



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



JAMDA

journal homepage: www.jamda.com

Original Study

Mortality of Care Home Residents and Community-Dwelling Controls During the COVID-19 Pandemic in 2020: Matched Cohort Study



Martin C. Gulliford FRCP, FFPH^{a,b,*}, A. Toby Prevost PhD^c, Andrew Clegg MD^{d,e},
Emma Rezel-Potts PhD^{a,b}

^a School of Population and Life Course Sciences, King's College London, Guy's Campus, London, United Kingdom

^b NIHR Biomedical Research Centre at Guy's and St Thomas' Hospitals London, Great Maze Pond, London, United Kingdom

^c Nightingale-Saunders Clinical Trials and Epidemiology Unit, Cicely Saunders Institute, King's College London, London, United Kingdom

^d Bradford Institute for Health Research, Bradford Teaching Hospitals NHS Foundation Trust, Bradford, United Kingdom

^e Academic Unit for Ageing and Stroke Research, University of Leeds, Leeds, United Kingdom

A B S T R A C T

Keywords:
COVID-19
SARS-CoV-2
mortality
care home
nursing home
primary care
pandemic

Objective: This study aimed to estimate and compare mortality of care home residents, and matched community-dwelling controls, during the COVID-19 pandemic from primary care electronic health records in England.

Design: Matched cohort study.

Setting and Participants: Family practices in England in the Clinical Practice Research Datalink Aurum database. There were 83,627 care home residents in 2020, with 26,923 deaths; 80,730 (97%) were matched on age, sex, and family practice with 300,445 community-dwelling adults.

Methods: All-cause mortality was evaluated and adjusted rate ratios by negative binomial regression were adjusted for age, sex, number of long-term conditions, frailty category, region, calendar month or week, and clustering by family practice.

Results: Underlying mortality of care home residents was higher than community controls (adjusted rate ratio 5.59, 95% confidence interval 5.23–5.99, $P < .001$). During April 2020, there was a net increase in mortality of care home residents over that of controls. The mortality rate of care home residents was 27.2 deaths per 1000 patients per week, compared with 2.31 per 1000 for controls. Excess deaths for care home residents, above that predicted from pre-pandemic years, peaked between April 13 and 19 (men, 27.7, 95% confidence interval 25.1–30.3; women, 17.4, 15.9–18.8 per 1000 per week). Compared with care home residents, long-term conditions and frailty were differentially associated with greater mortality in community-dwelling controls.

Conclusions and Implications: Individual-patient data from primary care electronic health records may be used to estimate mortality in care home residents. Mortality is substantially higher than for community-dwelling comparators and showed a disproportionate increase in the first wave of the COVID-19 pandemic. Care home residents require particular protection during periods of high infectious disease transmission.

© 2022 The Authors. Published by Elsevier Inc. on behalf of AMDA – The Society for Post-Acute and Long-Term Care Medicine. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

The protocol for the study was approved by the Clinical Practice Research Datalink (CPRD Independent Scientific Advisory Committee protocol number 20_000214).

The study is based in part on data from the CPRD obtained under license from the UK Medicines and Healthcare products Regulatory Agency. However, the interpretation and conclusions contained in this report are those of the authors alone. Requests for data sharing should be directed to the corresponding author.

The authors were supported by the National Institute for Health Research Biomedical Research Centre at Guy's and St Thomas' Hospitals and King's College London. The views expressed are those of the authors and not necessarily those of the National Health Service, the National Institute for Health Research, or the Department of Health. The funder of the study had no role in study design, data

collection, data analysis, data interpretation, or writing of the report. The authors had full access to all the data in the study and shared final responsibility for the decision to submit for publication.

All authors declare no support from any organization for the submitted work; no financial relationships with any organizations that might have an interest in the submitted work in the previous three years, and no other relationships or activities that could appear to have influenced the submitted work.

* Address correspondence to Martin C. Gulliford, FRCP, FFPH, School of Population and Life Course Sciences, King's College London, Guy's Campus, London SE1 1UL United Kingdom.

E-mail address: martin.gulliford@kcl.ac.uk (M.C. Gulliford).

<https://doi.org/10.1016/j.jamda.2022.04.003>

1525-8610/© 2022 The Authors. Published by Elsevier Inc. on behalf of AMDA – The Society for Post-Acute and Long-Term Care Medicine. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

The COVID-19 pandemic had major impacts during 2020.¹ The first wave of infections peaked during April 2020 in the United Kingdom (UK), with more than 1000 deaths per day within 28 days of a positive COVID-19 test. In the second wave, with more widespread testing, the number of people in the UK with a positive COVID-19 test result peaked at 81,525 on December 29, 2020.¹ Early studies identified deprivation,² household overcrowding,³ older age, male sex, obesity, comorbidity, and ethnic minority status as being important risk factors for severe disease and mortality.⁴ Residents of care homes, which in the UK include residential homes providing support with personal care, and nursing homes providing support with personal care and assistance from qualified nurses, were severely affected by the pandemic. Contributing factors included the discharge of hospital patients to care homes with risks of disease transmission,⁵ limited availability of COVID-19 testing,⁶ limited supply of personal protective equipment,⁷ and delayed development of guidance to ensure protection of the care home population.⁸ Data from the Office for National Statistics showed that weekly counts of deaths of care home residents in England and Wales increased from 2799 in the last week of February to 8476 and 9015 in the last 2 weeks of April 2020.⁹ Analysis of data reported to the Care Quality Commission in England suggested that excess deaths represented about 6.5% of care home beds.¹⁰ Care home residents typically have multiple risk markers of vulnerability for severe COVID-19, but susceptibility to infection may also have been increased because the care home environment had potential to facilitate transmission of COVID-19 and outbreaks were frequent. However, rigorous epidemiologic analysis has been limited. An editorial observed that “the COVID-19 pandemic has placed a spotlight on how little is known about this sector, and the lack of easily accessible, aggregated data on the UK care home population.”¹¹ To address this gap, we aimed to explore whether primary care electronic health records could be used to evaluate care home mortality during the pandemic.¹² We aimed to use primary care electronic health records to estimate all-cause mortality of care home residents in comparison with matched community-dwelling controls in England during 2020.

