REPORT OF A BREACH OF THE CODE OF PRACTICE FOR STATISTICS

Government Statistical Service

1. Core Information

Title and link to statistical output	COVID-19 deaths of patients with a learning disability notified to the Learning Disabilities Mortality Review (LeDeR) https://www.england.nhs.uk/publication/covid-19-deaths-of-patients-with-a-learning-disability-notified-to-leder/
Name of producer organisation	**NHS England**
Name and contact details of person dealing with report	**Clare McConnell (clare.mcconnell1@nhs.net) Alice Luetchford (alice.luetchford@nhs.net)**
Name and contact details of Head of Profession for Statistics/Lead Official	**Mark Svenson (mark.svenson@nhs.net) Clare Backhouse (clare.backhouse1@nhs.net)**
Link to published statement about the breach (if relevant)	
Date of breach report	**31 July 2020**

2. Circumstances of breach

Relevant principle(s) and practice(s)	**Trustworthiness:** Principle T6: Data governance Practice T6.3 Organisations and those acting on their behalf should apply best practice in the management of data and data services, including collection, transmission, access and analysis.
Date of occurrence of breach	**30 July 2020**

Each week an aggregate Excel table showing the number of weekly deaths of patients with a learning disability is published by NHS England. The data relates to deaths notified to the LeDeR programme and identifies whether deaths are due to suspected or confirmed COVID-19.

The file was published at 12pm on 30 July as usual but, unfortunately, the wrong file had been sent to the NHS England web team in error. Instead of the publication file, the working file containing patient level data was sent and subsequently published to the webpage. The file contained potentially patient identifiable data including date of death, age, ethnicity, region and whether cause of death was due to confirmed/suspected COVID-19. There were 1,758 records in the file. In practice, it would be difficult to identify an individual from this data, due to the limited amount of information on individuals contained within the file.

The file for publication was created by the Senior Analytical Manager. It was then checked by a second analyst to ensure figures were correct and that small number suppression had been applied. At around 1.30pm on the day before publication (29 July 2020) the correct file was sent to the National Transformation Lead for LeDeR and the Senior Analytical Lead for sign off.

Following sign off, and at 9.18am on 30 July, the Senior Analytical Manager sent the file to the web team for publication (cc'ing the Senior Analytical Lead and the National Transformation Lead for LeDeR). Unfortunately, the wrong file was attached to the email, as described above. Neither the Senior Analytical lead, the National Transformation Programme lead, nor the Web team identified that the wrong file had been sent for publication.

Although the file for publication is stored in a separate folder, and the file names of the working file and the publication file are different, the files names do not make it clear that one file is a working file and the other file is the final file. The files were named:

- LeDeR deaths v1.0 up to 24.07.2020.xlsx
- LeDeR COVID-19 weekly deaths 24.07.2020.xlsx

This meant that the fact the wrong file was sent, wasn't immediately apparent to the web team or others cc'd into web team email.

The University of Bristol alerted NHS England at 8.35am on 31 July 2020 to the potential information governance breach. As soon as the breach was identified the NHS England web team were asked to take down the file. The NHS England web team confirmed that the file had been taken down at 9.15am.

The incident has been reported to NHS England's Information Governance team: the reference number of the IG security incident report is S2020-07-31 TO91528.

3. Impact of the breach

The data released could potentially identify individuals, although there were no names, dates of birth, addresses or postcodes released. Therefore, due to the limited amount of information on individuals contained within the file, the analytical team at NHS England concluded it would be difficult to identify an individual from this dataset.

The Information Governance team at NHS England/NHS Improvement reviewed the incident, which does not fall under General Data Protection Regulation (GDPR) or the Data Protection Act 2018 legislation, as the data relates only to the deceased. The Information Governance team concluded it would be unlikely that anyone could be directly or indirectly identified from the data, without unreasonable effort and therefore there is no breach in terms of the common law duty of confidentiality which continues to apply.

The data was live for just under 24 hours. There were 34 downloads of the data file since its publication on 30 July, before it was taken down.

The release of the data was tweeted by an individual who provided further analysis of the data in further tweets (e.g. numbers of deaths, % breakdowns by age and ethnicity). The individual tweeted four tweets at 4:02pm on 30 July.

The text of the tweets is detailed below:

> Today's weekly release of LeDeR notification information on COVID-19 deaths among people with learning disabilities (STILL no easy-read summary) has a lot more information than previous versions. https://england.nhs.uk/publication/covid-19-deaths-of-patients-with-a-learning-disability-notified-to-leder/

> First, it reports the actual numbers of people who have died, rather than rounding to the nearest 5. In total, in 2020 up to 24th July 648 people with learning disabilities have died a confirmed/suspected COVID-19 death and 1,110 people have died due to a different cause.

> Second, it has information on how old people were when they died. Of people dying COVID-19 related deaths: 1.6% were under 18; 1.6% were 19-25; 42.4% were 26-60; 35.2% were 61-74; and 19.3% were 75+. ¾

> Finally, it has information on ethnicity - the recorded ethnicity of people with learning disabilities dying COVID-19 related deaths was: Asian 6.8%; Black 3.7%;

Mixed/multiple ethnicity 0.8%; White 79.3%; 'Other' 0.5%; Ethnicity not recorded 9/0%.

The four tweets were retweeted a total of 80 times, as of 5 August 2020.

A final tweet was sent by the individual at 10:36am the following day, on 31 July:

> This version looks like it's been withdrawn from the website this morning. Might have been a simple administrative error (perhaps people would be potentially identifiable with very small numbers) but it would be good for NHS England to explain why.

This tweet was retweeted 3 times, as of 5 August 2020.

The individual was subsequently contacted by NHS England and was asked to delete the data. The individual confirmed deletion of the data.

4. Corrective actions (taken or planned) to prevent re-occurrence

As soon as the breach was identified the NHS England web team were asked to take down the file. The NHS England web team confirmed that the file had been taken down by 9.15am.

The NHS England analytical team discussed whether they should put a note on their web site explaining that the file had been removed, but instead immediately focused on rectifying the problem and alerting both their own Information Governance team and the Government Statistical Service to the incident. NHS England will link to this report on its web page once this report has been published.

File and folder names have been reviewed and changed to ensure files for publication are more readily distinguishable from data or working files in future.

Aggregate tables for publication or onward sharing will no longer be created in the same file as the record level data. Instead data will be drawn from the data files through links. This will mean that there is no way the record level data could accidently be published.

Emails to the web team containing files for publication will be drafted and sent to a senior analytical manager (e.g. the Senior Analytical lead) in the first instance. This person will check the appropriate file is attached prior to forwarding the publication email to the web team.

NHS England is considering if any further automation to the process could be undertaken to avoid the potential for human error.

NHS England will ensure the learning from this incident is shared with other NHS England producers of statistics, to avoid the same mistake happening anywhere else.