

Article

# Impact of coronavirus in care homes in England: 26 May to 19 June 2020

First results from the Vivaldi study, a large scale survey which looked at coronavirus (COVID-19) infections in 9,081 care homes providing care for dementia patients and the elderly in England.

Contact:  
Becky Tinsley  
infection.survey.analysis@ons.  
gov.uk  
+44 (0)203 973 4761

Release date:  
3 July 2020

Next release:  
To be announced

## Correction

### 3 July 2020 14:31

A correction has been made in Section 3 to the odds ratio for the odds of infection in staff working in care homes regularly using bank or agency nurses or carers. This was due to a small error detected when quality assuring the publication content. You can see the original content in the superseded version. We apologise for any inconvenience.

### 8 July 2020 08:19

A correction has been made in Section 3 to the odds ratio and confidence intervals for the odds of infection for residents in care homes where staff receive sick pay. This was due to a small error detected when quality assuring the publication content. You can see the original content in the superseded version. We apologise for any inconvenience.

# Table of contents

1. [Main points](#)
2. [Collaboration](#)
3. [COVID-19 Surveillance Study in Care Homes \(Vivaldi\)](#)
4. [Strengths and limitations](#)
5. [Differences with other publications](#)
6. [Next steps](#)
7. [Glossary](#)
8. [Related links](#)
9. [Annex A: Calculating estimates from the survey](#)
10. [Annex B: Calculating risk factors](#)

# 1 . Main points

Between 26 May and 20 June 2020, as part of the Vivaldi project, 9,081 care homes in England (all with responsibility for providing dementia care or care for older residents (65 years and over)) were surveyed to collect information on their staff, residents and each setting to help understand the impact of the coronavirus (COVID-19) in these care homes and inform the public health response; 5,126 (56%) care homes responded to the survey.

We estimate that across the 9,081 care homes in the study there are 293,301 residents (95% [confidence interval](#): 293,168 to 293,434) and 441,498 staff<sup>1</sup> (95% confidence interval: 441,240 to 441,756). These estimates were produced by weighting the actual responses to take account of the care homes who did not respond to the survey.

Across the care homes included in the study, we estimate that 56% (95% Confidence Interval: 55% - 56%) reported at least one confirmed case of coronavirus (staff or resident).

Across the care homes that reported at least one case of coronavirus, we estimate that 20% of residents tested positive for COVID-19 (95% confidence interval: 19% to 21%), as reported by care home managers, since the start of the pandemic.

Across the care homes that reported at least one confirmed case of coronavirus, we estimate that 7% of staff tested positive for COVID-19 (95% confidence interval: 6% to 8%), as reported by care home managers, since the start of the pandemic.

These emerging findings reveal some common factors in care homes with higher levels of infections amongst residents.

These include prevalence of infection in staff, some care home practices such as more frequent use of bank or agency nurses or carers, and some regional differences (such as higher infection levels within care homes in London and the West Midlands). There is some evidence that in care homes where staff receive sick pay, there are lower levels of infection in residents.

Findings also include some common factors in care homes with higher levels of infection amongst staff. These include prevalence of infection in residents (although this is weaker than the effect of staff infection on residents), some care home practices (such as more frequent use of bank or agency nurses or carers, and care homes employing staff who work across multiple sites) and some regional differences (such as higher infection levels within care homes in the North East and Yorkshire and the Humber). However regional differences may be affected by different patterns of testing in staff and residents over time.

Care home managers were asked to report the total number of confirmed cases of infection in their staff and residents since the start of the pandemic, which may exclude anyone who had COVID-19 but had not been tested at the time of reporting. Estimates are therefore likely to underestimate the proportion of staff and residents who were infected. Future publications will incorporate data from the results of swabs taken during the whole care home testing programme. This will help to address some of the limitations described in this article.

## Statistician's comment

"These are the first results from the Vivaldi study, a large-scale survey which looked specifically at infections in care homes which provide care for people with dementia and older people across England. From this we've estimated that over half of these care homes have had at least one confirmed case of COVID-19 amongst their staff and residents.