Methods

Data Source and Participant Selection

The study drew on data from the Clinical Practice Research Data-link (CPRD) Aurum database, a database of longitudinal primary care electronic health records in England,¹³ including a total of 1473 general practices in England with approximately 14.8 million registered patients at January 1, 2020. The protocol for the study was approved by the CPRD Independent Scientific Advisory Committee protocol number 20_000214.

This study used data from the March 2021 release of CPRD Aurum, including all 215,110 patients registered in CPRD Aurum general practices in England between January 1, 2015 and December 31, 2020 who were recorded as being resident in a care home. The most frequently recorded index care home codes were “lives in a nursing home [or] care home” (Supplementary Table 1). There were 28,531 (13%) patients with index codes of “patient died in a nursing home [or] care home.” For these, patients we assumed that they were resident in the care home for 90 days before death. The median length of stay is 2 years for care home residents, and 1 year for nursing home residents,¹⁴ but we assumed that patients with first codes for “died in nursing/care home” would have lower than average lengths of stay. In sensitivity analyses, we found that varying this assumed duration between 14 and 365 days had negligible influence on estimates. For each patient, the start date was the latest of the patient’s start of

registration or the first care home code. The end of the patient’s record was the earliest of the end of patient registration, the death date recorded by CPRD and the last data collection date for the practice. There were 7584 care home residents and 16,861 controls whose records were censored by end of registration rather than by death or last data collection date. We included patients age 18–104 years of age.

For 83,627 care home residents contributing person-time during 2020, a matched comparison cohort of community-dwelling adults was sampled from the list of all patients registered in the CPRD Aurum March 2021 release after excluding care home residents. Control patients were matched for general practice, sex, and year of birth, and had a start date that was no later than 18 months after the start date for matched cases. Up to 4 community-dwelling control participants were randomly sampled with replacement¹⁵ for each care home resident. Care home residents were omitted from this analysis where there were no eligible matched controls.

Main Measures

The primary measure of interest was mortality from any cause based on the CPRD death date. Covariates were age, sex, region in England, multiple morbidity, and frailty category. Age in 2020 and was divided into the age-groups of 18–64, 65–74, 75–84, 85–94, and 95–104 years. Multiple morbidity was represented by a count of 20 conditions, ever recorded in each patient’s record up to the end of 2020, from the list of atrial fibrillation, cancer, chronic kidney disease, chronic obstructive pulmonary disease, dementia, depression, diabetes mellitus, epilepsy, frailty fractures, heart failure, hemorrhagic stroke, hypertension, ischemic heart disease, ischemic stroke, other mental health diagnoses, peripheral arterial disease, palliative care, rheumatoid arthritis, or transient ischemic attack. Frailty was evaluated from coded records of deficits noted in CPRD Observation files according to the e-Frailty index, as described by Clegg et al.¹⁶ The e-Frailty index is informed by the cumulative deficit model of frailty and includes 36 deficits across physical, mental, cognitive, and social functioning. Coded records of frailty index scores and frailty index categories were also analyzed to inform frailty classification with the highest recorded value being employed.

Statistical Analyses

We initially analyzed eligible care home patient records between January 1, 2015 and December 31, 2020. We divided records into calendar months, calculating the number of deaths, and person time at risk for each month. We fitted a negative binomial regression model using data up to the end of 2019 as the training dataset, with counts of observed deaths as dependent variable and age-group, sex, region, multiple morbidity, frailty category, calendar month, and calendar year as predictors. Month was fitted as a categorical variable, while year was fitted as a continuous predictor. Multiple morbidity was fitted with categories from 1 to 9 or more morbidities, with a separate category for “none recorded.” Frailty category was fitted as a categorical variable. The categories of “nonfrail,” “mild frailty,” “moderate frailty,” and “severe frailty” were employed for analysis. Robust standard errors were employed to allow for general practice clustering. The general practice effect was allowed to differ between care home residents and controls, by representing the care home residents and community controls of each practice as separate clusters, because the former were clustered within care homes. We estimated predicted deaths by month for pre-pandemic and pandemic periods (2015–2020), comparing predicted and observed deaths graphically. To evaluate mortality in 2020 in more detail, we analyzed data for care home residents and community controls, evaluating counts of deaths

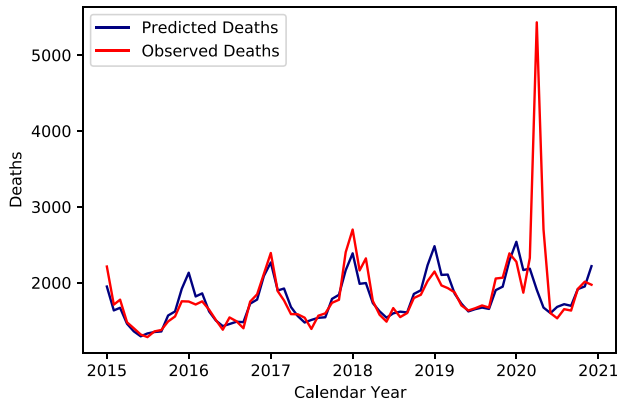


Fig. 1. Monthly counts of observed deaths of care home residents between 2015 and 2020 (red) with predicted deaths from Poisson model fitted to 2015 to 2019 data (blue).

and persons at risk by calendar week. We fitted a negative binomial model, with robust standard errors, now including interaction terms that allowed the associations of long-term conditions and frailty with mortality to differ between care home residents and community controls. To summarize the results, we fitted models separately for the periods of January to March, April, and May to December. However, we also present a difference-in-difference analysis that estimated the main effect of group (care home residence), time (January to March, April, and May to December) and the group by time interaction.

