A retrospective cohort study of risk factors for mortality among nursing homes with Covid-19

Clara Suñer PhD[1]† & Dan Ouchi MsC[1,2]†, Miquel Àngel Mas PhD[3,4], Rosa Lopez Alarcon[5], Mireia Massot Mesquida[6], Núria Prat MD[6], Josep Maria Bonet-Simó MD[6], Marta Expósito Izquierdo MD[6], Irene Garcia Sánchez MD[6], Sara Rodoreda Noguerola MD[6], Montserrat Teixidó Colet PhD[6], Joaquim Verdaguer Puigvendrelló MD[6], Norma Henríquez PhD[6], Ramon Miralles PhD[2,3,4], Eugènia Negredo PhD[1,7,8], Marc Noguera-Julian PhD,[9,10] Michael Marks PhD[11,12], Jordi Ara PhD[6] & Oriol Mitjà PhD[1,7,10,13]*

† Contributed equally.

1. Fight AIDS and Infectious Diseases Foundation, Badalona, Badalona, Spain

2. Universitat Autònoma de Barcelona, Barcelona, Spain

3. Direcció Clínica Territorial de Cronicitat Metropolitana Nord, Institut Català de la Salut, Barcelona, Catalonia, Spain

4. Geriatrics Department, Hospital Universitari Germans Trias i Pujol, Badalona, Spain

5. Direcció d'Organització i Sistemes d'Informació. Gerència Territorial Metropolitana Nord. Institut Català de la Salut, Barcelona, Catalonia, Spain

6. Gerència i Direcció d'Atenció Primaria Metropolitana Nord, Institut Català de la Salut, Barcelona, Catalonia, Spain

7. Infectious Diseases Department, Hospital Universitari Germans Trias i Pujol, Badalona, Spain

8. Centre for Health and Social Care Research, Faculty of Medicine, University of Vic - Central University of Catalonia, Barcelona, Spain

9. Institut de Recerca de la Sida, IrsiCaixa, Badalona, Spain

10. Universitat de Vic – Universitat Central de Catalunya, Vic, Spain

11. Clinical Research Department, London School of Hygiene & Tropical Medicine, London, United Kingdom

12. Hospital for Tropical Diseases, London, United Kingdom

13. Lihir Medical Centre- InternationalSOS, Lihir Island, Papua New Guinea

* Corresponding author: Oriol Mitjà, omitja@flsida.org, Hospital Universitari Germans Trias i Pujol, Carretera Canyet s/n, 08916, Badalona, Spain

Introductory paragraph

Long-term care (LTC) facilities have shown remarkable high mortality during the coronavirus disease 2019 (Covid-19) outbreak in many countries. Several risk factors for Covid-19 mortality in LTC facilities have been previously identified. We used multiple variables covering facility characteristics (e.g., size, type of care provided, and preparedness) and socioeconomic factors related to the geographic location (e.g., household income, and incidence of Covid-19 in the surrounding population) to more precisely identify risk factors for Covid-19-related excess mortality (i.e., laboratory-proven, and clinically suspected Covid-19) as well as the overall excess mortality. In addition, we used clustering approaches to detect patterns in data sets and generate hypotheses regarding potential relationships between types of nursing homes and mortality trends. We retrospectively analysed mortality in 167 nursing homes that provide long-term care to 8,716 residents during the Covid-19 outbreak in Catalonia (North-East Spain) using nursing homes as the unit of analyses. Between March 1 and June 1, 2020, 1,629 deaths were reported in the investigated nursing homes, 1,089 (66.9%) of them Covid-19-related. According to the multiple regression analysis, both outcomes, Covid-19 related mortality and overall mortality, at the facility level were significantly associated with a higher percentage of patients with complex diseases, lower scores on pandemic preparedness measures and, higher population incidence of Covid-19 in the surrounding population. Consistently with these findings, when grouping nursing homes into eight clusters of common specific characteristics, we found higher mortality rates in the cluster with higher levels of residents with complex chronic conditions or advanced diseases, the cluster with higher levels of low scores on pandemic preparedness, the cluster of nursing homes located in rural areas, and also in the cluster of nursing homes with larger size. Our results provide policymakers and clinicians caring for residential populations

with valuable information for prioritizing interventions in this setting in the advent of a Covid-19 outbreak.