"Future work will include more detailed analysis and will incorporate COVID-19 test results from the whole care home testing programme."

## Notes for: Main points

1. This includes all staff who work in care homes, including roles such as cleaning, catering and admin.

## 2 . Collaboration

The results from the COVID-19 Surveillance Study in Care Homes were produced by the Office for National Statistics (ONS), in partnership with the Department of Health and Social Care, Ipsos MORI, University College London and Public Health England, powered by NHS Foundry.

## 3 . COVID-19 Surveillance Study in Care Homes (Vivaldi)

This study was commissioned by the Department of Health and Social Care, and aims to measure the prevalence of the coronavirus (COVID-19) in care homes and the use of disease control measures in each setting. This will inform decisions around the best approach to care home testing in the future. Information on the use of disease control measures will help local public health teams provide effective guidance to care homes.

### More about coronavirus

- Find the latest on [coronavirus \(COVID-19\) in the UK](#).
- All ONS analysis, summarised in our [coronavirus roundup](#).
- View [all coronavirus data](#).
- Find out how we are [working safely during the pandemic](#).

## The data being collected

The Vivaldi study covers 9,081 care homes in England, which cater for dementia and those aged over 65 years. These care homes were part of the [whole care home testing programme](#), through which staff and residents are provided with throat and nose swabs (administered by staff) to test whether they currently have the virus.

It is important to note that the focus of this study is homes that provide care to older residents and those with dementia, and not all care homes. Overall, there are currently 15,000 care homes that are regulated by the Care Quality Commission (CQC) in England and who cover different aspects of care, such as convalescent care, palliative care and disability care. This study is therefore not directly comparable with the [Deaths involving COVID-19 in the care sector](#) publication, which covers all 15,000 care homes in England and also covers Wales.

In addition to testing all staff and residents for the virus, a survey of the 9,081 care homes was conducted by IPSOS-MORI. This 20 to 30 minute telephone interview with all care home managers was designed to collect information on:

- the number of staff and residents in each care home
- care home characteristics
- the period prevalence of infection among staff and residents by care home
- information on the use of disease control measures by care homes

The survey attempted to obtain a response from all 9,081 care homes, no sampling was involved.

The survey was conducted over a four-week period, from 26 May to 19 June 2020. Of the 9,081 care homes, 56% (5,126) responded.

## **Analysing the data**

The results from this article focus only on the results obtained from the telephone interviews with care home managers, and have not yet been linked to swab results from staff and residents tested through the whole care home testing programme. Therefore the analysis here is based upon the responses from the care home manager, rather than confirmed test results or diagnoses of COVID-19 infections.

Care home managers were asked to report on all confirmed infections, so this would be expected to include test results from whole care home testing and other sources such as Public Health England (PHE) outbreak surveillance, or testing that was undertaken in hospital. Further, these responses may have taken place before all care home staff and residents were tested, so may only relate to those who were tested because they were showing symptoms. It will therefore reflect a different infection rate than the results of the whole care home testing programme. This will be explored in future analysis.

Using the information that care home managers provided, we are able to estimate the percentage of care home residents and staff who, according to the care home manager, have been tested positive for COVID-19. The estimates are adjusted (weighted) to ensure they are representative of the total population within all 9,081 care homes. [Annex A](#) provides more information on how the estimates were calculated.

## **Describing the care homes included in the study**

### **Number of residents and staff**

We estimate that across the 9,081 care homes in the study there are 293,301 residents (95% [confidence interval](#): 293,168 to 293,434) and 441,498 staff (95% confidence interval: 441,240 to 441,756). All staff working in care homes were included in this reporting so this includes those with roles such as cleaning, catering and admin.

### **Size of care homes**

The size of care homes varies across those included in this study. Table 1 shows the number of care homes, according to the number of beds, in each size band.