Analyses were performed using the “statsmodels”¹⁷ package in Python 3.8.3 (Python Software Foundation). The “matplotlib”¹⁸ package in Python and the “ggplot2”¹⁹ package in the R program (R Foundation for Statistical Computing) were used for data visualization.

Results

We analyzed data for 215,110 patients who were registered at general practices in England and were recorded as resident in a care home, who contributed follow-up between January 1, 2015 and December 31, 2020. There were 137,024 (64%) women; 97,192 (45%) were age 85–94 years and 24,685 (11%) were age 95 years or older; 180,390 (84%) had 2 or more morbidities. Figure 1 shows the distribution of observed deaths (red line) by month from 2015 to 2020, compared with predicted values estimated from 2015 to 2019 data (blue line). It was clear that there was a substantial excess of observed over predicted deaths in early 2020, with a peak in April 2020.

Analyses were then restricted to 83,627 care home residents who were registered during 2020, of whom 80,730 (97%) were matched with 300,445 community-dwelling controls. Characteristics of the sample are shown in Supplementary Table 2. Care home residents and community controls were similar with respect to matching variables of sex and age-group, but care home residents generally showed higher counts of long-term conditions and greater levels of frailty.

There was a peak in observed deaths between April 6, 2020 and April 26, 2020 (Figure 2, upper panel). Mortality rates were higher in men than women and increased in successive age-groups. The highest age-specific

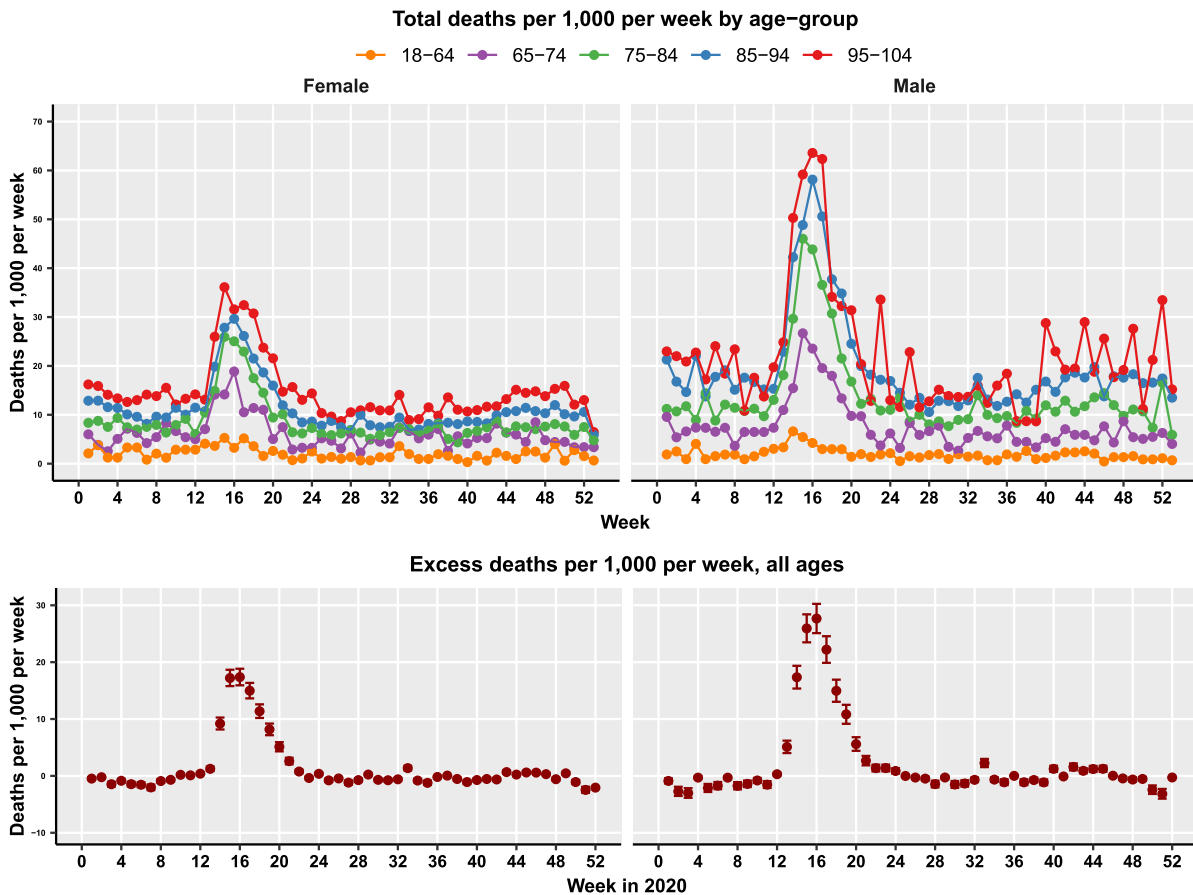


Fig. 2. Total deaths per 1000 patients per week during 2020 by age-group and sex (upper panel). Excess deaths (observed minus predicted) across all ages per 1000 patients per week (lower panel).