Main text

Twelve months after the first outbreak of the novel coronavirus disease (Covid-19), the global death toll associated with the pandemic amounted to more than two million people[1]. Residents of long-term care (LTC) facilities, such as nursing homes, are at high risk of Covid-19 infection and also at high risk of death[2]. The findings of several reports suggest that once Covid-19 has been introduced into a LTC facility it can spread rapidly and widely causing considerable impact among facility residents and local health care systems[3–6].

Large variations in Covid-19 death rates across LTC facilities have been observed. Whether the high death rates are exclusively linked to vulnerability of residents, owing to advanced ages and underlying health conditions, and whether the structural features or geographic location of LTC facilities are a contributing factor are still not fully answered questions. Because of the different policy implications of the relative influence of these two groups of characteristics, there is a need to more comprehensively understand the determinants of SARS-CoV-2 spread and mortality in LTC facilities[7,8].

Various predictors of higher Covid-19 infection rate in nursing homes have been reported, including larger facility size, higher proportion of residents with limited income, greater percentage of African Americans (in the US), urban location, and community-level rates of Covid-19[4,5,9,10]. Similarly, several possible reasons for the large percentage of deaths from Covid-19 among residents in LCTs have been identified in previous studies. First, the individual residents' characteristics like older age or high levels of multi-comorbidity[11,12]. Second, the nursing home characteristics, primarily size (i.e., larger number of residents and crowding)[13,14], and type of care provided (featured as the proportion of residents with complex or advanced diseases)[6,8,12,15,16]. This

variable can be used as a proxy to differentiate two types of senior housing centres: nursing homes that provide around the clock care and monitoring to residents with complex health care conditions, in contrast to nursing homes that provide assisted living and custodial care to residents who are mostly independent. There are also mixed facilities with both type of residents. Third, the level of pandemic preparedness by the nursing home. Residential settings need to follow specific guidance for managing the outbreak to prevent infections in communal living areas and shared rooms[6,14], to stop transmission from multiple caregivers[3] —who may work across multiple different facilities—, and to avoid shortage of healthcare resources (e.g., tests, and personal protective equipment)[6]. Fourth, the nursing home geographic location in higher risk areas. Some authors have identified risk factors associated with the characteristics of the geographic location and surrounding population, such as the mean household income, the population density of the area, the ethnicity profile, or the high population incidence of Covid-19[10,13,17,18].

In this retrospective analysis, we investigated the determinants of mortality in LCT facilities during the Covid-19 outbreak. A major component of the novelty of this study comes from the use of multiple variables covering both facility characteristics (e.g., size, type of care provided, and preparedness) and characteristics of the geographic location (e.g., household income, population density, and local Covid-19 incidence). Furthermore, we used a two-fold analysis to provide a more comprehensive view of determinants of mortality: first, multivariable analysis to more precisely identify the contribution of facility-specific factors and those of the communities they serve to Covid-19-related excess mortality and to overall excess mortality. Second, we used clustering approaches to detect patterns in our data set regarding potential relationships of nursing homes grouped according to common characteristics and describing mortality trends across clusters.

Results

Characteristics of the nursing homes

Our analysis of all public and private nursing homes for LTC in the study area included 167 nursing homes providing long-term care to 8,716 residents. Table 1 summarizes the characteristics of the nursing homes included in the analysis. Most nursing homes had high percentage of individuals older than 80 years (mean age at the facility level 87.1 years). An average of 46.1% of individuals were complex chronic patients (CCP), 10.5% were patients with advanced chronic disease (ACD), and 82.1% were highly dependent on daily activities. The median SNQ12 score—an *ad hoc* index of unmet needs among 12 essential items for pandemic preparedness and response (values 1-highest preparedness to 12-lowest preparedness) —was 1.4. The individual demographic, clinical, and epidemiological characteristics of included residents are summarized in the Supplementary Table 1.

