

# Deaths involving clostridium difficile: Wales QMI

Contact:

mortality@ons.gov.uk  
+44 (0)1633 455867

Release date:

3 October 2014


Next release:

To be announced

## Table of contents

1. [Methodology background](#)
2. [Overview](#)
3. [Executive summary](#)
4. [Output quality](#)
5. [About the output](#)
6. [How the output is created](#)
7. [Validation and quality assurance](#)
8. [Concepts and definitions](#)
9. [Other information](#)
10. [Sources for further information or advice](#)

# 1 . Methodology background

<b>National Statistic</b>	
<b>Survey name</b>	Deaths involving clostridium difficile: Wales QMI
<b>Frequency</b>	Annual
<b>How compiled</b>	Sample based survey
<b>Geographic coverage</b>	Wales
<b>Last revised</b>	3 October 2014

## 2 . Overview

- we have reported these statistics since 1999 to 2001 onwards
- the statistics are published in August each year (8 months after the reference period) information is supplied once a death is registered; each record is held on the Deaths Registration Database data is collected when Clostridium difficile (C. difficile) infection was mentioned in the death certificate of the underlying cause of death C. difficile is a spore forming bacterium, it can cause diarrhoea and severe inflammation of the bowel. Vulnerable people and people being treated with antibiotics are at the highest risk of getting C. difficile. It is normally spread on the hands of healthcare staff of other people that have been in contact with an infected patient.

The figures for Wales are broken down by sex, age group and place of death. They are based on deaths registered each calendar year, rather than those occurring in each year.

All deaths in Wales are coded according to the International Classification of Diseases (ICD). The specific code we use is A04.7 for Enterocolitis due to C. difficile. We have collated death certificates and the ICD codes since 1993, which has made it possible to identify the records where C. difficile is mentioned, but may not be coded.

The C. difficile statistics are used for monitoring trends in healthcare associated infections. They are used to inform policy, planning and research in both public and private sectors. This helps the policy makers and healthcare establishments to target their resources effectively.

## 3 . Executive summary

Clostridium difficile (C. difficile) is a spore-forming bacterium that can cause diarrhoea and severe inflammation of the bowel. It usually affects vulnerable people, particularly those who have been given antibiotic treatment, and is often spread on the hands of healthcare staff or other people who come into contact with infected patients. [Deaths involving Clostridium difficile, Wales](#) presents statistics on the number of deaths and the age-standardised mortality rates for deaths where C. difficile infection was mentioned on the death certificate.

[Deaths involving Clostridium difficile, Wales](#) is compiled using information supplied when a death is registered. The number of deaths where C. difficile is mentioned on the death certificate is extracted from our Deaths Registrations Database for Wales. The number of deaths and the mid-year population estimates we calculate are used to produce age-standardised mortality rates, standardised to the European Standard Population.

This report contains the following sections:

- Output quality
- About the output
- How the output is created
- Validation and quality assurance
- Concepts and definitions
- Other information, relating to quality trade-offs
- Sources for further information or advice

## 4 . Output quality

This report provides a range of information that describes the quality of the data and details any points that should be noted when using the output. We have developed [Guidelines for measuring statistical quality](#); these are based upon the five European Statistical System (ESS) quality dimensions. This report addresses these quality dimensions and other important quality characteristics, which are:

- relevance
- timeliness and punctuality
- accuracy
- coherence and comparability
- output quality trade-offs
- assessment of user needs and perceptions
- accessibility and clarity

\*Quality and Methodology Information (QMI) replaced Summary Quality Reports (SQR) from April 2011

More information is provided about these quality dimensions in the sections below.

## 5 . About the output

### Relevance

(The degree to which statistical outputs meet users' needs.)

C. difficile statistics are important for monitoring trends in healthcare-associated infections. The statistics are widely used to inform policy, planning and research in both public and private sectors and they enable policymakers and healthcare establishments to target their resources most effectively.

Figures on the number of deaths from *C. difficile* in Wales are used by various organisations, including the [Department of Health \(DH\)](#), [Public Health England \(PHE\)](#) and [Public Health Wales](#), for monitoring and evaluation purposes. For *C. difficile* infections occurring in Wales, surveillance of *C. difficile* is managed by the Welsh Healthcare Associated Infection Programme (WHAIP), which is part of Public Health Wales.