Table 1: Number of care homes by size (number of beds)

**Care home size (number of beds) Number of care homes**

0 to 40	5196
41 to 80	3390
81 to 120	436
121 to 160	43
More than 160	16

Source: Office for National Statistics

**Staff working arrangements**

Of the 9,081 care homes included in the study, we estimate that 93% (95% confidence interval: 93% to 93%) offer sick pay to their staff, 12% (95% confidence interval: 11% to 12%) of care homes have staff who work at more than one location, and 44% (95% confidence interval: 43% to 45%) do not employ any bank or agency staff. These results were reported as the circumstances since 1 March 2020.

**Care home closures**

Of the 9,081 care homes included in the study, we estimate that 97% (95% confidence interval: 97% to 98%) have been closed to visitors, while 19% (95% confidence interval: 19% to 20%) have been closed to new admissions.

We estimate that 1% of care home residents were in hospital at the time of the survey.

Table 2: Estimated proportion of care home characteristics and 95% confidence intervals

	<b>Estimated number (proportion)</b>	<b>95% confidence interval</b>	
		<b>Lower</b>	<b>Upper</b>
Proportion of care homes offering sick pay to staff	92.9%	92.5%	93.3%
Proportion of care homes closed to visitors	97.2%	96.8%	97.6%
Proportion of care homes closed to admissions	19.3%	18.5%	20.1%
Proportion of care home with staff at more than one location	11.5%	10.9%	12.1%
Proportion of care homes with no bank or agency staff	44.2%	43.4%	45.0%
Proportion of residents in hospital at the time of the survey	1.2%	1.0%	1.4%

Source: Office for National Statistics

**Coronavirus cases reported by care homes**

Of the 9,081 care homes included in the study, we estimate that 56% (95% Confidence Interval: 55% - 56%) reported at least one confirmed case of coronavirus (staff or resident).

Across those care homes where managers reported at least one case of coronavirus, we estimate that 20% (95% confidence interval: 19% to 21%) of residents in the care homes have tested positive for COVID-19, while 7% (95% confidence interval: 6% to 8%) of staff tested positive, as reported by care home managers, since the start of the pandemic. The level of infections seen in care home residents could be affected by pre-existing underlying health conditions, which are likely to be more prevalent among care home residents.

Across all 9,081 care homes, we estimate that 11% (95% confidence interval: 10% to 11%) of residents have tested positive for COVID-19, while 4% (95% confidence interval: 4% to 4%) of staff tested positive.

Care home managers were also asked how many residents had died having contracted COVID-19. We estimate this to be 15,606 deaths (95% confidence interval: 15,566 to 15,647).

This will differ to the number of deaths reported in the [Deaths involving COVID-19 in the care sector](#) publication because of differences in reporting periods, care home coverage and information source. For more information please refer to the [Differences with other publications](#) section and a [blog](#) we have released on care homes outlining the differences in the data.

Table 3: Estimated proportion of coronavirus cases reported by care homes, with 95% confidence intervals

	Estimated proportion (number)	95% confidence interval	
		Lower	Upper
Proportion of care homes with at least one case of coronavirus (staff or resident)	55.6%	54.8%	56.4%
Proportion of care home residents testing positive for COVID-19, in care homes with at least one case of coronavirus	19.9%	18.5%	21.3%
Proportion of care home residents testing positive for COVID-19 across all 9,081 care homes	10.7%	10.1%	11.3%
Proportion of care home staff testing positive for COVID-19, in care homes with at least one case of coronavirus	6.9%	5.9%	7.9%
Proportion of care home staff testing positive for COVID-19 across all 9,081 care homes	4.0%	3.6%	4.4%

Source: Office for National Statistics

## Factors affecting the risk of infection in care homes

Using the information collected by the survey, we are able to analyse which different factors may explain higher or lower levels of infection seen in care homes. [Annex B](#) provides more information about how these risk factors were identified. Here, we describe some of the main risk factors affecting the likelihood of infection in both residents and staff.

The forest plots in this section show a visual representation of the change in likelihood for these different factors when compared with a baseline category. When a factor has an [odds ratio](#) of 1, this means that there is neither an increase nor a decrease in the likelihood of infection compared with the baseline. An odds ratio of higher than 1 means that there is an increased likelihood of infection compared with the baseline. An odds ratio of lower than 1 means that there is a reduced likelihood of infection compared with the baseline.