Mortality rate

Between March 1 and June 1, 2020, a total of 1,629 deaths were reported among people institutionalized in the nursing homes included in the analysis (Supplementary Figure 1). Of these, 1,089 (66.9%) were registered as Covid-19-related deaths in the mortality registry of the Department of Health. The cause of the death of the remaining 540 deaths could not be confirmed. Overall, the excess deaths in the analysed nursing homes compared with the same period in the four previous years were estimated to be 971 deaths; Covid-19-related deaths accounted for 89.2% of all excess mortality. At the nursing home level, the median (IQR) mortality rate was 14.3 (7.6

– 26.1) deaths/100 residents/3-month study period for all-cause death, and 3.9 (0.0 – 18.4) for Covid-19 related deaths.

Multivariable analysis of risk factors for mortality

Owing to the presence of deaths recorded as undetermined cause, we investigated the characteristics of nursing homes that may contribute to explaining mortality using two outcomes: Covid-19-related death (i.e., laboratory confirmed and clinically suspected) and all-cause death. According to the multivariable analysis (Table 2), the risk of Covid-19 related deaths was higher in nursing homes with a higher percentage of complex CCP patients (incidence rate ratio [IRR] 1.07; 95%CI 1.04 – 1.11 per 10 units increase), lower scores on the pandemic preparedness measures (IRR 1.05; 1.02 – 1.09 per unit increase), and geographic location in areas with higher Covid-19 14-day incidence rate per 100,000 population (IRR 2.39; 1.95 – 2.93 per 100 unit increase). The risk factors for all-cause mortality were the same as those for Covid-19 related death, and two additional factors: a higher percentage of male residents (IRR 1.07; 1.01 - 1.13 per 10 units increase) and nursing homes with larger number of residents (IRR 1.02; 1.01 - 1.03 per 10 units increase). At the individual level age, gender, comorbidities and the underlying incidence of SARS-COV-2 in the surrounding population were the dominant risk factors for Covid-19 related death (Supplementary Table 2).

Comparison of characteristics among nursing home clusters

Complementary to the multivariable analysis, we conducted a cluster analysis of nursing homes grouped according to common characteristics and assessed mortality trends across clusters. The cluster analysis based on the k-nearest neighbour classifier identified eight groups of nursing

homes that were significantly different from each other. Figure 1 illustrates the intensity of each characteristic (i.e., the extent of the difference between the mean of a given cluster and that of the entire sample) in the resulting clusters and the contribution of each characteristic to their definition. Key characteristics of each cluster were as follows:

Nursing homes in **cluster 1** were placed in low densely populated areas with high population incidence of Covid-19, and high household income; **cluster 2** were facilities with a high proportion of CCP and ACD patients, and located in areas with low population incidence of Covid-19; **cluster 3** had low proportion of CCPs and highly dependent residents, and high proportion of male residents; **cluster 4** had higher number of residents than the median, although with a very low proportion of CCPs; nursing homes in this cluster were placed in areas with low household income; **cluster 5** had low proportion of ACD patients and dependent residents, and had higher number of residents that returned home with their relatives; **cluster 6** were placed in areas with high household income and low population incidence of Covid-19; **cluster 7** had high proportion of CCP and ACD patients; nursing homes in this cluster were located in densely populated areas; **cluster 8** had a high SNQ12 score—indicating very limited capacity for pandemic preparedness and response—and high proportion of CCPs and younger residents than the median.

Association of nursing home clusters with mortality

The ANOVA analysis revealed significant differences in mortality (both all-cause and Covid-19-related mortality) among clusters (Figure 2). Clusters 1, 7, and 8, particularly characterized by a geographic location with low population density, higher percentages of residents with complex or advanced chronic diseases, and lower scores on the pandemic preparedness measure, respectively, had a greater Covid-19-related mortality rate than the median. Clusters 1 and 8—but not cluster

7—also reported higher all-cause mortality rates than the median, along with cluster 4, characterized by higher number of residents.