We have produced Deaths involving *C. difficile* statistics for 1999 and for 2001 onwards (see the Coherence and comparability section for more information). Figures for recent years show a large decrease in the number of deaths and the mortality rate for deaths where *C. difficile* was the underlying cause or was mentioned anywhere on the death certificate among both males and females. This finding is consistent with the incidence figures reported by PHE and Public Health Wales. These decreases may be due to the actions taken to reduce healthcare-associated infections described above.

*C. difficile* is a healthcare-associated infection, which means that it is likely to be picked up in communal establishments where patients are often being treated for other, unrelated illnesses. This means that there is a high level of interest in where these deaths occurred, despite the fact that death certificates only tell us where a person died and rarely specify the place where the infection was acquired.

Due to improvements in the classification and coding of communal establishments, the place of death definition used we use has been revised. In particular, the NHS and non-NHS nursing home and private residential home categories have been replaced with local authority and non-local authority care home categories, which reflects current user needs. The allocation and coding of individual establishments to place of death categories is a continual exercise that will improve the quality of this new classification.

Figures are presented in [Deaths involving Clostridium difficile, Wales](#) for Wales only combined. Sub-national figures are available on request and this is noted in the statistical bulletin.

## Timeliness and punctuality

(Timeliness refers to the lapse of time between publication and the period to which the data refer. Punctuality refers to the gap between planned and actual publication dates.)

The annual release of [Deaths involving Clostridium difficile, Wales](#) is announced on the GOV.UK [Statistics Release Calendar](#) 12 months in advance and on [our release calendar](#).

Statistics are published in August each year (8 months after the end of the reference period), following the release of the provisional annual death registrations data.

For more details on related releases, the GOV.UK [Statistics Release Calendar](#) and our release calendar provide 12 months' advance notice of release dates. In the unlikely event of a change to the pre-announced release schedule, public attention will be drawn to the change and the reasons for the change will be explained fully at the same time, as set out in the [Code of Practice for Official Statistics](#).

## 6 . How the output is created

[Deaths involving Clostridium difficile, Wales](#) is compiled using information supplied when a death is registered. A record for each death registered in Wales is held on our Deaths Registrations Database. This information goes through a series of quality assurance checks before it is considered ready for publication. These checks include manual and automated checks carried out by the Registration Service such as: ensuring that the death certificate is complete; that data fall within a realistic range of values; and that the death occurred within the last 12 months.

Upon receipt of the data, we carry out some basic checks including checking for missing or duplicate values, followed by more detailed validation checks; for example, checking for consistency between dates of birth, death and registration and between age and marital status. Further details about the information held on the Deaths Registrations Database and the methods used to quality assure the data can be found in [Mortality Statistics: Metadata](#).

We code all deaths in Wales according to the [International Classification of Diseases \(ICD\)](#). In the [10th Revision \(ICD-10\)](#), used from 2001 onwards, there is a specific code (A04.7) for Enterocolitis due to Clostridium difficile. While this code identifies the vast majority of deaths involving C. difficile, a small number of C. difficile-related deaths are not captured by this code alone.

Since 1993, we have stored the text of death certificates in addition to the ICD codes relating to causes identified on the death certificate. This means that it is possible to identify records where C. difficile is mentioned, but is not coded under the specific ICD-10 code.

In addition to extracting all deaths related to the specific ICD-10 code (A04.7), deaths mentioning a number of other ICD categories to which diseases including C. difficile could be coded are also extracted. The text of these death certificates is then searched manually for mentions of Clostridium difficile, C. difficile or pseudomembranous colitis. The ICD-10 codes used to select deaths in order to search manually are shown in Table 1.

**Table Specific and non-specific ICD-10 codes related to C. difficile**

<b>Specific codes <sup>1</sup></b>	<b>Non-specific codes <sup>1</sup></b>
A04.7 (Enterocolitis due to Clostridium difficile)	A05.8 (Other specified bacterial food borne intoxications)
	A41.4 (Septicaemia due to anaerobes, excludes gas gangrene)
	A48.0 (Gas gangrene: Clostridial; cellulites, myonecrosis)
	A49.8 (Other bacterial infections of unspecified site)
	P36.5 (Sepsis of newborn due to anaerobes)

Source: Office for National Statistics

Notes:

<sup>1</sup> Codes used to identify deaths where C. difficile was the underlying cause of death (on deaths where C. difficile was mentioned): A04.7, A09, A41.4 and A49.8.

