COVID-19 diagnosis who have died.
Deaths identified as follows:
- Prior to 2\(^{nd}\) April (when number of deaths was low), based on notification of deaths from NHS Boards after family liaison had occurred
- From 2\(^{nd}\) April (when numbers started to rise), based on death registrations (a more robust method as the number of deaths increases).

Figures continue to be published daily and differ from the NRS weekly figures because NRS include suspected (but not confirmed) cases.
Trend in daily lab-confirmed COVID-19 deaths
Cumulative deaths involving COVID-19 using different data sources
Deaths by week of registration

- All deaths 2020
- All deaths, average of previous 5 years
- COVID-19 deaths 2020

Week number: 1 to 19
Deaths: 0 to 2,500
Excess deaths by underlying cause

![Chart showing excess deaths by underlying cause across different weeks.](image-url)
COVID-19 death rate by SIMD Quintile

- Death rate from all causes is 1.9 times higher in the most deprived areas than the least deprived areas.
- Death rate from COVID-19 is 2.3 times higher in the most deprived areas than the least deprived areas.
COVID-19 death rate by urban/rural classification

- Large Urban Areas
- Other Urban Areas
- Accessible Small Towns
- Remote Small Towns
- Accessible Rural Areas
- Remote Rural Areas

Death rate per 100,000 population

All causes
COVID-19
Mortality reporting in Wales and estimation of total numbers of deaths using capture-recapture methods

Gareth John
NHS Wales Informatics Service
How are COVID-19 deaths reported in Wales?

2 primary sources of data for published COVID-19 deaths:

• Office for National Statistics registered deaths
• Public Heath Wales rapid surveillance

Other sources used within NHS Wales:

• Electronic Master Patient Index (eMPI)
• Mortality surveillance e-form
<table>
<thead>
<tr>
<th>Local Health Board of Residence</th>
<th>U071 Underlying</th>
<th>U071 Any mention</th>
<th>U071 or U072 Underlying</th>
<th>U071 or U072 Any mention</th>
<th>U071 any mention or MPI with +ive COVID19 test result</th>
<th>U071 or U072 any mention or MPI with +ive COVID19 test result</th>
<th>Public Health Wales dashboard*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Betsi Cadwaladr University Health Board</td>
<td>253</td>
<td>268</td>
<td>272</td>
<td>289</td>
<td>304</td>
<td>325</td>
<td>207</td>
</tr>
<tr>
<td>Aneurin Bevan University Health Board</td>
<td>396</td>
<td>420</td>
<td>414</td>
<td>443</td>
<td>427</td>
<td>450</td>
<td>260</td>
</tr>
<tr>
<td>Hywel Dda University Health Board</td>
<td>87</td>
<td>95</td>
<td>96</td>
<td>105</td>
<td>99</td>
<td>109</td>
<td>56</td>
</tr>
<tr>
<td>Cwm Taf Morgannwg University Health Board</td>
<td>282</td>
<td>304</td>
<td>339</td>
<td>369</td>
<td>324</td>
<td>389</td>
<td>256</td>
</tr>
<tr>
<td>Cardiff and Vale University Health Board</td>
<td>298</td>
<td>321</td>
<td>367</td>
<td>399</td>
<td>331</td>
<td>409</td>
<td>236</td>
</tr>
<tr>
<td>Powys Teaching Health Board</td>
<td>53</td>
<td>56</td>
<td>75</td>
<td>80</td>
<td>59</td>
<td>83</td>
<td>13</td>
</tr>
<tr>
<td>Swansea Bay University Health Board</td>
<td>215</td>
<td>232</td>
<td>246</td>
<td>266</td>
<td>236</td>
<td>270</td>
<td>193</td>
</tr>
<tr>
<td>Wales</td>
<td>1584</td>
<td>1696</td>
<td>1809</td>
<td>1951</td>
<td>1780</td>
<td>2035</td>
<td>1221</td>
</tr>
</tbody>
</table>
More information:

- Chief Statistician’s blog: “explaining COVID-19 mortality data sources for Wales”
  
  [Link to blog post](https://digitalanddata.blog.gov.wales/2020/05/05/chief-statisticians-update-explaining-covid-19-mortality-data-sources-for-wales/)
Providing timely death figures to support COVID-19 modelling work
Estimating numbers of death for the latest period using Capture-Recapture methods?

• “Capture and recapture” or “Mark and recapture” is a method commonly used in ecology to estimate an animal population's size, where it is impractical to count every individual.

• Involves an initial capture, mark, and release of a sample of animals, and then a repeat of the process, in which the proportion of animals caught in both samples is used to estimate the total population numbers.

• In the context of estimating numbers of deaths, we use the fact that we have our two different sources of death notifications, ONS and MPI.

• We equate
  • the number of animals captured in the first sample to the number of ONS Death records received (n)
  • the number of animals caught in the second sample to the number of MPI death notifications received (K)
  • the number of animals caught in both samples to the number of deaths independently notified by both the ONS and MPI (k).

• Given these assumptions, the most simple estimate for the total numbers of deaths (N) is given as follows:

\[
N \approx \mu \pm \sqrt{\mu \epsilon}
\]

\[
\mu = \frac{(K - 1)(n - 1)}{k - 2} \quad \text{for } k > 2
\]

\[
\epsilon = \frac{(K - k + 1)(n - k + 1)}{(k - 2)(k - 3)} \quad \text{for } k > 3
\]
Estimate for Wales deaths (all causes) as at 19th May 2020

- 12th May
  - Estimate = 115 (78-151)
  - Combined ONS/MPI = 77
  - ONS = 25

- 8th May
  - Estimate = 104 (93-115)
  - Combined ONS/MPI = 92
  - ONS = 75
Estimate for deaths in Wales (all causes) as at 12\textsuperscript{th} May 2020

- 5th May
  - Estimate = 110 (84-137) \textbf{112}
  - Combined ONS/MPI = 81 \textbf{109}
  - ONS = 31 \textbf{104}

- 2nd May
  - Estimate = 137 (123-149) \textbf{138}
  - Combined ONS/MPI = 119 \textbf{133}
  - ONS = 98 \textbf{127}
Conclusion and next steps

- The Capture-Recapture methodology offers a simple but seemingly robust and accurate way of tracking the number of deaths in a population in a timely fashion.

- It does not need the data sources to be complete, but it does require the records to be as representative a sample as possible, and for those two sources to be as independent as possible.

- It is difficult to ascertain how independent the MPI and ONS sources are, and we have seen that as the relative proportions of community to hospital deaths change, this can affect the accuracy of our estimates.

- The only sure way of taking into account independence is to introduce a third source of data, and there are well established 3-source capture-recapture methods, and a package in R (Rcapture), which are able to produce more sophisticated estimates and confidence intervals in these cases.
gareth.john@wales.nhs.uk
What local users need from mortality data on COVID-19

- Timeliness
- Comparability and consistency
- Granularity
- Content
- Flexibility
- Freedom to share