Discussion

In this analysis of 167 nursing homes, we provide a novel approach to understanding differences in mortality between nursing homes during the Covid-19 outbreak. Most studies on risk factors for Covid-19 mortality in the LTC setting did not simultaneously address potential confounding factors at the resident, organizational, or community level that may influence measured outcomes. Alternatively, we used multiple variables covering both facility characteristics and socioeconomic factors related to the geographic location to investigate the determinants of mortality in LTC facilities using a multivariable model and a clustering approach to detect patterns of LTC facilities that are at higher risk.

Our analysis revealed that a ten percent increase in the proportion of residents with complex conditions increased the mortality risk by 7%; a 1-point increase in the 12-points score of unmet measures for containing SARS-CoV-2 spread increased the mortality risk by 5%, an increase of 100 in the Covid-19 14-day incidence rate per 100,000 population increased the mortality risk by 240%.

In addition to the risk factors for mortality at the facility level, we aimed to identify nursing home profiles that were at higher risk of mortality to assist policy and healthcare planning in the advent of future outbreaks. Consistently with our regression analysis, clusters with greater mortality than the median (cluster numbers 1,4,7, and 8) were all located in neighbourhoods with high incidence

of Covid-19. These results align with previous studies that reported a significant relationship between Covid-19 incidence in the catchment area and the risk of an outbreak[9,10] and mortality[18] at the facility level. The influence of the local incidence of Covid-19 on mortality underscores the paramount importance of preventing SARS-CoV-2 entry into facilities—often with new residents, staff, or visitors—[8,18,19] and suggests that efforts to contain Covid-19 transmission in the community may also contribute to reducing Covid-19 deaths at their local nursing homes. The multiple regression and cluster analysis were also consistent regarding the importance of the capacity of the nursing home for pandemic preparedness and response, as described previously[3,6]. Although most nursing homes showed low SNQ12 scores—indicating few unmet needs for applying containment measures—facilities in cluster 8, characterized by higher SNQ12 scores (mean of 5 unmet items over a total of 12 essential requirements) than the median, experienced high mortality levels.

Finally, our multivariable analysis revealed a significant relationship between higher percentages of CCP patients and increased mortality risk. According to local clinical guidelines, CCP patients are characterized by high clinical complexity [20], suggesting an increased likelihood of death in the advent of any infection or acquired disease. Similar to our findings on clinical complexity profile and vulnerability, a cohort study of US nursing home residents with Covid-19, found that impaired cognitive and physical function were independently associated with mortality[11]. Interestingly, cluster no. 2, characterized by the higher health risk of its residents, had similar mortality than clusters 4, 5, and 6, with a more favourable resident health profile. These conflicting results suggest that the mechanisms driving mortality risk in nursing homes are complex and may depend on the conjunction of various factors.

Our analysis had the intrinsic limitations of retrospective studies, particularly regarding data completeness. Owing to the overload of the healthcare system during the investigated period, a large number of deaths could not be tested for SARS-CoV-2 PCR and remained unconfirmed. Another information potentially relevant for our analysis that could not be collected was the worker profiles in each nursing home. Unlike skilled nursing homes aimed at intermediate care or mental health resources, which tend to be coordinated by the healthcare authorities, non-specialized nursing homes aimed at long-term stay are a case-mix of organizational models. Hence, the inclusion of the characteristics of the work team profile (e.g., skills, resident/worker ratio, and presence of physicians) might have provided interesting insights regarding the capacity of the residence to cope with the outbreak[18]. In the absence of this information, we used the proportion of residents with complex or advanced diseases as a proxy to determine whether the nursing facility mostly provided intensive monitoring to residents with complex conditions or just assisted living to independent residents. Finally, it is noteworthy that, owing to the lack of validated scores for measuring preparedness to deal with Covid-19 outbreaks, we used a non-validated measure of preparedness.

Our results raise important policy implications by suggesting structural factors of the nursing homes and their surrounding districts that are important drivers of Covid-19-related mortality in this setting. Identification of facilities with low capacity for pandemic response and located in areas with high incidence of Covid-19 could help public health officers to identify facilities where preventative interventions need to be prioritized. The presence of complex patients also increased mortality risk.

