Age and Ageing 2022; **51:** 1–10
https://doi.org/10.1093/ageing/afac250

RESEARCH PAPER

The impact of dementia, frailty and care home characteristics on SARS-CoV-2 incidence in a national cohort of Welsh care home residents during a period of high community prevalence

Chris Emmerson[1], Joseph Hollinghurst[2], Laura North[2], Richard Fry[2], Ashley Akbari[2], Ciarán Humphreys[1], Mike B. Gravenor[3], Ronan A. Lyons[2]

[1]Public Health Wales, Cardiff, UK
[2]Population Data Science and Health Data Research UK, Swansea University, Swansea, UK
[3]Swansea University Medical School, Swansea University, Swansea, UK

Address correspondence to: Chris Emmerson. Email: chris.emmerson@wales.nhs.uk

Abstract

Background: dementia may increase care home residents' risk of COVID-19, but there is a lack of evidence on this effect and on interactions with individual and care home-level factors.
Methods: we created a national cross-sectional retrospective cohort of care home residents in Wales for 1 September to 31 December 2020. Risk factors were analysed using multi-level logistic regression to model the likelihood of SARS-CoV-2 infection and mortality.
Results: the cohort included 9,571 individuals in 673 homes. Dementia was diagnosed in 5,647 individuals (59%); 1,488 (15.5%) individuals tested positive for SARS-CoV-2. We estimated the effects of age, dementia, frailty, care home size, proportion of residents with dementia, nursing and dementia services, communal space and region. The final model included the proportion of residents with dementia (OR for positive test 4.54 (95% CIs 1.55–13.27) where 75% of residents had dementia compared to no residents with dementia) and frailty (OR 1.29 (95% CIs 1.05–1.59) for severe frailty compared with no frailty). Analysis suggested 76% of the variation was due to setting rather than individual factors. Additional analysis suggested severe frailty and proportion of residents with dementia was associated with all-cause mortality, as was dementia diagnosis. Mortality analyses were challenging to interpret.
Discussion: whilst individual frailty increased the risk of COVID-19 infection, dementia was a risk factor at care home but not individual level. These findings suggest whole-setting interventions, particularly in homes with high proportions of residents with dementia and including those with low/no individual risk factors may reduce the impact of COVID-19.

Keywords: COVID-19, dementia, care homes, SARS-CoV-2, frailty, older people

Key Points

- Multi-level modelling of SARS-CoV-2 testing in care homes suggests setting is more strongly associated with infection than individual risk factors.
- Risks of COVID-19 infection appear higher in care homes with higher proportions of residents with dementia.
- Analysis of a national cohort of care home residents in Wales suggests dementia diagnosis does not increase risk of COVID-19.

Background

Residents of care homes in the UK were recognised at an early stage in the COVID-19 pandemic as being highly vulnerable to infection [1] and at risk of worse clinical outcomes [2] compared with the general population. Care home size, community prevalence, some types of dementia care provision and frailty have all been identified as affecting the risk of individual infection or care home outbreak [3–5].

Dementia was the most frequently identified pre-existing condition amongst care home residents who died of COVID-19 between March 2020 and April 2021 [6] and was also associated with increased risk of hospitalisation with a positive SARS-CoV-2 test in the general population prior to vaccine roll-out (OR 3.5, 95% CIs 1.93–6.34) [7]. Dementia itself may make individuals more vulnerable, but the impact on cognitive function and judgement can additionally make it more difficult to implement infection control measures, such as isolation [8]. However, it is not clear what the specific impact of dementia may be on individuals' risk, nor how dementia may interact with other individual and setting-level factors.

In this study, we used a national cohort dataset, including individual-level dementia diagnosis, SARS-CoV-2 testing and care home of residency to evaluate factors affecting the likelihood of a positive SARS-CoV-2 test. The chosen time period of September–December 2020 was during the 'second wave' of the pandemic and predated the UK vaccination programme [6]. This permitted a substantial number of cases to be included and reduced complexities associated with analysing differences between homes at different stages of vaccine roll-out. Given the clustering of residents in care homes that experience different levels of infection has an impact on each individual's risk we used a multi-level design, with individuals nested within care homes. Analysis included individual-level factors (e.g. demographics), compositional factors (e.g. care home size) and contextual factors (e.g. provision of nursing).

Methods

Design

We used anonymised individual-level population-scale linkable health and demographic records to create a cross-sectional retrospective cohort of care home residents in Wales. The outcome was a positive SARS-CoV-2 PCR test between 1 September and 31 December 2020. Individual-level variables were age, sex, dementia and frailty. Setting-level variables were compositional (registered places, residents identified, proportion not previously infected (susceptible), proportion with dementia) and contextual (dementia service available, nursing care provided, communal space, Health Board).

Participants

We created the cohort using data within the Secure Anonymised Information Linkage (SAIL) Databank. SAIL contains multiple anonymised individual-level, population-scale healthcare data sources, including SARS-CoV-2 testing results, linked using an anonymous linking field (ALF) created by a trusted third party (Digital Health and Care Wales (DHCW)) [9–12].

SAIL includes address data, which allowed care home residents to be identified through linkage to Care Inspectorate Wales (CIW) registration data [13]. Dementia diagnosis data were available via the SAIL Dementia eCohort (SDeC) [14]. SDeC links records from primary care, hospital admissions and mortality to identify individuals in Wales who have received a formal diagnosis of dementia [14, 15]. SAIL and SDeC include data from approximately 80% of General Practices (GPs) in Wales [15].