## Care home residents

The following factors were found to increase the risk of infection in care home residents from the multivariable analysis:

- the number of infected staff members: for each additional member of infected staff working at the care home, the odds of infection for residents increase by 11% (95% confidence interval: 10% to 11%)
- the number of bank or agency nurses or carers employed by the care home: care homes using bank or agency nurses or carers most days or every day are more likely to have more cases in residents (odds ratio 1.58, 95% confidence interval: 1.50 to 1.65), compared with those care homes who never use bank or agency staff
- some regional differences: for this analysis, the level of infection was compared with the level observed in London; this analysis showed that residents who were at care homes in most regions had a lower chance of infection than those at care homes in London
- the only exception to this was for care homes in the West Midlands, where the odds of infection for residents were increased by 9% compared with London (95% confidence interval: 0% to 17%), this regional effect is likely to exist due to other factors, associated with region, which we have not yet incorporated into the models

Conversely, care homes in which staff receive sick pay are less likely to have cases of coronavirus in residents (odds ratio 0.87, 95% confidence interval 0.82 to 0.93), compared with those care homes where staff do not receive sick pay.

There are other factors that explain the level of infection seen in care homes, but they are not as strong as those reported here.

### **Figure 1: Odds of COVID-19 infection in care home residents by staff infection, region, use of bank or agency nursing staff, and staff working at other locations**

#### Notes

1. Estimated odds ratios are adjusted for care home size, closure status for new resident admissions, timing of closure to visitors, use of other bank or agency staff, Index of Multiple Deprivation, number of care homes in provider group, whether the care home pays sick pay, care home cleaning level, staff training level, PPE usage, and frequency of staff caring for both COVID-19 and non-COVID-19 residents.

[Download the data](#)

## Care home staff

The following factors were found to increase the chance of infection in care home staff:

- the number of infected residents: for each additional infected resident at the care home, the odds of infection for staff increase by 4% (95% confidence interval: 4% to 4%); this relationship is not as strong as the chance of infection for residents when staff are infected
- the number of bank or agency nurses or carers employed by the care home: care homes using bank or agency nurses or carers most or every day increase the odds of infection in staff (odds ratio 1.88, 95% confidence interval: 1.77 to 2.00), compared with those care homes not using bank or agency staff
- staff who regularly work elsewhere: care homes where staff regularly work elsewhere (most or every day) increase the odds of infection in staff (odds ratio 2.40, 95% confidence interval: 1.92 to 3.00) compared with those care homes who have staff who never work elsewhere
- some regional areas: for this analysis, the level of infection was compared with the levels observed in London, This shows a different picture to the chance of infection for residents; this analysis showed that staff who were at care homes in all regions outside of London had a higher odds of infection than those at care homes in London
- the regions with the most increased odds of infection for staff were the North East (odds ratio 3.86, 95% confidence interval: 3.38 to 4.41) and Yorkshire and The Humber (odds ratio 2.88, 95% confidence interval: 2.54 to 3.28); this regional effect is likely to exist because of other factors, associated with region, which we have not yet incorporated into the models

**Figure 2: Odds of COVID-19 infection in care home staff by resident infection, region, use of bank or agency nursing staff, and staff working at other locations.**

#### Notes:

1. Estimated odds ratios are adjusted for care home size, closure status for new resident admissions, timing of closure to visitors, use of other bank or agency staff, Index of Multiple Deprivation, number of care homes in provider group, whether the care home pays sick pay, care home cleaning level, staff training level, PPE usage, and frequency of staff caring for both COVID-19 and non-COVID-19 residents.

[Download the data](#)

## 4 . Strengths and limitations

A total of 9,081 care homes were invited to participate in this study, based on their selection into the whole care home testing programme (LaingBuisson, 2020<sup>1</sup> ). This is a census of care homes providing dementia care and care for the older people.

The response rate of care homes to the survey was 56% (5,126). This is a good response rate for a voluntary survey, when compared with response rates obtained from other voluntary surveys.