References

1. World Health Organization (WHO). Coronavirus disease (COVID-19) situation report - 163. (2020). Available at: https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200701-covid-19-sitrep-163.pdf?sfvrsn=c202f05b_2. (Accessed: 24th July 2020)
2. Trabucchi, M. & De Leo, D. Nursing homes or besieged castles: COVID-19 in northern Italy. *The Lancet Psychiatry* **7**, 387–388 (2020).
3. Belmin, J. *et al.* Coronavirus Disease 2019 Outcomes in French Nursing Homes That Implemented Staff Confinement With Residents. *JAMA Netw. open* **3**, e2017533 (2020).
4. Abrams, H. R., Loomer, L., Gandhi, A. & Grabowski, D. C. Characteristics of U.S. Nursing Homes with COVID-19 Cases. *J. Am. Geriatr. Soc.* **68**, 1653–1656 (2020).
5. Chatterjee, P., Kelly, S., Qi, M. & Werner, R. M. Characteristics and Quality of US Nursing Homes Reporting Cases of Coronavirus Disease 2019 (COVID-19). *JAMA Netw. open* **3**, e2016930 (2020).
6. Lipsitz, L. A. *et al.* Stemming the Tide of COVID-19 Infections in Massachusetts Nursing Homes. *J. Am. Geriatr. Soc.* **68**, 2447–2453 (2020).
7. Lau-Ng, R., Caruso, L. B. & Perls, T. T. COVID-19 Deaths in Long-Term Care Facilities: A Critical Piece of the Pandemic Puzzle. *J. Am. Geriatr. Soc.* (2020). doi:10.1111/jgs.16669
8. McMichael, T. M. *et al.* Epidemiology of covid-19 in a long-term care facility in King County, Washington. *N. Engl. J. Med.* **382**, 2008–2011 (2020).
9. White, E. M. *et al.* Variation in SARS-CoV-2 Prevalence in U.S. Skilled Nursing Facilities. *J. Am. Geriatr. Soc.* 1–7 (2020). doi:10.1111/jgs.16752

10. Sugg, M. M. *et al.* Mapping community-level determinants of COVID-19 transmission in nursing homes: A multi-scale approach. *Sci. Total Environ.* **752**, 141946 (2020).

11. Panagiotou, O. A. *et al.* Risk Factors Associated with All-Cause 30-Day Mortality in Nursing Home Residents with COVID-19. *JAMA Intern. Med.* **02912**, 1–10 (2021).

12. Liu, K., Chen, Y., Lin, R. & Han, K. Clinical features of COVID-19 in elderly patients: A comparison with young and middle-aged patients. *J. Infect.* **80**, e14–e18 (2020).

13. He, M., Li, Y. & Fang, F. Is There a Link between Nursing Home Reported Quality and COVID-19 Cases? Evidence from California Skilled Nursing Facilities. *J. Am. Med. Dir. Assoc.* **21**, 905–908 (2020).

14. Brown, K. A. *et al.* Association between Nursing Home Crowding and COVID-19 Infection and Mortality in Ontario, Canada. *JAMA Intern. Med.* **181**, 229–236 (2021).

15. Li, Y., Temkin-Greener, H., Shan, G. & Cai, X. COVID-19 Infections and Deaths among Connecticut Nursing Home Residents: Facility Correlates. *J. Am. Geriatr. Soc.* **68**, 1899–1906 (2020).

16. Unruh, M. A., Yun, H., Zhang, Y., Braun, R. T. & Jung, H. Y. Nursing Home Characteristics Associated With COVID-19 Deaths in Connecticut, New Jersey, and New York. *J. Am. Med. Dir. Assoc.* **21**, 1001–1003 (2020).

17. Brandén, M. *et al.* Residential context and COVID-19 mortality among adults aged 70 years and older in Stockholm : a population-based , observational study using individual-level data. *Lancet Heal. Longev.* **1**, 80–88 (2020).