Frailty data are available from the electronic Frailty Index (eFI) using primary care records from the Welsh Longitudinal General Practice (WLGP) dataset, which are updated monthly within SAIL [16]. The eFI assigns individuals to one of four categories based on a model of cumulative deficits [17].

Care homes

A total of 1,073 adult care homes were registered with Care Inspectorate Wales (CIW) in May 2020. Their recorded capacity was 25,661 places [3]. Data available from CIW included the number of registered places, provision of nursing care and details of dementia service availability based on a 2019 review of registrations. It was also possible to use Unique Property Reference Numbers [18] and CIW audits to describe communal space [18].

Statistical analysis

We carried out univariate logistic regression to assess associations between a positive SARS-CoV-2 test and individual and setting-level risk factors. We then used univariate multi-level logistic regression modelling to assess associations between each individual and care-home level variable and a positive test for SARS-CoV-2, accounting for the non-independence of results for individuals in the same care home.

Finally, we built a multivariable multi-level logistic regression model. First, we defined a 'null' baseline model including only individual-level intercepts. This provided a value for the variance within the data in terms of outcomes between individuals, without considering how these individuals were clustered within homes or how specific individual or group-level characteristics might affect outcomes. We then defined a null model including care home and health board levels to assess whether clustering individuals within these higher order units reduced model variance, and so improved model fit. Finally, we added individual and group level characteristics to evaluate whether these further improved model fit, with all characteristics significantly associated with outcome in the univariate multi-level modelling regression added in order based on the size of the recorded odds ratio. With each addition of a variable, the variance associated with the new

model was compared to the variance for the previous model using ANOVA. Those models showing reduced variance with a *p* value less than 0.05 were considered as better fitting and retained.

We checked for multicollinearity between age, frailty and dementia by calculating the variance inflation factor (VIF) and considering a VIF of 5 or more as a threshold requiring amendment of the analysis plan [20]. The variance partition coefficient (VPC) was used to attribute variance between individuals and settings [21–23]. The VPC can be interpreted as the proportion of variance explained by the clustering of individuals into settings [21].

Additional analysis was carried out to consider whether factors associated with risk of infection were also associated with mortality. We modelled COVID-19-related mortality (i.e. deaths within 28 days of a positive test for SARS-CoV-2) and all-cause mortality.

Statistical analysis was carried out in R (4.1.0 [24]) using the lme4 package (v1.1-27 [25]). VIF was calculated using the car package (3.0-12 [26]). Equations for all models are presented in the Supplementary Data, available in *Age and Ageing* online.

Results

Cohort

A total of 12,211 individuals were identified as resident in 909 care homes on 1 September 2020. Participants with no record of GP registration were removed, as were those having a positive SARS-CoV-2 test result prior to 1 September 2020, as evidence suggests reinfection over this period was rare [6] and these residents would have different susceptibility to infection. Where mortality or hospital records indicated less than 14 days (i.e. two incubation periods) of residence individuals were excluded. Homes in which every resident was under 65 years old were also excluded, as these included homes where individuals had needs (e.g. historic brain injury) unrelated to age, dementia or frailty and therefore experienced different risks. Finally, homes for which the status of nursing care provision was unknown were excluded. Exclusions are detailed in Figure 1.

The final cohort included 9,571 individuals in 673 care homes. There were 5,647 individuals (59%) with a dementia diagnosis. The mean time since diagnosis was 61.1 months. A total of 2,761 residents were recorded with Alzheimer's disease (49.9% of all those with a dementia diagnosis) and 1,781 with a diagnosis of vascular dementia (31.5%). Some 1,488 (15.5%) individuals recorded a positive test for SARS-CoV-2 between 1 September and 31 December 2020. Of those 973 (65.4%) had a dementia diagnosis.

Statistical analysis

Descriptive statistics, results of univariate logistic regression (no levels included) and univariate multi-level logistic regression (individuals clustered by care home of residence and care homes clustered by Health Board) are shown in Table 1.

The VIFs for age, frailty and dementia flag were 1.05, 1.02 and 1.02, respectively. Multicollinearity was therefore not considered an issue. Residuals were plotted for continuous measures of age, residents identified and proportion of residents with dementia. Plots are included in the Supplementary Data (Supplementary Figures 1–3 are available in *Age and Ageing* online) and suggested linearity.

Initial intercept-only models indicated that clustering individuals by care homes significantly improved model fit ($p < 0.001$). However, including Health Boards as an additional level did not improve fit and this level was not included in subsequent analyses.

Variables describing individual or group-level characteristics were then added. Communal space and individual-level frailty were dichotomised to reflect thresholds suggested by multi-level regression modelling. The proportion of individuals with dementia significantly reduced variance. No significant reduction in variance was associated with age, resident numbers or communal space. The addition of severe frailty as a binary variable did reduce variance. This process is shown in Table 2.

Therefore, the final model included two factors associated with infection risk: the care home-level proportion of residents with a dementia diagnosis and individual-level severe frailty. This model suggests that the odds of an individual testing positive in a home in which 50–74% of residents were diagnosed with dementia were 3.36 times (95% CIs 1.19–9.48) that of an individual in a home in which no residents had dementia. For an individual in a home where more than 75% of individuals were recorded with a dementia diagnosis the odds ratio for testing positive was 4.54 (95% CIs 1.55–13.27). The odds ratio of a positive test for an individual who was severely frail was 1.29 (95% CIs 1.05–1.59) compared with an individual who was mildly, moderately or not frail.

The VPC for the final model was 0.76, suggesting that 76% of variation in the infection risk for a given care home resident is due to their residency in a particular care home rather than individual factors.