Table 4 shows the response rates by size of care home (number of beds). This indicates that the survey results are representative of all sizes of care homes in the sample.

Table 4: Response rates by size of care home (number of beds)

Care home size (number of beds)	Response rate
---------------------------------	---------------

0 to 40	56%
41 to 80	58%
81 to 120	57%
121 to 160	65%
More than 160	50%

Source: Office for National Statistics

We have also examined the response rates of care homes by local authority (LA). Although there is some regional variation, response rates are good. Some LAs have very small numbers of care homes. Overall, these results do not indicate any particular regional patterns or suggest any geographical bias within the results.

## Uncertainty in these data

The estimates presented in this bulletin contain [uncertainty](#). There are many sources of uncertainty, but the main sources in the information presented include each of the following.

### Quality of data collected in the questionnaire

As in any survey, some data can be incorrect or missing. For example, participants and interviewers sometimes misinterpret questions or skip them by accident, and these results were reliant on care home managers recalling information on behalf of their residents and staff, and recalling events in the past. (The reference period for the study was “since 1 March 2020” and the survey took place between 26 May and 20 June 2020.)

To minimise the impact of this, we clean the data, editing or removing things that are clearly incorrect. Because the survey was via telephone, and care home managers were extremely knowledgeable and responsive there is very little inconsistency in the data and limited editing was required. There were some missing responses, but this was minimal and no imputation was necessary.

Two topics were added midway through the survey. One was to ask questions about residents who had returned from hospital, and the other was to ask about whether the whole care home testing had taken place and on what date. Analysis involving these questions will take the reduced sample size into account.

The survey collection used a reference period rather than a point in time. Therefore, some of the responses will refer to a total number of incidences over that time period (for instance, number of positive cases). The estimates and modelling assume that any associations during that time period are negligible, whereas in reality there may be some causal effects – for instance, a number of positive infections at the beginning of the period may have led to actions taken within the care home such as additional preventative measures, which in turn reduced the likelihood of additional positive infections. This effect could potentially distort the relationships we report.

### Non-response bias

The results presented in this report are based on the 56% of care homes who responded to the survey. To ensure that they are representative of all care homes in the study, the results are based on weighted data, or used weighting factors in their calculation, to account for those who have not responded to the survey.

We apply weighting to ensure the responding sample is representative of all 9,081 care homes in terms of size, deprivation of local area, and provider. [Annex A](#) describes this process on more detail.

## Notes for: Strengths and limitations

1. LaingBuisson provided the data to produce the sampling frame. This ensured consistency with the whole care home testing programme. CareSearch, LaingBuisson (2020).

# 5 . Differences with other publications

## COVID-19 Infection Survey

In partnership with others, the Office for National Statistics (ONS) is conducting a pilot of the [COVID-19 Infection Survey](#).

To date, over 32,000 people in England have enrolled in the survey, with plans to extend this to over 300,000 over the next 12 months and look to cover people from all four UK nations. All individuals aged two years and over in sampled households are invited to provide nose and throat swabs for testing. Adults aged 16 years or older were also asked to give a sample of blood to test for antibodies to COVID-19 from around 10% of enrolled households.

Care homes are not classified as households, and are therefore not sampled by ONS surveys of private households. This means care home residents are not sampled by the COVID-19 Infection Survey, and are not represented in the published results. However, staff who work in care homes are residents of private households, and may also be sampled by the Infection Survey.

This means that there may be double-counting in the estimates of those testing positive in both studies. It is important therefore no inferences are made about the infection rates of private households and care homes by adding the statistics from the separate studies together. Further work will be done to consider how this may be improved.

## Deaths involving COVID-19 in the care sector

There are a number of differences between the death figures reported here and those reported in the [Deaths involving COVID-19 in the care sector](#) publication, these are outlined in this section.

### Definition of care home resident

- The “Deaths involving COVID-19 in the care sector” publication describes deaths involving COVID-19 in England and Wales; the term “care home residents” used in this article refers to all deaths where either the death occurred in a care home, or the death occurred elsewhere but the place of residence of the deceased was recorded as a care home.
- This is different from the term “care home residents” in the Vivaldi study described in this article, which refers only to those resident in a sub-sector of care homes (those providing dementia care or care for older people) at the time that the survey was taken; the Vivaldi study also only covers care homes in England.