18. Gorges, R. J. & Konetzka, R. T. Staffing Levels and COVID -19 Cases and Outbreaks in US Nursing Homes. *J. Am. Geriatr. Soc.* jgs.16787 (2020). doi:10.1111/jgs.16787

19. Centers for Disease Control and Prevention (CDC). Discontinuation of Transmission-

Based Precautions and Disposition of Patients with COVID-19 in Healthcare Settings (Interim Guidance). (2020). Available at: https://www.cdc.gov/coronavirus/2019-ncov/hcp/disposition-hospitalized-patients.html. (Accessed: 27th July 2020)

20. Catalan Ministry of Health. Health Plan for Catalonia 2016-2020. A person-centred system: public, universal and fair. (2016). Available at: https://salutweb.gencat.cat/web/.content/_departament/pla-de-salut/Pla-de-salut-2016-2020/documents/health-plan-catalonia_2016_2020.pdf. (Accessed: 3rd August 2020)
21. Catalan Health Service. eCAP. *2018* Available at: https://salutweb.gencat.cat/ca/ambits_actuacio/linies_dactuacio/tecnologies_informacio_i_comunicacio/ecap/. (Accessed: 9th May 2020)
22. Seematter-Bagnoud, L. *et al.* Predictors of functional recovery in patients admitted to geriatric postacute rehabilitation. *Arch. Phys. Med. Rehabil.* **94**, 2373–2380 (2013).
23. Catalan Health Service. Pla d'acció per la gestió de persones en l'àmbit residencial i l'assistència sanitària en situació de pandèmia per COVID-19. (2020). Available at: https://scientiasalut.gencat.cat/handle/11351/4847. (Accessed: 3rd July 2020)
24. Echeverría, P. *et al.* COVIDApp as an innovative strategy for the management and follow-up of COVID-19 cases in long-term care facilities in Catalonia: Implementation study. *JMIR Public Heal. Surveill.* **6**, 1–8 (2020).
25. Levine, J. H. *et al.* Data-Driven Phenotypic Dissection of AML Reveals Progenitor-like Cells that Correlate with Prognosis. *Cell* **162**, 184–197 (2015).
26. Blondel, V. D., Guillaume, J. L., Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* **2008**, P10008 (2008).
27. Breiman, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).

28. R Core Team. R: A language and environment for statistical com-puting. *R Foundation for Statistical Computing, Vienna, Austria.* (2017). Available at: https://www.r-project.org. (Accessed: 25th May 2020)

Acknowledgments

The authors would like to thank Gerard Carot-Sans (PhD) for providing medical writing support during the preparation of the manuscript. Funding sources: Crowdfunding campaign YoMeCorono (*https://www.yomecorono.com/*), and Generalitat de Catalunya. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Authors Contributions

CS, DO, MAM, RLA and OM designed the study. DO, MAM, RLA, EN, MMM, NPG, JMB-S, MEI, IGS, SRN, RM, MTC, JVP, NH, JA collected the data. CS, DO, MM, MNJ, OM analysed the data. CS, DO, MM, MNJ, MAM, MMM, RM and OM interpreted the data. CS, DO, OM wrote the manuscript. All authors reviewed and approved the final version of the manuscript.

Declaration of interests

The authors declare no competing interests.

Figure Legends

Figure 1. Heatmap of nursing home clusters. (A) Heatmap of nursing home clusters. For each characteristic (x-axis), the extent of the difference between the mean of a given cluster and the median of the entire sample is illustrated with the following colour code: green tones indicate a mean of the cluster below the median of the entire sample, whereas purple tones indicate a mean of the cluster above the median of the entire sample. In both cases, more intense colours represent greater differences between the cluster and the whole sample. For each cluster, the number of nursing homes is indicated (n, % of total). Supplementary Figure 2 provides further details regarding the mean (SD) of the characteristics in each cluster. (**B**) Contribution of variables to each cluster. Percentage is shown for the three variables with greater weight. Black squares highlight the most important variable.