Deaths with an underlying cause of C. difficile were identified by selecting those deaths with a mention of C. difficile that also had an underlying cause of one of the following ICD-10 codes: A04.7, A41.4 and A49.8. Death certificates that mention C. difficile and record the code A09 (diarrhoea and gastroenteritis of presumed infectious origin) as the underlying cause of death, are also taken to indicate that C. difficile was the underlying cause of death.

Deaths are extracted where C. difficile was coded either as the underlying cause of death or as a contributory factor and manual searching of the text of each death record is used to ensure that every mention is captured. Further information about the extraction of C. difficile deaths can be found in the Concepts and definitions section. Since information is held for all deaths registered, it is possible to extract actual counts of deaths, representative of the whole Wales population. No modelling or imputation of the number of deaths is necessary.

Since 1986, we have used the internationally recommended death certificate for neonatal deaths (infants under 28 days old). This certificate is designed to record all conditions found at death, but means that neonates cannot be assigned an underlying cause of death. However, as the data required are based on deaths where *C. difficile* is mentioned on the death certificate, neonates have been included. Neonatal deaths are extracted in the same way as described above for post-neonatal deaths.

Mortality rates are calculated using the number of deaths and latest mid-year population estimates provided by our Population Estimates Unit. Population estimates are based on the decennial UK census estimates and use information on births, deaths and migration to estimate the mid-year population in non-census years. Further information about the methods used to calculate mid-year population estimates can be found in the [Mid-year population estimates short methods guide](#).

The mortality rates published in Deaths involving *Clostridium difficile*, Wales are age-standardised rates. Age-standardised rates allow for differences of the age structure of populations and therefore allow valid comparisons to be made between geographical areas. The rates presented in Deaths involving *Clostridium difficile*, Wales are standardised using the direct method, which gives the mortality rate that would have occurred if the mortality rates observed had been applied to a standard population. Mortality rates are not calculated where there are fewer than three deaths in a cell. This is because rates based on such low numbers are susceptible to inaccurate interpretation. Rates that were calculated from less than 20 deaths are distinguished by italic type as a warning to the user that their reliability as a measure may be affected by the small number of events.

Age-standardised rates are calculated as follows:

$$\text{Age-standardised rate} = \frac{\sum (P_k m_k)}{\sum P_k}$$

Where: Standard population in sex/age group k Observed mortality rate (deaths per 100,000 persons) in sex/age group age/sex group 0, 1-4, 5-9, ... , 80-84, 85 years and over

Age-standardised rates are standardised to the European Standard Population. This is a hypothetical population and assumes that the age structure is the same in both sexes, therefore allowing comparisons to be made between the sexes as well as between geographical areas. The European Standard Population was first introduced in 1976 and its suitability as a standard population has not been reviewed since its introduction. Demographic changes since the population was developed may mean that it is no longer representative of the European population structure and work has begun to update it.

## Impact of the 2013 European Standard Population (ESP)

The ESP has become an accepted methodological standard in health statistics in the UK and the rest of Europe. It is used in the calculation of age-standardised rates by us, other government departments, the NHS and academic health researchers. The ESP used in the previous publications of *C. difficile* statistics was first introduced in 1976 but it has since been recognised that it is no longer representative of the age structure of the population of EU Member States. In light of this, [Eurostat implemented a new version of the ESP in 2013](#). In addition, we, on behalf of the Government Statistical Service (GSS), has carried out a [public consultation on how to implement the new ESP in the UK](#).

The 2013 ESP takes into account changes in the EU population, providing a more current, methodologically sound and widely acceptable basis for the calculation of age-standardised rates (Eurostat, 2013). The 1976 and 2013 ESPs differ in two ways. First, the 2013 ESP gives the populations in older age groups greater weighting than the 1976 ESP. Second, the age distribution of the 1976 ESP has an upper limit of 85 years and over, while the 2013 ESP is further disaggregated to include age groups 85 to 89 years, 90 to 94 years and 95 years and over.