Mortality

A total of 1,075 residents (11.2%) died during the period, of whom 316 (3.3%) recorded a positive SARS-CoV-2 test within the preceding 28 days. Age, frailty, dementia diagnosis and proportion of residents with a dementia diagnosis were associated with mortality in univariate multilevel models for both COVID-19-related deaths and all deaths. In the final multivariate multilevel model for COVID-19-related mortality, only dementia and age remained associated (VPC = 0.64). The final multivariate multilevel model for all-cause mortality included severe frailty, dementia diagnosis and proportion of residents with dementia (VPC = 0.11). Descriptive statistics and analysis for mortality are summarised in the Supplementary Data, available in *Age and Ageing* online.

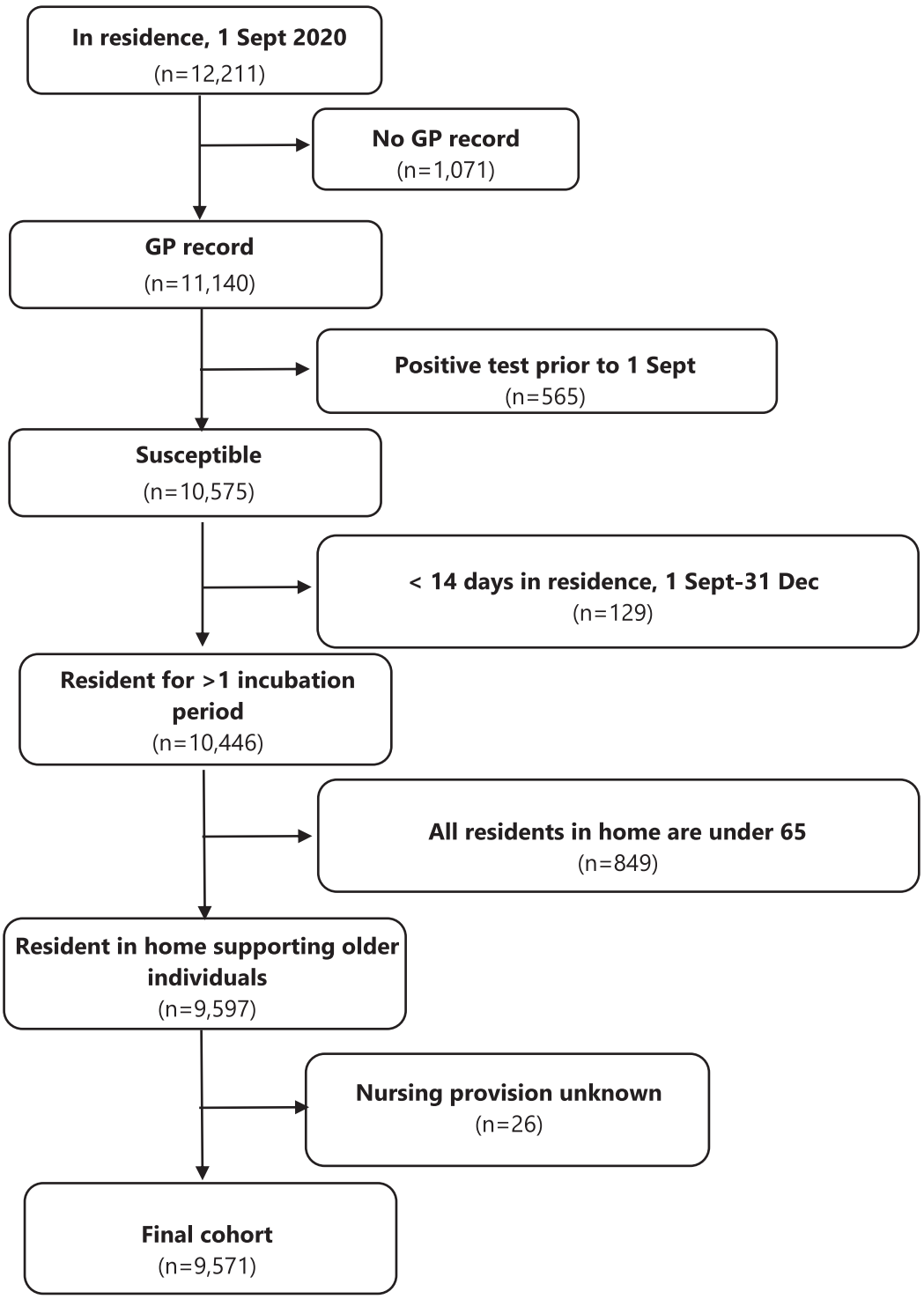

Figure 1. Study flow diagram of participants and exclusions.

Discussion

Summary

Our multi-level analysis of individual and care home-level factors and the likelihood of COVID-19 in residents has produced interesting findings. Individual frailty and the proportion of residents with dementia remained significant in the final model. The odds ratios of a positive SARS-CoV-2 test for individuals in homes with different proportions of residents with dementia suggests a gradient of risk; however, overlapping confidence intervals also suggest the possibility of a threshold at which risks may increase substantially. It is important to note that these risks were identified regardless of an individual's own dementia status.