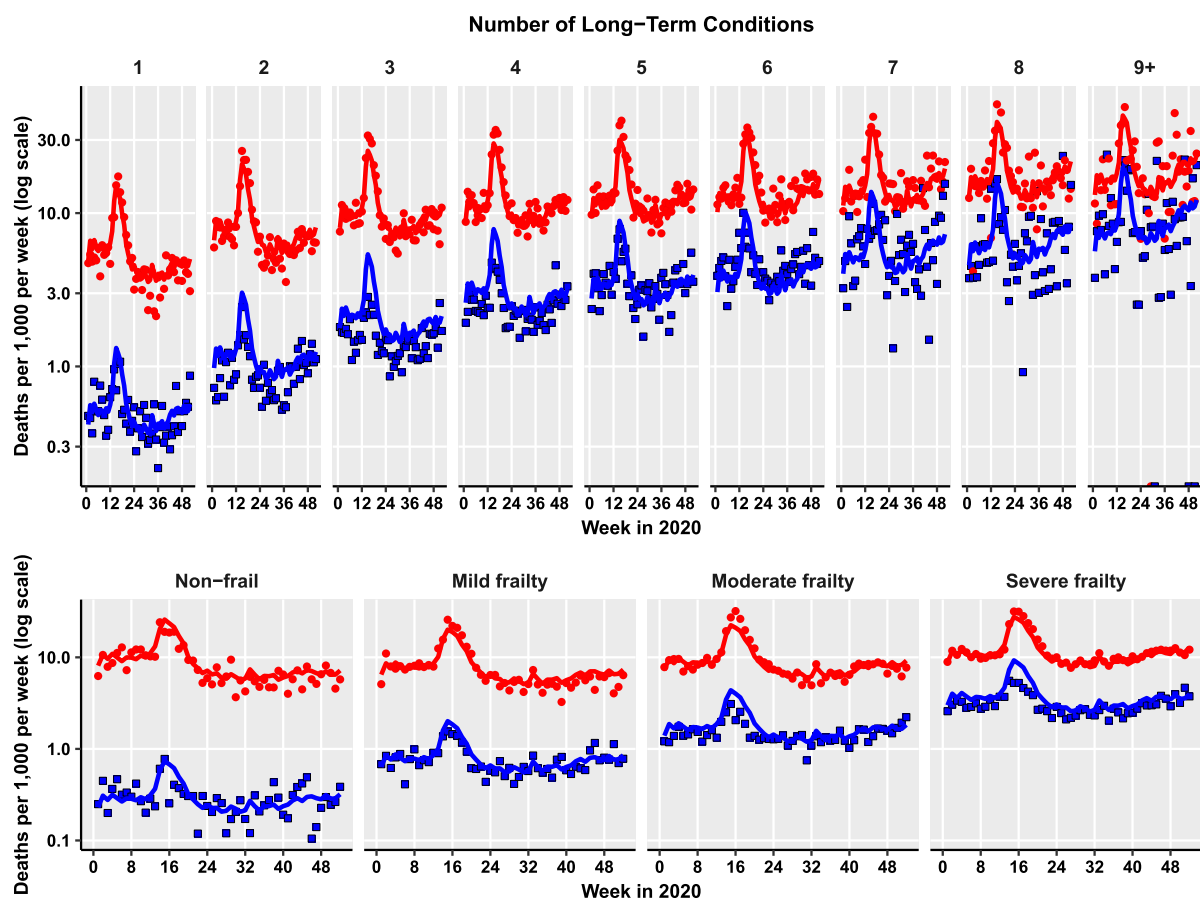


Fig. 3. Total deaths per 1000 patients per week during 2020 by number of long-term conditions (upper panel) and frailty category (lower panel). Red symbols, care homes residents; blue symbols, community dwelling controls. Points represent crude rates; lines represent predictions from adjusted regression model.

mortality rate was 63.6 (95% confidence interval 47.8–79.4) per 1000 patients per week in men age 95–104 years between April 13 and 19, 2020. Excess deaths, calculated as the difference between observed and predicted deaths, were summed across all age-groups (Figure 2, lower panel). Across all ages, excess deaths peaked in men between April 13 and 19, 2020 (men, 27.7, 95% confidence interval 25.1–30.3; women, 17.4, 15.9–18.8 per 1000 per week). During the first wave of the pandemic (weeks 12–24), there were an estimated 2125 excess deaths in men (137 per 1000 patients) and 2727 (89 per 1000 patients) in women.

Figure 3 shows mortality rates per 1000 patients per week for each week of 2020 for care home residents (red) and community controls (blue). Data are presented separately by number of long-term conditions (upper panel) and frailty category (lower panel). There was a peak in observed deaths between April 6, 2020 and April 26, 2020 that was evident in both care home residents and community controls. Mortality of care home residents was always higher than for community controls. Mortality also increased with number of long-term conditions and frailty category. However, the effect of increasing long-term condition count or frailty category was greater for community controls than for care home residents.

Table 1 shows data aggregated for the periods January to March, April, and May to December 2020. Mortality of care home residents was higher in the April period than the other periods; this increase was evident at each level of frailty with absolute risks of mortality increasing with frailty level. Mortality of community controls was also higher in April compared with the other periods; the increase was proportionately smaller than for care home residents but, in absolute terms, the increase was greatest for patients with the most advanced level of frailty. Comparing care home

residents and community controls, the adjusted relative mortality rate decreased with increasing level of frailty, reflecting the higher mortality of frail community controls. However, relative risks were higher in April period than in other periods of 2020.

Table 2 presents the unadjusted and adjusted estimates from a difference-in-difference analysis. After adjustment for covariates, the rate ratio for care home residence overall was 5.59 (5.23–5.99). Mortality for controls showed a 66% (55%–78%) relative increase during April 2020 compared with January to March 2020. After allowing for the underlying difference between care home residents and controls, and the increase shown by controls in April 2020, care home residents showed a further 76% (60%–93%) relative increase in mortality during April 2020.

Supplementary Figure 1 presents a forest plot of the adjusted mortality rate ratios, comparing care home residents with community controls. At each level of morbidity or frailty, the relative risk of mortality for care home residents was higher during the COVID-related peak of mortality in April 2020 compared with the mostly prepandemic period of January to March or later-pandemic period of May to December.