An [ONS report examining the impact of the change in ESP on mortality data](#) showed that sex- specific rates, for causes where deaths predominantly occur at older ages, are significantly higher under the 2013 ESP compared with the 1976 ESP. This is because the larger number of older people in the 2013 ESP exerts more influence on such rates than the 1976 ESP. Since deaths involving C. difficile occur predominantly at older ages, the rates presented here are greater in magnitude than those previously published using the 1976 ESP for the same periods. However, it is important to note that the difference between death rates based on the old and new ESP is purely methodological and does not indicate an actual increase in previously published number of deaths or death rates. Historical data from 1999 onwards have also been recalculated using the 2013 ESP.

## Distribution of the European Standard Population

**Table 2: The 2013 European Standard Population**

Age group (years)	Population	Abridged version
Under 1	1,000	1,000
1 to 4	4,000	4,000
5 to 9	5,500	5,500
10 to 14	5,500	5,500
15 to 19	5,500	5,500
20 to 24	6,000	6,000
25 to 29	6,000	6,000
30 to 34	6,500	6,500
35 to 39	7,000	7,000
40 to 44	7,000	7,000
45 to 49	7,000	7,000
50 to 54	7,000	7,000
55 to 59	6,500	6,500
60 to 64	6,000	6,000
65 to 69	5,500	5,500
70 to 74	5,000	5,000
75 to 79	4,000	4,000
80 to 84	2,500	2,500
85 to 89	1,500	1,500
90 to 94	800	-
95 and over	200	-
90 and over	-	1000

Source: Eurostat

We provide a Microsoft Excel template that demonstrates how [age-standardised rates](#) are calculated.

## Modifications to standard error and confidence interval calculations

## Age-standardised rates

### Standard error

In previous publications, the standard error for age-standardised rates has been calculated using a simple approximation method as shown below. The standard error is denoted as SE(ASR) and calculated as:

$$SE(ASR) = \frac{ASR}{\sqrt{N}}$$

Where:

*ASR* is the age-standardised rate

*N* is the total number of deaths in all age groups in each year

The age-standardised rate is a weighted sum of age-specific death rates where the age-specific weights represent the relative age distribution of the standard population (in this case the 2013 ESP). Therefore, it is more accurate to calculate its variance as the sum of the age-specific variances and to estimate its standard error as the square root of the variance as shown below.

$$SE(ASR) = \sqrt{\frac{\sum_i \left( w_i^2 \cdot \frac{r_i^2}{d_i} \right)}{\left( \sum_i w_i \right)^2}}$$

Where:  $w_i$  is the number of individuals in the standard population in age group  $i$ ,  $r_i$  is the crude age-specific rate in the local population in age group  $i$ ,  $d_i$  is the number of deaths in the local population in age group  $i$

The standard error calculation has now modified such that it takes into account the form of the estimator, which is a weighted sum of age-specific rates.

### Confidence intervals

The mortality data in this release are not subject to sampling variation as they were not drawn from a sample. Nevertheless, they may be affected by random variation, particularly where the number of deaths or probability of dying is small. To help assess the variability in the rates, they have been presented alongside 95% confidence intervals.

The choice of the method used in calculating confidence intervals for rates will in part depend on the assumptions made about the distribution of the deaths data these rates are based on. Traditionally, a normal approximation method has been used in calculated confidence intervals on the assumption that *C. difficile* deaths are normally distributed. However, the annual number of deaths involving *C. difficile* are relatively small (usually fewer than 100) and may be assumed to follow a Poisson probability distribution. In such cases, it is more appropriate to use the confidence limit factors from a Poisson distribution table to calculate the confidence intervals instead of a normal approximation method.