Table 1. Descriptive analysis of risk factors, odds ratios and associated confidence intervals: those testing positive for SARS-CoV-2 in the period compared with those who did not. All odds ratios are unadjusted univariate estimates, from either the basic unstructured logistic regression, or from the multi-level logistic regression (structured at individual, care home and Health Board levels). Note that separate analyses have been included for age as a continuous variable (which minimises loss of information in statistical analysis) and as a categorical variable (for ease of interpretation)

	Descriptive data				Univariate logistic regression			Univariate multi-level logistic regression (individual, care homes, Health Board levels)			
	Residents with a positive test		Residents with no positive test		OR	CIs	*p*	OR	CIs	*p*	VPC
Total	1,488	15.5%	8,083	84.5%							
Individual level characteristics											
Female	1,019	15.7%	5,456	84.3%			BASE			BASE	0.71
Male	469	15.1%	2,628	84.9%	1.05	0.93–1.18	0.45	0.89	0.75–1.06	0.2	
Mean days in residence during study period	110		114		0.99	0.99–0.99	<0.001	1	1–1	0.45	0.7
Mean age	83.4		81.2		1.01	1.01–1.02	<0.001	1.01	1–1.02	0.03	0.71
Age category											
Under 60	62	7.6%	756	92.4%			BASE			BASE	0.69
60–64	32	10.8%	265	89.2%	1.47	0.93–2.29	0.09	1.48	0.79–2.78	0.23	
65–69	54	12.1%	393	87.9%	1.68	1.14–2.46	<0.01	1.61	0.92–2.81	0.09	
70–74	113	16.3%	582	83.7%	2.37	1.71–3.31	<0.001	2.14	1.29–3.55	<0.01	
75–79	166	17.4%	786	82.6%	2.58	1.90–3.53	<0.001	2.54	1.54–4.16	<0.001	
80–84	265	17.4%	1,257	82.6%	2.57	1.94–3.47	<0.001	2.48	1.54–4.01	<0.001	
85–89	351	18.0%	1,602	82.0%	2.68	2.03–3.58	<0.001	2.86	1.78–4.59	<0.001	
over 90	445	15.4%	2,442	84.6%	2.22	1.70–2.96	<0.001	2.21	1.39–3.53	<0.01	
Dementia diagnosis											
No	515	13.1%	3,409	86.9%			BASE			BASE	0.7
Yes	973	17.3%	4,674	82.7%	1.39	1.24–1.56	<0.001	1.03	0.86–1.23	0.76	
Frailty category											
Fit	184	11.8%	1,377	88.2%			BASE			BASE	0.7
Mild	458	15.6%	2,485	84.4%	1.38	1.15–1.66	<0.01	1.02	0.78–1.34	0.65	
Moderate	560	17.1%	2,710	82.9%	1.55	1.30–1.85	<0.001	1.28	0.97–1.69	0.22	
Severe	286	15.9%	1,511	84.1%	1.42	1.16–1.73	<0.01	1.53	1.12–2.09	0.03	
Care home level factors: compositional											
Registered maximum places					1	1–1	0.03	1.02	1.01–1.04	<0.01	0.72
<10 places	66	12.2%	476	87.8%			BASE			BASE	0.72
10–24 places	222	12.4%	1,564	87.6%	1.05	0.79–1.41	0.74	1.58	0.56–4.43	0.39	
25–49 places	835	18.2%	3,771	81.1%	1.65	1.27–2.16	<0.001	4.95	2.03–12.11	<0.001	
50+ places	365	13.8%	2,272	86.2%	1.19	0.91–1.58	0.22	2.42	0.82–7.11	0.11	
Residents identified on SAIL					1	1–1.01	0.03	1.03	1.01–1.05	<0.01	0.72
<10	100	9.9%	913	90.1%			BASE			BASE	0.72
10–24	735	16.3%	3,766	83.7%	1.78	1.43–2.23	<0.001	4.25	2.11–8.55	<0.001	
25–49	522	18.1%	2,726	81.9%	1.75	1.40–2.20	<0.001	3.24	1.39–7.54	<0.01	
50+	131	16.2%	678	83.8%	1.76	1.34–2.33	<0.001	8.07	1.35–48.25	0.02	
Any resident susceptible											
Some suscept.	473	13.0%	3,176	87.0%			BASE			BASE	0.7
All susceptible	1,015	17.1%	4,907	82.9%	1.39	1.23–1.56	<0.001	1.51	0.82–2.79	0.19	
% of residents with dementia diagnosis											
0%	43	9.7%	402	90.3%			BASE			BASE	0.71
1–24%	46	5.6%	782	94.4%	0.55	0.36–0.84	<0.01	0.86	0.25–2.94	0.81	
25–49%	322	15.0%	1,829	85.0%	1.57	1.14–2.22	<0.01	3.62	1.30–10.09	0.01	
50–74%	558	16.0%	2,938	84.0%	1.83	1.34–2.56	<0.001	4.28	1.60–11.42	<0.01	
75%+	519	19.6%	2,132	80.4%	2.35	1.72–3.30	<0.001	6.10	2.21–16.87	<0.01	
Care home level factors: contextual											
Dementia service											
None	220	11.3%	1,727	88.7%			BASE			BASE	0.72
Non-specialist	890	17.2%	4,288	82.8%	1.65	1.41–1.93	<0.001	3.15	1.53–6.47	<0.01	
Specialist	346	16.7%	1,724	83.3%	1.59	1.32–1.90	<0.001	2.46	0.97–6.25	0.06	
Unknown	32	8.5%	344	91.5%	0.74	0.49–1.07	0.12	0.76	0.16–3.42	0.58	

(*Continued*)