Discussion

Main Findings

This analysis shows that primary care electronic health records have potential to provide timely and relevant information concerning the care home population. There was evidence of a substantial

Table 1
Deaths and Person-Time for Care Home Residents and Matched Controls by Year

| Months | Frailty Category | Care Home Residents | | | Community-Dwelling Controls | | | Adjusted Rate Ratio* (95% CI) | | |
|-----------------------|---------------------------|---------------------|---------------------------|--------------------------|-----------------------------|---------------------------|--------------------------|-------------------------------|-------|-------|
| | | Deaths | Time at Risk (Patient Wk) | Rate per 1000 Patient Wk | Deaths | Time at Risk (Patient Wk) | Rate per 1000 Patient Wk | RR | LL | UL |
| January to March 2020 | All patients [†] | 6134 | 595,571.9 | 10.30 | 5291 | 3,737,963 | 1.42 | 6.22 | 5.79 | 6.69 |
| | Nonfrail | 323 | 29,897.0 | 10.80 | 257 | 776,819.7 | 0.33 | 36.04 | 28.10 | 46.22 |
| | Mild | 845 | 97,690.0 | 8.65 | 823 | 1,039,730.0 | 0.79 | 11.59 | 10.46 | 12.84 |
| | Moderate | 1474 | 159,954.7 | 9.22 | 1420 | 942,300.1 | 1.51 | 6.39 | 5.92 | 6.91 |
| | Severe | 3459 | 307,435.0 | 11.25 | 2740 | 846,803.6 | 3.24 | 3.55 | 3.28 | 3.83 |
| April 2020 | All patients [†] | 5169 | 189,975.7 | 27.2 | 2776 | 1,200,610 | 2.31 | 11.1 | 10.1 | 12.2 |
| | Nonfrail | 195 | 9971.8 | 19.56 | 125 | 253,107.3 | 0.49 | 44.78 | 31.74 | 63.19 |
| | Mild | 691 | 31,542.3 | 21.91 | 486 | 336,445.6 | 1.44 | 18.51 | 15.49 | 22.13 |
| | Moderate | 1385 | 51,132.3 | 27.09 | 776 | 301,813.3 | 2.57 | 12.21 | 10.67 | 13.97 |
| | Severe | 2871 | 97,116.7 | 29.56 | 1362 | 266,028.7 | 5.12 | 6.17 | 5.74 | 6.62 |
| May to December 2020 | All patients [‡] | 14,340 | 1,617,320 | 8.87 | 11,452 | 9,379,935 | 1.22 | 6.17 | 5.85 | 6.51 |
| | Nonfrail | 675 | 107,097.3 | 6.30 | 528 | 2,025,972.0 | 0.26 | 29.01 | 24.60 | 34.22 |
| | Mild | 1707 | 288,639.6 | 5.91 | 1849 | 2,664,259.0 | 0.69 | 9.96 | 9.04 | 10.97 |
| | Moderate | 3401 | 439,903.3 | 7.73 | 3266 | 2,348,550.0 | 1.39 | 6.19 | 5.74 | 6.66 |
| | Severe | 8496 | 779,945.0 | 10.89 | 5735 | 1,993,801.0 | 2.88 | 3.88 | 3.75 | 4.02 |

CI, confidence interval; LL, lower bounds of 95% confidence interval; RR, adjusted rate ratio; UL, upper bounds of 95% confidence interval.

*Care home residents compared to community controls, adjusted for age-group, sex, frailty, number of long-term conditions, region, month and log of patient months as offset.

[†]Includes patients with “not classified” frailty category (273 deaths and 525,420.3 person weeks).

underlying mortality difference between care home residents and community-dwelling controls that were matched for age, sex, and general practice. This difference persisted after further adjustment for frailty category, number of long-term conditions and region. We caution that, because of residual confounding from unmeasured and incompletely measured confounders, this analysis cannot determine to what extent the underlying mortality difference between care home residents and community controls is determined by the health status of residents, or the shared environment of the care home. Analyses quantified the first wave of COVID-19 mortality in April 2020 and showed that mortality peaked between April 6 and 26, 2020, being strongly associated with advanced age, male sex, multiple morbidity, and frailty category. Compared with community-dwelling control patients, mortality for care home residents was 4 to 5 times higher before the onset of the pandemic. Care home residents were disproportionately affected and during the month of April 2020 after allowing for differences in case-mix; mortality of care home residents was more than 10 times higher than for community-dwelling patients overall. Mortality remained high during the remainder of 2020 while the pandemic continued. The level of frailty and number of long-term

conditions were found to be effect modifiers, being more strongly associated with mortality of community-dwelling patients than those living in care homes.

Strengths and Limitations

We drew on a well-described database,¹³ and the quality of data offered by electronic health records has been shown to be generally high.²⁰ However, we acknowledge that there could be misclassification of care home status and it is possible that care home residence might be under-recorded. Community controls were matched on a small number of well recorded variables including age, sex, and general practice. Community controls were exactly matched with care home residents on year of birth, to allow for the important confounder of age, results were summarized over age groups. Controls might have been more closely matched for health status, but this might lead to problems of bias from over-matching. We compared unadjusted and covariate-adjusted estimates, as well as stratifying analyses by health status. We also acknowledge that limited testing for COVID-19, and recording of COVID-19 diagnoses, might have underestimated the