In this bulletin, the method used in calculating confidence intervals for rates based on fewer than 100 deaths was proposed in [Confidence intervals for weighted sums of poisson parameters](#) and described in [a 2008 APHO technical briefing](#). In this method, confidence intervals are obtained by scaling and shifting (weighting) the exact interval for the Poisson distributed counts (number of deaths in each year). The weight used is the ratio of the standard error of the age-standardised rate to the standard error of the number of deaths. The lower and upper 95% confidence intervals are denoted as ASR lower and ASR upper, respectively, and calculated as:

$$ASR_{lower} = ASR + (D_l - D) \cdot \sqrt{\frac{v(ASR)}{v(D)}}$$

$$ASR_{upper} = ASR + (D_u - D) \cdot \sqrt{\frac{v(ASR)}{v(D)}}$$

Where:

$D_l$  and  $D_u$  are the exact lower and upper confidence limits for the number of deaths, calculated using confidence limit factors from a Poisson probability distribution table  $D$  is the number of deaths in each year  $v(ASR)$  is the variance of the age-standardised rate  $v(D)$  is the variance of the number of deaths Where there were 100 or more deaths the 95% confidence intervals for age-standardised rates were calculated using the normal approximation method:

$$ASR_{LL/UL} = ASR \pm 1.96 \cdot SE$$

Where:

$ASR_{LL/UL}$  represents the upper and lower 95% confidence limits, respectively, for the age- standardised rate

As a general rule, if the confidence interval around one figure overlaps with the interval around another, we cannot say with certainty that there is more than a chance difference between the two figures.

## Age-specific rates

For age-specific rates, the exact Poisson limit factors for the number of deaths was multiplied by the rate to calculate the 95% confidence intervals where there were fewer than 100 deaths in a particular age group.

$$LL(R) = L \cdot R \text{ and } UL(R) = U \cdot R$$

Conversely, the normal approximation method below was used where there were 100 or more deaths.

$$R_{LL/UL} = R \pm 1.96 \cdot \frac{R}{\sqrt{N}}$$

Where:

$LL$  and  $UL$  are the lower and upper 95% confidence limits, respectively  $R$  is the age-specific rate  $L$  and  $U$  are the exact lower and upper Poisson confidence limit factors for the age-specific number of deaths.

# 7 . Validation and quality assurance

## Accuracy

(The degree of closeness between an estimate and the true value.)

It is a legal requirement that all deaths are registered; however, the accuracy of statistics on mortality is dependent on the quality of information supplied when the death is registered. An incorrect cause of death may be provided by the doctor completing the death certificate or a contributory condition may be omitted. Many thousands of practising doctors complete death certificates and the nature and amount of training they have had in death certification varies greatly. Inaccurate information may also be supplied by the informant (usually a relative of the deceased) who must use the death certificate to register the death with the registrar. It is not possible to measure the magnitude of errors such as these.

[Guidance on death certification](#), with specific reference to healthcare-associated infections, was issued to doctors in May 2005 (revised in 2010). This was followed by [a message from the Chief Medical Officer](#) to all doctors reminding them of their responsibilities with respect to death certification and drawing their attention to the guidance. More recently, the Department of Health and Public Health England released a [report, Clostridium difficile infection: how to deal with the problem](#), detailing good practice and recommendations on completing death certificates for deaths involving C. difficile.

Deaths should be registered within 5 days from the date of death but, in some cases, the registration may be delayed so that registration may not take place in the same calendar year as the death occurred. Since mortality statistics are generally reported by the year of registration, this may lead to the under-reporting of deaths. This is most likely to occur in cases where the death is referred to a coroner and an inquest is held. Deaths are referred to a coroner in cases including where the cause of death is unknown, where the deceased was not seen by a doctor before or after death or where the death was violent, unnatural or suspicious. If the coroner chooses to hold an inquest, the death can only be registered once the inquest has taken place. Further information about the process of death registration can be found in [Mortality Statistics: Metadata](#).

The number of deaths due to C. difficile is difficult to estimate. Trends in mortality are normally monitored using the underlying cause of death (the disease that initiated the train of events leading directly to death). However, C. difficile (and other healthcare-associated infections) are often not the underlying cause of death. Those who die with C. difficile are usually patients who were already very ill, and it may be their existing illness, rather than C. difficile, which is designated as the underlying cause of death. There is therefore an interest in the number of deaths where C. difficile contributed to the death: only conditions that contribute directly to the death should be recorded on the death certificate. Results presented in [Deaths involving Clostridium difficile, Wales](#) identify deaths where the underlying cause was C. difficile and also where C. difficile was mentioned as the underlying cause or as a contributory factor in the death.

Information about the checks carried out to ensure the quality of the data we hold can be found in [Mortality Statistics: Metadata](#).