Table 1. Continued

	Descriptive data				Univariate logistic regression			Univariate multi-level logistic regression (individual, care homes, Health Board levels)			
	Residents with a positive test		Residents with no positive test		OR	CIs	*p*	OR	CIs	*p*	VPC
Nursing											
No	797	16.2%	4,132	83.8%			BASE			BASE	0.71
Yes	691	14.9%	3,951	85.1%	0.91	0.82–1.01	0.11	1.04	0.56–1.93	0.90	
Communal space, data available											
<50 m^2	35	14.5%	207	85.5%			BASE			BASE	0.78
50–99 m^2	178	14.9%	1,014	85.1%	1.02	0.70–1.53	0.91	2.00	0.45–8.90	0.36	
100–199 m^2	434	15.3%	2,406	84.7%	1.07	0.75–1.57	0.73	4.11	0.99–17.12	0.52	
200–499 m^2	401	13.8%	2,495	86.2%	0.96	0.67–1.41	0.81	3.17	0.76–13.15	0.11	
500 m^2+	195	18.2%	877	81.8%	1.32	0.90–1.97	0.17	8.01	1.40–45.70	0.02	
Not recorded	245	18.4%	1,084	81.6%	1.34	0.92–1.99	0.14	4.95	1.12–21.93	0.04	
Health Board											
Aneurin Bevan	313	22.3%	1,090	77.7%			BASE	Health Board added as level			
BCU	200	7.3%	2,533	92.7%	0.27	0.23–0.33	<0.001				
C&V	170	13.4%	1,094	86.6%	0.54	0.44–0.66	<0.001				
CTM	300	29.2%	726	70.8%	1.44	1.20–1.73	<0.001				
Hywel Dda	174	12.2%	1,256	87.8%	0.48	0.39–0.59	<0.001				
Powys	37	11.7%	278	88.3%	0.46	0.32–0.66	<0.001				
Swansea Bay	294	21.0%	1,106	79.0%	0.92	0.77–1.11	0.40				

The high value for the VPC demonstrates that the care home environment makes a substantially greater contribution to a resident's likelihood of contracting COVID-19 than individual risk factors.

Area of residence and prevalence

Initial analysis without including levels suggested Health Board was a meaningful predictor of the likelihood of a positive test. Additional evidence on COVID-19 prevalence in local authorities (see Supplementary Table 2 available in *Age and Ageing* online) did not suggest any clear patterns by rurality or population density (e.g. during week 44 of 2020, the midpoint of the study period, prevalence in Swansea was double that of Neath Port Talbot, an adjacent urban area). Comparison between care home infection rates within and between lower super output areas (LSOAs) [4] suggests community prevalence was a more relevant factor than rates across higher level geographies.

Individual-level factors

Although dementia diagnosis appeared to be associated with the chances of an individual testing positive for SARS-CoV-2 in the initial univariate logistic regression, this effect disappeared once the clustering of individuals within care homes was modelled. Associations between age and risk of infection were similar in basic and multi-level univariate regression.

It was surprising that dementia did not emerge as a risk factor at an individual level, particularly given previous clinical evidence that neurodegenerative diseases may break down the blood–brain barrier, providing one plausible explanation for associations between COVID-19 and dementia [27] and for increasing vulnerability to neurological complications and mortality amongst those with dementia diagnoses [28, 29]. No analysis was carried out on subgroups of those with different forms of dementia, which may be associated with different types and levels of infection risk, and other conditions associated with cognitive impairment were not included. We cannot rule out the possibility that high prevalence of cognitive decline in the study population masked the effect of dementia, and further research might include a wider range of diagnoses.

The relationship between frailty and a positive test for SARS-CoV-2, apparent in the initial analysis, was substantially attenuated within the univariate multi-level model, with only severe frailty remaining as a risk factor (OR 1.29, 95% CIs, 1.05–1.59).

Compositional factors

We tested associations for both maximum registered capacity and the number of residents identified in SAIL. Modelled in isolation, both variables suggested an association between resident numbers and infection risk. However, the effect was no longer evident when included with the proportion of residents with dementia in the multivariable model. The proportion of residents with a dementia diagnosis remained associated with the likelihood of infection in the multivariable model, with the risk increasing at any level of dementia prevalence above 50% compared with living in a home with no diagnosed residents.

Table 2. Multivariable multi-level logistic regression models for positive test for SARS-CoV-2 in the study period. Odds ratios, confidence intervals and p values. p value of ANOVA refers to the p value returned when each model is compared to the previous model, with $p < 0.05$ used as a threshold for considering the model as an improved fit

Variables	VPC	AIC p value of ANOVA	Variable	OR	CIs	p	Notes
Intercept-only (null models) (ANOVA to compare variance with preceding model)							
Intercept only (no levels)		8,273.2					Variance accounted for in intercept-only (null) model with no levels and no variables
Intercept only (care homes)	0.71	5,277.1 $p < 0.001$					Variance accounted for in intercept-only (null) model with care home as the only level and no variables
Intercept only (care homes + HBs)	0.71	5,276.9 $p = 1$					Variance accounted for in intercept-only (null) model with care home and HBs as levels and no variables No improvement on model with care homes as only level, so model dropped.
Models including care homes as levels and variables (ANOVA to compare variance with preceding model)							
% residents with dementia diagnosis	0.71	5,259.8 $p < 0.001$	0%			**BASE**	'% of residents with dementia diagnosis' added to model. Model variance significantly reduced compared to previous model, therefore variable retained
			0–24%	0.58	0.17–2.03	0.4	
			25–49%	2.86	0.98–8.32	0.05	
			50–74%	3.53	1.26–9.88	0.02	
			75%+	4.77	1.64–13.87	<0.01	
% residents with dementia diagnosis Age	0.76	5,260.3 $p = 0.22$					'Age' added to model (as continuous variable) No significant reduction in variance compared to previous model, therefore variable dropped
% residents with dementia diagnosis Residents identified	0.77	5,259.0 $p = 0.13$					'% of residents identified' added to model (as continuous variable) No significant reduction in model variance compared to previous model, therefore variable dropped
% residents with dementia diagnosis Residents in home with non-specialist dementia	0.76	5,258.5 $p = 0.07$					Binary category ('non-specialist dementia service' v 'other category') added to model No significant reduction in model variance, therefore variable dropped
% residents with dementia diagnosis Residents in homes with > 500 m^3 of communal space	0.72	5,259.9 $p = 0.16$					Binary category ('500 m3+' v 'other') added to model No significant reduction model variance, therefore variable dropped
% residents with dementia diagnosis Frailty - severe	0.76	5,256.1 $p = 0.02$	0%			BASE	Binary category of severe frailty no/mild/moderate frailty added to model Significant reduction in variance, therefore variable retained Final model
			0–24%	0.57	0.16–2.01	0.39	
			25–49%	2.75	0.94–8.03	0.07	
			50–74%	3.36	1.19–9.48	0.02	
			75%+	4.54	1.55–13.27	<0.01	
			Severe frailty	1.29	1.05–1.59	0.02	