Table 2
Results of Poisson Regression Models Showing Main Effects of Group and Time Period and Group by Time Interaction

| | Unadjusted Rate Ratio | | P value | Adjusted Rate Ratio | |
|--|-----------------------|--|---------|---------------------|---------|
| | 95% CI | | | 95% CI | P Value |
| Net effect of care home residence during 2020 ^{*†} | | | | | |
| January: March 2020 | Ref. | | | Ref. | |
| April 2020 | 1.68 (1.53–1.85) | | <.001 | 1.76 (1.60–1.93) | <.001 |
| May: December 2020 | 1.00 (0.93–1.07) | | .966 | 1.05 (0.98–1.12) | .183 |
| Controls during 2020 [‡] | | | | | |
| January: March 2020 | Ref. | | | Ref. | |
| April 2020 | 1.64 (1.53–1.77) | | <.001 | 1.66 (1.55–1.78) | <.001 |
| May: December 2020 | 0.86 (0.81–0.91) | | <.001 | 0.90 (0.85–0.95) | <.001 |
| Overall difference between care home residents and controls [§] | | | | | |
| Control | Ref. | | | Ref. | |
| Care home residents | 7.38 (6.90–7.88) | | <.001 | 5.59 (5.23–5.99) | <.001 |

CI, confidence interval; RR, adjusted rate ratio.

Estimates were adjusted for age-group, sex, region, frailty category, and number of long-term conditions.

*Group by time interaction.

[†]Additional effect of care home residence during 2020, net of underlying difference between care home residence and controls and rate in controls in same period.

[‡]“time” effect.

[§]“group” effect.

burden of illness during the early stages of the pandemic. We addressed this by comparing the mortality of care home residents in 2020, with the mortality experienced in the preceding 5-year period (2015–2019). We also evaluated mortality for each week from January 1, 2020 onward. For control participants, the Office for National Statistics Coronavirus (COVID-19) Infection Survey showed that at the height of the first wave of infection from April 27 to May 10, 2020, an average of 0.27% (95% confidence interval: 0.17%–0.41%) of the general population had COVID-19.²¹ We did not have data concerning whether control participants were receiving social or nursing care support in their own homes, which might have been associated with frailty status. We included a count of important long-term conditions as well as analyzing frailty category. In the cumulative deficit model, frailty and multiple morbidity are closely related concepts,²² but more accurate phenotypic characterization of patients frailty status over time would have added to the study.²³ Deprivation is associated with reduced healthy life expectancy, which could lead to care home admission. Patients were matched for general practice, so it was not possible to adjust for deprivation at the general practice-level. We did not employ individual postcode-level deprivation scores as these might have presented difficulties if the care home postcode did not reflect deprivation exposures over the life-course. COVID-19 mortality is associated with deprivation, as well as age, but the effect of deprivation diminishes with age.²⁴ Ethnic minorities make up about 3% of the English population age 80 years and over²⁵ and, while ethnic minorities may be under-represented in care homes, mortality of minorities from COVID-19 was generally higher than in the white population.² Future studies should aim to include ethnicity and socioeconomic measures. Control sampling was with replacement and duplicated controls were included to reduce bias.¹⁵ Matching for family practice ensured that care home cases and community controls were resident in similar local areas and exposed to similar community prevalence of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection.

Comparison with Other Studies

Previous studies of care home mortality during the COVID-19 pandemic have mainly drawn on data from care home records.^{10,26} Morciano et al¹⁰ analyzed data for numbers of deaths reported to the care quality commission and estimated that over the first 7 months of 2020, deaths accounted for 6.5% of care home beds. The estimates from our analyses are not directly comparable because we estimated the mortality rate per 1000 residents per week. Dutey-Magni et al²⁶ analyzed data collected by care homes for incidence of COVID-19 and mortality. Their findings, like our study, suggested that deaths were frequent among residents who were probably infected with SARS-CoV-2 but were not tested. Burton et al²⁷ found that outbreaks of COVID-19 were frequent within care homes and most deaths occur in the context of outbreaks.^{10,27} In the United States, mortality in care homes was consistently associated with facility-size, community-incidence of COVID-19, and poverty.²⁸ We did not have data to identify individuals at the same care homes and the possible clustering of deaths at care homes could not be investigated in our data. Hollinghurst et al²⁹ analyzed linked primary care and administrative records for the population of Wales and found that care homes showed increased mortality during the first wave of the pandemic. Other studies confirm that background mortality is very high in care home residents. Vossius et al³⁰ found that annual mortality of nursing home residents was 31.8%. Shah et al³¹ analyzing the The Health Improvement Network primary care database for 2009 found that the age and sex standardized mortality ratio for nursing home residents was 419 and for residential home residents was 284, consistent with the elevated relative rates observed in the present analyses.

Conclusions and Implications

This study shows that individual-patient data from primary care electronic health records may be used to estimate mortality in care home residents in comparison with community-dwelling comparators. Mortality of care home residents is substantially higher than for community-dwelling comparators and showed a disproportionate increase in the first wave of the COVID-19 pandemic. Care home residents require particular protection during periods of high infectious disease transmission.