Figure 2. Mortality at the cluster level for all-cause deaths and Covid-19 related deaths. The dotted line shows the median mortality rate for the entire sample. Bars show the mean mortality rate of each cluster; error lines represent the standard error of the mean (SEM). Red and green indicate mean cluster mortality higher and lower than the total median, respectively. For both colours, light tones indicate that the SD of the cluster encompasses the overall median, whereas intense tones indicate that the whole SD range is above (red) or below (green) the total median. Differences between groups were determined by one-way ANOVA. All-cause deaths: ($F(7,167) = 2.83$; *p*-value = 0.009. Covid-19-related deaths: ($F(7,167) = 3.22$; *p*-value = 0.003. Number of nursing homes per cluster [Cluster (n)]: 1(13), 2(16), 3(25), 4(33), 5(12), 6(18), 7(36), 8(14).

Tables

Table 1. Characteristics of nursing homes

	Total (N = 167)
Individual residents' characteristics	
Age of residents, in years	87.1 (2.1)
Percentage of male residents	26.4 (9.6)
Number of comorbidities	1.5 (0.6)
Nursing home characteristics and type of care provided	
Percentage of dependent residents*	82.1 (9.5)
Percentage of CCPs	46.1 (17.3)
Percentage of ACD patients	10.5 (8.9)
Current number of residents	46.2 (29.8)
Pandemic preparedness by nursing home	
SNQ12 score	1.4 (1.7)
Percentage of residents who return home	1.4 (3.0)
Geographic location characteristics	
Mean household income, in Euro	36099.6 (5527.5)
Density of population per km^2, in log_{10}	17.9 (9.5)
Incidence rate of Covid-19, in 14-day cases per 100,000 population	900 (300)

Variables are described as mean (SD), computed as mean of the percentages in each facility (categorical variables) or mean of means in each facility (continuous variables). CCP=complex chronic patient. ACD=advanced chronic disease. SNQ12=number of unmet epidemic and infection control preparedness requirements (0-12 scale). *Barthel score < 50.

Table 2. Estimated effect of long-term care facilities' characteristics in all deaths and Covid-19 related deaths at facility level.

	All-cause deaths			Covid-19-related deaths		
	Univariate analysis	Multivariable analysis		Univariate analysis	Multivariable analysis	
	IRR (95% CI)	IRR (95% CI)	*p*-value	IRR (95% CI)	IRR (95% CI)	*p*-value
Individual residents' characteristics						
Age of residents, in years†	1.00 (0.97-1.02)	..	..	0.98 (0.95-1.02)	..	..
% of male residents	1.06 (1.00-1.12)*	1.07 (1.01-1.13)	0.0263	1.03 (0.95-1.12)	..	..
Number of comorbidities†	1.18 (1.08-1.28)*	..	..	1.31 (1.18-1.46)*	..	..
Nursing home characteristics and type of care provided						
% Dependent Residents	1.00 (0.95-1.05)	..	..	1.00 (0.93-1.07)	..	..
% CCP Residents	1.04 (1.01-1.06)*	1.07 (1.03-1.10)	<0.0001	1.06 (1.02-1.09)*	1.07 (1.04-1.11)	0.0001
% ACD Residents	1.06 (1.00-1.11)	..	..	1.10 (1.02-1.18)*	..	..
Current number of residents	1.02 (1.00-1.03)*	1.02 (1.01-1.03)	0.0098	0.95 (0.98-1.01)	..	..
Pandemic preparedness by nursing home						
SNQ12†	1.06 (1.04-1.09)*	1.05 (1.03-1.08)	0.0001	1.07 (1.04-1.10)*	1.05 (1.02-1.09)	0.0005
% of residents who return home	0.95 (0.79-1.13)	..	..	0.86 (0.67-1.10)	..	..
Geographic location characteristics						
Mean household income, in Euro‡	1.00 (0.99-1.00)	..	..	0.99 (0.98-1.00)	..	..
Population density, log10 people/km2§	1.06 (0.89-1.27)	..	..	1.00 (0.80-1.26)	..	..
Population incidence of Covid-19‖	1.89 (1.63-2.20)*	1.88 (1.61-2.19)	<0.0001	2.25 (1.85-2.73)*	2.39 (1.95-2.93)	<0.0001