Contextual factors

There was only very limited evidence for the impact of different types of care within homes. However, it is important to note that categories such as 'specialist dementia care' may mitigate risks at an individual-level but be less effective at addressing risks associated with supporting a large resident population with high needs. The size of communal areas is likely to correlate with registered places, further suggesting care home size was not a risk factor. Public Health Wales guidance, including requiring the closing of communal areas during an outbreak [30], is also likely to have mitigated risks.

Mortality analysis

Analysis of mortality suggested that risk factors for all deaths were similar to those for COVID-19 infection (with the addition of individual-level dementia diagnosis) but that the proportion of variation accounted for by residency in a given care home was considerably lower. In contrast, only individual-level factors remained associated for

COVID-19-related deaths, but the clustering of individuals in care homes accounted a higher proportion of variation. This divergence in findings may result from under-ascertainment of COVID-19-related deaths, (due to positive prior tests or subsequent deaths falling outside the period) or due to COVID-19 infection carrying a longer-term risk for older people, meaning all-cause mortality risk factors more closely mirroring those for infection in these settings. In addition, there may be more complexity in causal pathways associated with mortality (e.g. dementia diagnosis may affect mortality risk both directly and through increasing risks of infection) and therefore odds ratios may not interpretable as straightforward adjusted effects [31]. Whilst these results are difficult to interpret, the evidence on all-cause mortality suggests the possibility that the prevalence of dementia in care homes would benefit from further analysis.

Implications

The fact that both individual factors (frailty) and compositional factors (proportion of residents with dementia) remained in the final model suggests both are relevant to risks of infection in care home settings. It is striking that dementia was not found to be a significant individual-level but was found to be a major factor at the care home level. This may reflect a distinction between the risks of an individual contracting infection and the risks of infection spread within the environment. For example, frailty may render an individual more susceptible to infection, and/or require greater 'hands on' contact from carers with consequent increased infection risk. However, settings with substantial numbers of individuals with dementia may have increased risk of spread due both to the increased numbers of staff required to support residents and to the challenges of maintaining social distancing and isolation where more physical contact between staff and residents is required. Further research to explore the relationship between individual and environmental vulnerability in relation to infection might consider the impact of these factors on infection rates amongst residents and also amongst staff who are not themselves frail or diagnosed with dementia but work in settings with different proportions of residents with dementia. Further research might also analyse patterns of infection amongst residents to establish whether initial cases were more or less likely to have a diagnosis of dementia and whether initial cases with dementia were associated with more rapid or widespread infection within the setting.

The COVID-19 pandemic has presented enormous challenges to the care sector and those working within it. Reviews have identified a need to prepare whole system strategies based on high-quality evidence that considers environmental and social factors within care home settings and not only individual level vulnerabilities. This evidence implies clinicians, management and front line social care staff need to collaborate effectively together to plan preventative approaches to improve the physical and social environment (for example, to reduce instances where those with dementia move between different parts of a care home) and to manage limited resources, such as staff availability and protective equipment during outbreaks.

It is important to note emerging evidence that despite the effectiveness of vaccination programmes [32–35], some residents, particularly those who are frail, continue to face higher risks of infection [32] and that vaccine effectiveness wanes within months for both residents and staff [31, 35]. This evidence suggests findings presented in this paper are likely to remain relevant for COVID-19 and may also be applicable to efforts to reduce transmission of other respiratory diseases such as influenza.

Strengths and weaknesses

This study benefits from a large, national population cohort of care home residents. The opportunity to use existing, robust measures of dementia and frailty was also a strength. The study demonstrates the value of using multi-level modelling designs within research on care homes.

Weaknesses include imperfect ascertainment of resident individuals, with no opportunities to evaluate potential biases introduced, and the requirement to remove residents with no GP records. 'Dementia diagnosis' may include individuals with a broad range of cognitive impairments. A proportion of residents will also have cognitive impairments unrelated to dementia. The assumption that all individuals with dementia (and only those with dementia) face equal challenges in following advice protective measures such as isolation is a limitation. Using data from the winter 2020–2021 COVID-19 'second wave' time period, before vaccination was available may limit generalisability of the analysis.

Conclusions

These analyses suggest there was a substantial increased risk of COVID-19 in this period amongst those residents in care homes with a large proportion of residents with a dementia diagnosis, and who were severely frail. An individual's chances of a positive test were not related to whether they themselves had a diagnosis of dementia, their age or the size of their care home. Care homes and the public health system should consider how these findings can be incorporated into guidance for preventing and mitigating the spread of infection within these settings.