### Reporting periods

- The “Deaths involving COVID-19 in the care sector” publication includes provisional deaths data occurring up to 12 June 2020 and registered up to 20 June 2020.
- The reference period for the Vivaldi study was “since 1 March 2020”, and the survey took place between 26 May and 20 June 2020.

## Coverage of care homes

- The “Deaths involving COVID-19 in the care sector” publication includes provisional deaths data occurring from 2 March 2020 up to 12 June 2020 and registered up to 20 June 2020
- The reference period for the Vivaldi study was “since 1 March 2020”, and the survey took place between 26 May and 20 June 2020.

## Source of information

- Data in the “Deaths involving COVID-19 in the care sector” publication are based on death occurrence (date of death), not date of registration and are obtained from details collected when deaths are certified and registered; more information can be found in the [Mortality statistics in England and Wales QMI](#).
- Data in the Vivaldi study are based on reported information provided by the care home managers).

Therefore, caution should be applied when drawing any inferences between these two publications.

## Coronavirus (COVID-19) related deaths by occupation, England and Wales

[Coronavirus \(COVID-19\) related deaths by occupation, England and Wales](#) reports on COVID-19 related deaths among health and social care workers. It does not report on the number of deaths specifically for care home staff as these are not listed separately in the [Standard Occupational Classification 2010 \(SOC 2010\)](#), which was used to define occupations in this publication. Therefore comparisons should not be made between the results from the Vivaldi care home study and the publication about COVID-19 related deaths by occupation.

## 6 . Next steps

This article will be followed with more detailed analysis in a forthcoming publication, which will also incorporate data from the results of swabs taken during the whole care home testing programme. This will help to explain more of the variability in the models, which identify factors associated with high or low levels of infection by including variables related to individuals, such as age and sex.

## Other studies

This is one of a suite of surveillance studies, which are being conducted to understand how many people have the coronavirus (COVID-19), and how the virus spreads in non-household populations (such as care homes and prisons) and other settings (such as schools and hospitals). This is complementary to the [COVID-19 Infection Survey](#), which aims to find out more about how many people in residential private households have the coronavirus in the UK.

The Vivaldi 2 study is a cohort study in staff and residents from over 100 care homes. This study is collecting serial blood samples and swabs to investigate the incidence and prevalence of current and past infection in this population. These results will be linked to individual-level data from residents and to information about care homes to investigate why some individuals become infected and why some care homes are more likely to have outbreaks than others.

## 7 . Glossary

### Confidence interval

A confidence interval gives an indication of the degree of uncertainty of an estimate and helps to decide how precise a sample estimate is. Confidence intervals give a range of values between which we are 95% certain that the true value lies. A wider interval indicates more uncertainty in the estimate. For more information, see our [latest methodology article on statistical uncertainty](#).

### Odds ratio

An odds ratio is a measure of association between a characteristic and an outcome. The odds ratio represents the likelihood that an outcome will occur given a particular characteristic, compared with the likelihood of the outcome occurring in the absence of that characteristic. The odds ratio can be used to determine whether a particular characteristic is a risk factor for a particular outcome, and to compare the magnitude of various risk factors for that outcome.

## 8 . Related links

### [Coronavirus \(COVID-19\) latest data and analysis](#)

Web page | Updated as and when data become available

Latest data and analysis on the coronavirus (COVID-19) in the UK and its effect on the economy and society.

### [Deaths registered weekly in England and Wales, provisional](#)

Bulletin | Weekly

Provisional counts of the number of deaths registered in England and Wales, including deaths involving the coronavirus (COVID-19), by age, sex and region, in the latest weeks for which data are available.

### [Deaths involving COVID-19 in the care sector](#)

Article | Released 3 July 2020

Provisional figures on deaths involving the coronavirus (COVID-19) within the care sector, in England and Wales.