## Coherence and comparability

(Coherence is the degree to which data that are derived from different sources or methods, but refer to the same topic, are similar. Comparability is the degree to which data can be compared over time and domain, for example, geographic level.)

We produce statistics for C. difficile for Wales. The National Records of Scotland (NRS), formerly the General Register Office for Scotland (GROS) produces statistics for [Scotland](#), and the Northern Ireland Statistics and Research Agency (NISRA) produces statistics for [Northern Ireland](#).

In the UK, causes of death are coded according to the [ICD](#). In Wales, deaths are coded using [ICD-10](#) from 2001 onwards. While there is a specific code for deaths involving C. difficile in ICD-10, there is no equivalent ICD code in the ninth revision (ICD-9). Under ICD-9, any deaths related to C. difficile would most likely have been coded to some other unspecified infection, but it would be difficult to tell. Deaths registered in 1999 were coded to both ICD-9 and ICD-10 as part of a special study to compare the two ICD revisions, and have therefore been used to give an additional year of data on deaths involving C. difficile to add to information collected since the introduction of ICD-10 in 2001.

In [Deaths involving Clostridium difficile, Wales](#), deaths of non-residents are included in the overall figure for Wales. Non-residents are excluded where figures are presented for Wales separately so that the sum of the number of deaths in Wales separately does not equal the overall figure for Wales.

Deaths involving Clostridium difficile, Wales presents the number of cases where C. difficile is mentioned on the death certificate and the number of cases where C. difficile is the underlying cause of death. Since the number of cases where C. difficile is mentioned on the death certificate includes those cases where C. difficile is the underlying cause, these numbers cannot be summed.

## 8 . Concepts and definitions

(Concepts and definitions describe the legislation governing the output and a description of the classifications used in the output.)

ICD – a standard classification used by all World Health Organisation (WHO) member states to classify diseases and health problems on death certificates and for other types of health and vital statistics. We have used the 10th revision (ICD-10) since 2001.

Clostridium difficile (C. difficile) – a spore-forming bacterium that can cause diarrhoea and severe inflammation of the bowel. It usually affects vulnerable people, particularly those who have been given antibiotic treatment, and is often spread on the hands of healthcare staff or other people who come into contact with infected patients.

Underlying cause – defined by WHO as “the disease or injury which initiated the train of morbid events leading directly to death, or the circumstances of the accident or violence which produced the fatal injury”, in accordance with the rules of the International Classification of Diseases.

Mention – any reference to C. difficile on a death certificate regardless of whether either was also identified as the underlying cause of death.

## 9 . Other information

### Output quality trade-offs

(Trade-offs are the extent to which different dimensions of quality are balanced against each other.)

Due to the availability of mid-year population estimates, it is often necessary to calculate age- standardised mortality rates using population projections rather than final estimates. This means that the rates published in [Deaths involving Clostridium difficile, Wales](#)<sup>1</sup> need to be marked as provisional. In May 2010, we published revised mid-year population estimates for Wales so that revisions were required to the C. difficile rates published for 2002 to 2011 in light of the 2011 census. Rates may therefore differ from previous editions.

Deaths involving Clostridium difficile, Wales is published as early as possible each year to meet user need for timely figures. Users have indicated a preference for this balance between timeliness and quality.

## 10 . Sources for further information or advice

## Accessibility and clarity

(Accessibility is the ease with which users are able to access the data, also reflecting the format in which the data are available and the availability of supporting information. Clarity refers to the quality and sufficiency of the release details, illustrations and accompanying advice.)

Our recommended format for accessible content is a combination of HTML web pages for narrative, charts and graphs, with data being provided in usable formats such as CSV and Excel. We also offer users the option to download the narrative in PDF format. In some instances other software may be used, or may be available on request. Available formats for content published on our website but not produced by us, or referenced on our website but stored elsewhere, may vary. For further information please refer to the contact details at the top of the page.

For information regarding conditions of access to data, please refer to the following links:

- [terms and conditions \(for data on the website\)](#)
- [copyright and reuse of published data](#)
- [pre-release access \(including conditions of access\)](#)
- [accessibility](#)

In addition to this Quality and Methodology Information, basic quality information relevant to each release is available in the background notes of the relevant [statistical bulletin](#).