Data Availability: The data used in this study are available in the SAIL Databank at Swansea University, Swansea, UK, but as restrictions apply they are not publicly available. All proposals to use SAIL data are subject to review by an independent Information Governance Review Panel (IGRP). Before any data can be accessed, approval must be given by the IGRP. The IGRP gives careful consideration to each project to ensure proper and appropriate use of SAIL data.

When access has been granted, it is gained through a privacy protecting safe haven and remote access system referred to as the SAIL Gateway. SAIL has established an application process to be followed by anyone who would like to access data via SAIL at https://www.saildatabank.com/application-process.

The R code used to create the study dataset and carry out statistical analysis is publicly available at https://github.com/ChrisEmmerson/Dementia_frailty_carehome_SARS-CoV-2_study.

Acknowledgements This study makes use of anonymised data held in the Secure Anonymised Information Linkage (SAIL) Databank. This work uses data provided by patients and collected by the NHS as part of their care and support. We would also like to acknowledge all data providers who make anonymised data available for research. We wish to acknowledge the collaborative partnership that enabled acquisition and access to the de-identified data, which led to this output. The collaboration was led by the Swansea University Health Data Research UK team under the direction of the Welsh Government Technical Advisory Cell (TAC) and includes the following groups and organisations: the SAIL Databank, Administrative Data Research (ADR) Wales, Digital Health and Care Wales (DHCW), Public Health Wales, NHS Shared Services Partnership (NWSSP) and the Welsh Ambulance Service Trust (WAST). All research conducted has been completed under the permission and approval of the SAIL independent Information Governance Review Panel (IGRP) project number 0911.

Declaration of Sources of Funding: This work was supported by the Con-COV team funded by the Medical Research Council (grant number: MR/V028367/1). This work was supported by Health Data Research (HDR) UK, which receives its funding from HDR UK Ltd (HDR-9006) funded by the UK Medical Research Council, Engineering and Physical Sciences Research Council, Economic and Social Research Council, Department of Health and Social Care (England), Chief Scientist Office of the Scottish Government Health and Social Care Directorates, Health and Social Care Research and Development Division (Welsh Government), Public Health Agency (Northern Ireland), British Heart Foundation (BHF) and the Wellcome Trust. This work was supported by the ADR (Administrative Data Research) Wales programme of work. The ADR Wales programme of work is aligned to the priority themes as identified in the Welsh Government's national strategy: Prosperity for All. ADR Wales brings together data science experts at Swansea University Medical School, staff from the Wales Institute of Social and Economic Research, Data and Methods (WISERD) at Cardiff University and specialist teams within the Welsh Government to develop new evidence, which supports Prosperity for All by using the SAIL Databank at Swansea University, to link and analyse anonymised data. ADR Wales is part of the Economic and Social Research Council (part of UK Research and Innovation) funded Administrative Data Research UK (grant ES/S007393/1). This work was supported by the Wales COVID-19 Evidence Centre, funded by Health and Care Research Wales.

Declaration of Conflicts of Interest: R.A.L. is an unremunerated member of the Welsh Government COVID-19 Technical Advisory Group.

References

1. Scientific Advisory Group for Emergencies. Scientific Pandemic Influenza Group on Modelling (SPI-M) Consensus Statement, Government Office for Science, 2020. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/888763/S0201_SAGE27_200420_SPI-M-O_consensus_statement.pdf.
2. Zhou F, Wang Y, Liu Y *et al.* Disease severity and clinical outcomes of community-acquired pneumonia caused by non-influenza respiratory viruses in adults: a multicentre prospective registry study from the CAP-China Network. Eur Respir J 2019; 54: 1802406. https://doi.org/10.1183/13993003.02406-2018.
3. Emmerson C, Adamson JP, Turner D *et al.* Risk factors for outbreaks of COVID-19 in care homes following hospital discharge: A national cohort analysis. Influenza Other Respi Viruses 2021; 15: 371–80.
4. Hollinghurst J, Hollinghurst R, North L *et al.* COVID-19 risk factors amongst 14,786 care home residents: an observational longitudinal analysis including daily community positive test rates of COVID-19, hospital stays and vaccination status in Wales (UK) between 1 September 2020 and 1 May 2021. Age Ageing 2022; 51: 1–9.
5. Burton JK, McMinn M, Vaughan JE, Fleuriot J, Guthrie B. Care-home outbreaks of COVID-19 in Scotland March to May 2020: national linked data cohort analysis. Age Ageing 2021; 50: 1482–92.
6. Office for National Statistics. Deaths Involving COVID-19 in the Care Sector, England and Wales: Deaths Registered between Week Ending 20 March 2020 and Week Ending 2 April 2021. Office for National Statistics. 2021. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/articles/deathsinvolvingcovid19inthecaresectorenglandandwales/latest (24 November 2021, date last accessed).
7. Atkins JL, Masoli JAH, Delgado J *et al.* Preexisting comorbidities predicting COVID-19 and mortality in the UK Biobank community cohort. J Gerontol Ser A Biol Sci Med Sci 2020; 75: 2224–30.
8. Iaboni A, Cockburn A, Marcil M *et al.* Achieving safe, effective, and compassionate quarantine or isolation of older adults with dementia in nursing homes. Am J Geriatr Psychiatry 2020; 28: 835–8.
9. Lyons RA, Jones KH, John G *et al.* The SAIL databank: linking multiple health and social care datasets. BMC Med Inform Decis Mak 2009; 9: 1–8. https://doi.org/10.1186/1472-6947-9-3.
10. Ford DV, Jones KH, Verplancke J-P *et al.* The SAIL databank: building a national architecture for e-health research and evaluation. BMC Health Serv Res 2009; 9: 157. https://doi.org/10.1186/1472-6963-9-157.