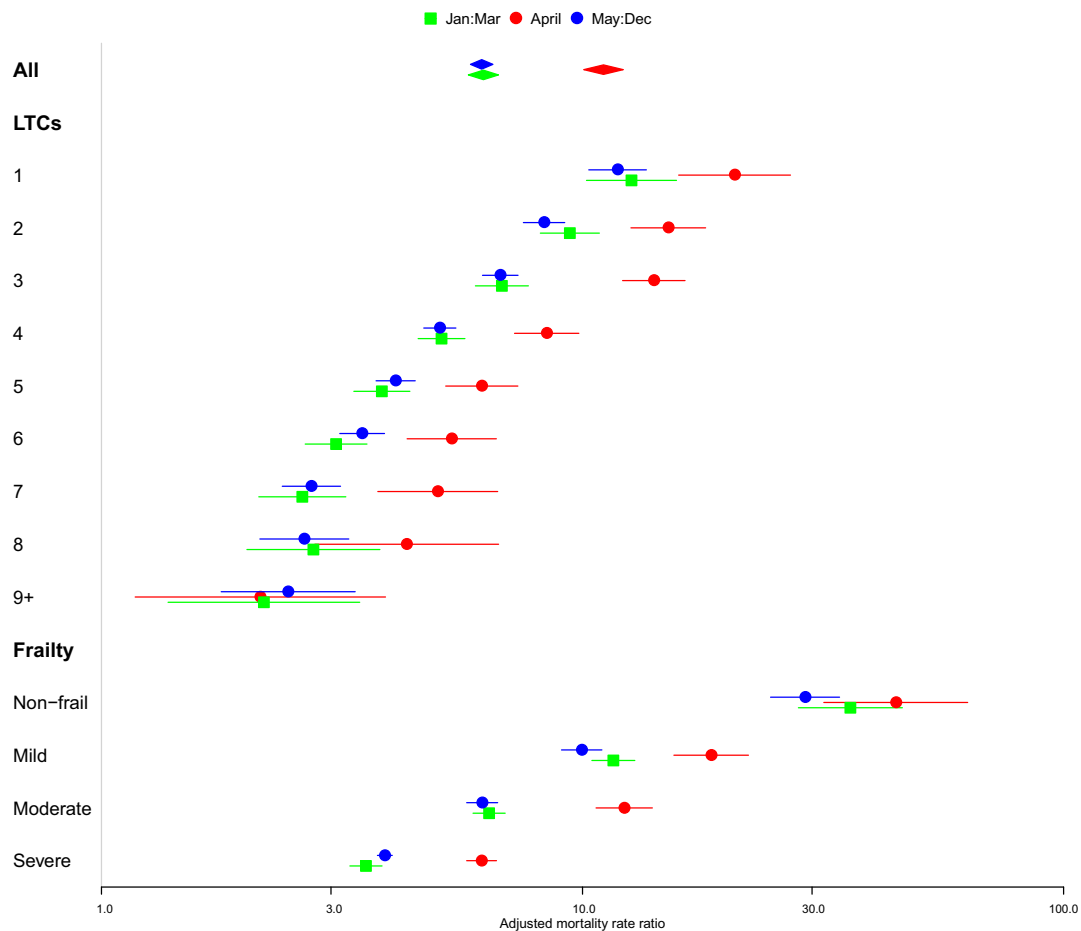
Acknowledgments

MCG wrote the study protocol and conducted data analyses; ERP provided advice on study design, data analysis and interpretation. ATP advised on statistical modeling and interpretation of results. AC advised on frailty classification and interpretation of results. All authors contributed to drafting the article and approved the final draft. MCG is guarantor.

References

1. UK Government. Coronavirus (COVID-19) in the UK. London: UK Government. 2021. <https://coronavirus.data.gov.uk/>. Accessed May 11, 2022.
2. Patel AP, Paranjpe MD, Kathiresan NP, et al. Race, socioeconomic deprivation, and hospitalization for COVID-19 in English participants of a national biobank. *Int J Equity Health*. 2020;19:114.
3. Harris R. Exploring the neighbourhood-level correlates of COVID-19 deaths in London using a difference across spatial boundaries method. *Health Place*. 2020;66:102446.
4. Williamson EJ, Walker AJ, Bhaskaran K, et al. Factors associated with COVID-19-related death using OpenSAFELY. *Nature*. 2020;584:430–436.
5. NHS England. Next steps on NHS response to COVID-19. London: NHS England. 2020. <https://www.england.nhs.uk/coronavirus/wp-content/uploads/sites/52/2020/03/20200317-NHS-COVID-letter-FINAL.pdf>. Accessed May 11, 2022.
6. Oliver D. Let's be open and honest about covid-19 deaths in care homes. *BMJ*. 2020;369:m2334.
7. Burki T. England and Wales see 20,000 excess deaths in care homes. *The Lancet*. 2020;395:1602.
8. Department for Health and Social Care. Coronavirus (COVID-19): adult social care action plan. London: Department for Health and Social Care. 2020. <https://www.gov.uk/government/publications/coronavirus-covid-19-adult-social-care-action-plan>. Accessed May 11, 2022.
9. Office for National Statistics. Care home resident deaths registered in England and Wales, provisional. London: Office for National Statistics. 2021. <https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/datasets/carehomeresidentdeathsregisteredinenglandandwalesprovisional>. Accessed May 11, 2022.
10. Morciano M, Stokes J, Kontopantelis E, et al. Excess mortality for care home residents during the first 23 weeks of the COVID-19 pandemic in England: a national cohort study. *BMC Med*. 2021;19:71.
11. Hanratty B, Burton JK, Goodman C, et al. COVID-19 and lack of linked datasets for care homes. *BMJ*. 2020;369:m2463.
12. Jain A, van Hoek AJ, Walker JL, et al. Identifying social factors amongst older individuals in linked electronic health records: an assessment in a population based study. *PLoS One*. 2017;12:e0189038.
13. Wolf A, Dedman D, Campbell J, et al. Data resource profile: Clinical Practice Research Datalink (CPRD) Aurum. *Int J Epidemiol*. 2019;48:1740.
14. Forder J, Fernandez JL. Length of stay in care homes. Report commissioned by Bupa Care Services, PSSRU Discussion Paper 2769. Canterbury: PSSRU. 2011. <https://eprints.lse.ac.uk/33895/1/dp2769.pdf>. Accessed May 11, 2022.
15. Heide-Jørgensen U, Adelborg K, Kahlert J, et al. Sampling strategies for selecting general population comparison cohorts. *Clin Epidemiol*. 2018;10:1325–1337.
16. Clegg A, Bates C, Young J, et al. Development and validation of an electronic frailty index using routine primary care electronic health record data. *Age Ageing*. 2016;45:353–360.
17. Seabold SPJ. Statsmodels: econometric and statistical modeling with python. Proceedings of the 9th Python in Science Conference. 2010. <https://www.statsmodels.org/stable/index.html>. Accessed May 11, 2022.
18. Hunter JD. Matplotlib: a 2D graphics environment. *Computing Sci Eng*. 2007;9:90–95.
19. Wickham H. *ggplot2: Elegant graphics for data analysis*. Heidelberg: Springer; 2009.
20. Herrett EL, Thomas SL, Smeeth L. Validity of diagnoses in the general practice research database. *Br J Gen Pract*. 2011;61:438–439.
21. Office for National Statistics. Coronavirus (COVID-19) Infection Survey pilot: England, 14 May 2020. London: Office for National Statistics. 2020. <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/>