### [COVID-19 Infection Survey pilot](#)

Bulletin | Weekly

Initial data from the COVID-19 Infection Survey, delivered in partnership with IQVIA, Oxford University and UK Biocentre.

## 9 . Annex A: Calculating estimates from the survey

## Sample design

All care homes that were part of the whole care home testing programme (that is, those which primarily provide dementia care or care services for those aged 65 years and over) were invited to take part in this survey; 9,081 care homes were contacted in total with 5,126 care homes responding.

Since the likelihood of care homes responding is variable based on certain characteristics of the care home, we have used weights to adjust for non-response when calculating the estimates. Characteristics of care homes identified to have an impact on response include: number of beds within the care home, deprivation of local area, and the number of care home locations run by the umbrella provider (if there is one). These characteristics were available for all the 9,081 care homes, regardless if they had responded or not.

## Post-stratification

Care homes have been grouped into post-strata based on combinations of the following:

Number of beds	Index of Multiple Deprivation	Number of care homes run by the provider
0 to 14	Below 20th centile	1
15 to 29	Above 20th centile	2 to 9
30 to 44		10 or more
45 to 59		
60 and over		

This results in 5 multiplied by 2 multiplied by 3 equals 30 post-strata. To ensure the non-response weights were robust, some post-strata were collapsed to achieve a minimum of 50 responding care homes in each stratum. This resulted in a total of 21 post-strata, each with a separate non-response weight adjustment contributing to the overall estimates.

## Method for estimating proportions

From an unweighted sample, a proportion ( $p$ ) would be estimated by  $x/n$ , where  $x$  is the number of care homes in the sample with the characteristic of interest, and  $n$  is the total number of care homes in the sample.

To produce the weighted estimates of proportions in this study we use the following calculation, where  $p_h$  refers to proportions estimated within a given post-stratum  $h$ ,  $N_h$  the total number of care homes within a given post-stratum  $h$ , and  $N$  the total number of care homes overall:

$$\hat{p}_{weighted} = \sum_{h=1}^H \frac{N_h}{N} \hat{p}_h$$

## Method for estimating population totals

Estimates of care home population totals are also reported in this analysis, for example, the number residents testing positive for COVID-19. Weights across post-strata have been used to adjust for care home non-response using the following calculation, where  $n_h$  refers to the number of care homes responding to the survey in a given post-stratum  $h$ ,  $N_h$  the total number of care homes within a given post-stratum  $h$ , and  $t_h$  the total population with the characteristic of interest from the survey responses in given post-stratum  $h$ .

$$\hat{t}_{weighted} = \sum_{h=1}^H \frac{N_h}{n_h} t_h$$

## 10 . Annex B: Calculating risk factors

The risk factors reported in this study have been derived from logistic regression models, using two approaches.

### A series of “univariable” models

These measure the extent to which individual questions asked in the survey (for example, provision of personal protective equipment (PPE) impact on the prevalence of COVID-19 in care homes.

The univariable risk factors give an indication of how prevalence is likely to increase or decrease when fitting models based on single variable that has been collected from the survey. A series of risk factors at univariable level have been produced for the prevalence of COVID-19 for both care home residents and staff.

### Multivariable models

These models explore the effects on prevalence of COVID-19 in care homes when multiple variables collected in the survey are combined simultaneously in a single model.

Risk factors estimated from the multivariable models highlight which variables have the biggest effects on COVID-19 prevalence when all predictors are being taken into account. Risk factors were selected based on theoretical assumptions about what may increase the spread (for example, region and use of bank workers), clinical importance of individual risk factors (for example, number of staff or residents testing positive and how PPE is being used), and statistical significance in the univariable models.

Various standard diagnostic measures were examined to provide assurance that the logistic regression model was appropriate and any model assumptions were well approximated. For instance, a variance inflation factor (VIF) was calculated for each risk factor to test the severity of multicollinearity, a measure of whether a factor is highly associated with another factor already present in the model. VIF identified no causes for concern. Risk factors at multivariable level have also been produced for both resident populations and staff populations in care homes.