11. Jones KH, Ford DV, Jones C *et al.* A case study of the Secure Anonymous Information Linkage (SAIL) gateway: a privacy-protecting remote access system for health-related research and evaluation. J Biomed Inform 2014; 50: 196. https://doi.org/10.1016/J.JBI.2014.01.003.
12. Lyons J, Akbari A, Torabi F *et al.* Understanding and responding to COVID-19 in Wales: protocol for a privacy-protecting data platform for enhanced epidemiology and evaluation of interventions. BMJ Open 2020; 10: 1–7.
13. Rodgers SE, Lyons RA, Dsilva R *et al.* Residential Anonymous Linking Fields (RALFs): a novel information infrastructure to study the interaction between the environment and individuals' health. J Public Health (Bangkok) 2009; 31: 582–8.
14. Improving social care and childcare in Wales, Chief Inspector's Annual Report 2020-21. Care Inspectorate Wales. https://careinspectorate.wales/about-us/what-we-do.
15. Schnier C, Wilkinson T, Akbari A *et al.* The Secure Anonymised Information Linkage databank dementia e-cohort (SAIL-DeC). Int J Popul Data Sci 2020; 5. https://doi.org/10.23889/ijpds.v5i1.1121.
16. Hollinghurst J, Fry R, Akbari A *et al.* External validation of the electronic Frailty Index using the population of Wales within the Secure Anonymised Information Linkage Databank. Age Ageing 2019; 48: 922–6.
17. 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: 353–60.
18. Power of UPRNs. Unique Property Reference Number. GeoPlace LLP. https://www.geoplace.co.uk/addresses-streets/location-data/the-uprn (27 September 2021, date last accessed).
19. Geoplace. Addressing the UK. Local Government Association. Improvement service. Ordinance Survey. 2021. www.geoplace.co.uk
20. Akinwande MO, Dikko HG, Samson A. Variance inflation factor: as a condition for the inclusion of suppressor variable(s) in regression analysis. Open J Stat 2015; 5: 754–67.
21. Leyland AH, Groenewegen PP. Multilevel Modelling for Public Health and Health Services Research. Springer. 2020. https://doi.org/10.1007/978-3-030-34801-4.
22. Merlo J, Viciana-Fernández FJ, Ramiro-Fariñas D. Bringing the individual back to small-area variation studies: a multilevel analysis of all-cause mortality in Andalusia. Spain Soc Sci Med 2012; 75: 1477–87.
23. Goldstein H, Browne W, Rasbash J. Partitioning variation in multilevel models. Underst Stat 2002; 1: 223–31.
24. R Core Team. R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing; 2018. https://www.r-project.org.
25. Bates D, Mächler M, Bolker BM, Walker SC. Fitting linear mixed-effects models using lme4. J Stat Softw 2015; 67: 1–48.
26. Fox J, Weisberg S. An R Companion to Applied Regression. 3rd edition. Thousand Oaks: Sage, 2019.
27. Kuo CL, Pilling LC, Atkins JL *et al.* APOE e4 genotype predicts severe COVID-19 in the UK Biobank community cohort. J Gerontol Ser A Biol Sci Med Sci 2020; 75: 2231. https://doi.org/10.1093/GERONA/GLAA131.
28. Sweeney MD, Sagare AP, Zlokovic BV. Blood–brain barrier breakdown in Alzheimer's disease and other neurodegenerative disorders. Nat Rev Neurol 2018; 14: 133. https://doi.org/10.1038/NRNEUROL.2017.188.
29. Ueno M, Chiba Y, Matsumoto K *et al.* Blood–brain barrier damage in vascular dementia. Neuropathology 2016; 36: 115–24.
30. Guidance to Prevent COVID-19 Among Care Home Residents and Manage Cases, Incidents and Outbreaks in Residential Care Settings in Wales. version 4.2. Cardiff: Public Health Wales, 2020.
31. Westreich D, Greenland S. The Table 2 fallacy: presenting and interpreting confounder and modifier coefficients. Am J Epidemiol 2013; 177: 292–8.
32. Hollinghurst J, North L, Perry M *et al.* COVID-19 infection risk amongst 14,104 vaccinated care home residents: a national observational longitudinal cohort study in Wales, UK, December 2020–March 2021. Age Ageing 2022; 51: 1–7.
33. Charles, A and Ewbank, L. The road to renewal: five priorities for health and care. The King's Fund. 2021. https://www.kingsfund.org.uk/publications/covid-19-road-renewalhealth-and-care
34. Bedston S, Akbari A, Jarvis CI *et al.* COVID-19 vaccine uptake, effectiveness, and waning in 82,959 health care workers: a national prospective cohort study in Wales. Vaccine 2022; 40: 1180. https://doi.org/10.1016/J.VACCINE.2021.11.061.
35. Shrotri M, Krutikov M, Nacer-Laidi H *et al.* Duration of vaccine effectiveness against SARS-CoV-2 infection, hospitalisation, and death in residents and staff of long-term care facilities in England (VIVALDI): a prospective cohort study. The Lancet Healthy Longevity 2022; 3: e470–80. https://doi.org/10.1016/S2666-7568(22)00147-7.

Received 18 February 2022; editorial decision 22 August 2022