- conditionsanddiseases/bulletins/coronaviruscovid19infectionsurvey/pilot/england14may2020. Accessed May 11, 2022.
22. Clegg A, Young J, Iliffe S, et al. Frailty in elderly people. *The Lancet*. 2013;381:752–762.
 23. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. *J Gerontol Series A Biol Sci Med Sci*. 2001;56A:M146–M157.
 24. Tinson A. What geographic inequalities in COVID-19 mortality rates and health can tell us about levelling up. London: Health Foundation. 2021. <https://www.health.org.uk/news-and-comment/charts-and-infographics/what-geographic-inequalities-in-covid-19-mortality-rates-can-tell-us-about-levelling-up#:~:text=The%20'deprivation%20gap'%20in%20COVID,those%20aged%2065%20and%20older>. Accessed May 11, 2022.
 25. NOMIS Official Labour Market Statistics. LC2109 - Ethnic group by age. London: Office for National Statistics. 2011. https://www.nomisweb.co.uk/census/2011/LC2109EWLS/view/2092957703?rows=c_age&cols=c_ethpuk11. Accessed May 11, 2022.
 26. Dutey-Magni PF, Williams H, Jhass A, et al. COVID-19 infection and attributable mortality in UK care homes: Cohort study using active surveillance and electronic records (March–June 2020). *medRxiv*. 2021. 2020.07.14.20152629.
 27. Burton JK, Bayne G, Evans C, et al. Evolution and effects of COVID-19 outbreaks in care homes: a population analysis in 189 care homes in one geographical region of the UK. *The Lancet Healthy Longevity*. 2020;1:e21–e31.
 28. Abrams HR, Loomer L, Gandhi A, et al. Characteristics of U.S. Nursing Homes with COVID-19 Cases. *J Am Geriatr Soc*. 2020;68:1653–1656.
 29. Hollinghurst J, Lyons J, Fry R, et al. The impact of COVID-19 on adjusted mortality risk in care homes for older adults in Wales, UK: a retrospective population-based cohort study for mortality in 2016–2020. *Age Ageing*. 2020;50:25–31.
 30. Vossius C, Selbæk G, Šaltytė Benth J, et al. Mortality in nursing home residents: A longitudinal study over three years. *PLoS One*. 2018;13:e0203480.
 31. Shah SM, Carey IM, Harris T, et al. Mortality in older care home residents in England and Wales. *Age Ageing*. 2013;42:209–215.



Supplementary Fig. 1. Forest plot showing adjusted mortality rate ratios, comparing care home residents and community controls, at each level number of long-term conditions and frailty category. January to March 2020 (green); April 2020 (red); May to December 2020 (blue).

Supplementary Table 1

First Recorded Medical Codes for Care Home Residence for 215,110 Patients Registered from 2015 to 2020 (14 Most Frequent Codes Shown).

| Medical Terms | Frequency |
|--|-----------|
| Lives in a nursing home | 78,897 |
| Lives in care home | 54,217 |
| Patient died in nursing home | 25,652 |
| Care home visit | 9383 |
| Seen in nursing home | 8198 |
| Care home visit for initial patient assessment | 6724 |
| Nursing home visit note | 5715 |
| Care home enhanced services administration | 5054 |
| Care home visit for follow-up patient review | 4097 |
| Weekly care home ward round | 3509 |
| Patient died in care home | 2879 |
| Admission to nursing home | 2717 |
| Nursing home | 2642 |
| Discharge to nursing home | 1312 |
| All other codes | 4114 |

Supplementary Table 2

Characteristics of Participants

| Variables | Category | Care Home Residents | Community Controls |
|------------------|---------------|---------------------|--------------------|
| Total | | 80,730 | 300,445 |
| Gender | Female | 51,873 (64) | 191,523 (64) |
| | Male | 28,857 (36) | 108,922 (36) |
| Age-group (y) | 18–64 | 9250 (11) | 36,801 (12) |
| | 64–74 | 8234 (10) | 32,699 (11) |
| | 74–84 | 20,777 (26) | 81,402 (27) |
| | 84–94 | 34,277 (42) | 126,977 (42) |
| | 94–104 | 8190 (10) | 22,563 (8) |
| Number of LTCs | 1 | 10,466 (13) | 64,481 (21) |
| | 2 | 14,264 (18) | 64,935 (22) |
| | 3 | 16,199 (20) | 51,240 (17) |
| | 4 | 14,448 (18) | 33,622 (11) |
| | 5 | 10,084 (12) | 18,773 (6) |
| | 6 | 5986 (7) | 8996 (3) |
| | 7 | 3052 (4) | 3849 (1) |
| | 8 | 1288 (2) | 1387 (0) |
| | 9+ | 614 (1) | 560 (0) |
| Frailty category | None recorded | 4349 (5) | 52,602 (18) |
| | Nonfrail | 5284 (7) | 61,829 (21) |
| | Mild | 13,467 (17) | 82,645 (28) |
| | Moderate | 21,400 (27) | 75,600 (25) |
| | Severe | 40,367 (50) | 69,337 (23) |
| | Not recorded | 212 (0) | 11,034 (4) |

Figures are frequencies (percent of column total).