CCP= complex chronic patient; ACD= advanced chronic disease; Dependent resident=Barthel score < 50; SNQ12= number of unmet essential items for implementing preventive measures (0-12).
Incidence rate ratio (IRR) and 95% CI are shown for the zero inflated Poisson model.
IRR represents the estimated effect for an increase of 10 units, unless otherwise indicated.
† IRR for an increase in 1 unit.
‡ IRR for an increase of 10,000€/annum in mean household income.
§ IRR for an increase in 1 $\log_{10}$ people/km^2.
‖ IRR for an increase in 100 cases of Covid-19 14-day incidence rate per 100,000 population.
*p-value for univariate analysis <0.05
Omnibus test: All-cause death model (χ^2-statistic: 1332; *p*-value<0.0001) and Covid-19-related death model (χ^2-statistic: 2211; *p*-value<0.0001).

Methods

Study setting and participants

This was a retrospective cohort study of Covid-19 mortality risk factors in the residential setting in Catalonia (North-East Spain). The study included clinical, mortality, and structural information corresponding to all public and private nursing homes in the administrative health region *Metropolitana Nord* (population 1,986,032 people) in Barcelona, Spain between March 1 and June 1, 2020, during the Covid-19 outbreak. Skilled nursing facilities (i.e., intermediate care) and mental health facilities were excluded from the analysis.

On March 1, 2020, the Department of Health of Catalonia launched a comprehensive disease control program to minimize Covid-19 spread and mortality among residents in nursing homes. The containment strategy was implemented in all LTC facilities in the study area and involved 64 primary care teams that reported daily information regarding the epidemiological status of each nursing home. The primary care teams provided preventive epidemiological recommendations, including the partition of communal living areas, isolation of suspected cases and contacts, guidance on personal protective measures to nursing home workers. In the advent of a confirmed or suspected case of Covid-19, the teams also conducted systematic screening of close contacts—or all residents, in centres with high incidence—using real-time reverse transcription–polymerase chain reaction (rt-PCR) from nasopharyngeal swabs. No systematic screenings with PCR were conducted in nursing homes without suspected or confirmed Covid-19 cases.

The study protocol was approved by the institutional review board of Hospital Universitari Germans Trias Pujol.

Data collection

Demographic and clinical data of residents were extracted from electronic medical records using a standardized data collection form [21]. The structural and organizational features of each nursing home were gathered at facility assessment visits by the study team. The demographic and epidemiological profile of the district where the nursing home was located was retrieved from the Statistical Institute of Catalonia. Deaths were identified from the Mortality Registry of the Department of Health of Catalonia where public all deaths at a country-level are systematically registered, irrespective of the place of death. All data were handled according to the General Data Protection Regulation 2016/679 on data protection and privacy for all individuals within the European Union and the local regulatory framework regarding data protection. All data is available in Supplementary Appendix 2.

Definitions

Variables regarding residents' health characteristics in each nursing home included demographic characteristics (i.e., age and gender), and clinical characteristics (i.e., number of comorbidities). Comorbidities were codified according to the ICD-10 system and included dementia, asthma or chronic obstructive pulmonary disease, hypertension, type-1 diabetes mellitus, type-2 diabetes mellitus, chronic kidney disease, cerebrovascular disease, cardiovascular disease.

We also recorded, nursing home characteristics according to the total number of residents and type of care provided. The latter was assessed based on the proportion of residents classified by their primary care teams as residents with high dependence in activities of daily living (i.e., defined as a Barthel score<50)[22], complex chronic patients (CCP), and patients with advanced chronic disease

(ACD), according to clinical guidelines of the Catalan Health Department [20]. The Catalan guidelines define CCP based on their clinical condition (e.g., multimorbidity, disability, difficult symptom control) and/or social environment (e.g., lack of support from family or caregivers, isolated household). Patients with ACD are those with advanced and irreversible chronic conditions that limit their life expectancy to approximately 12 months.

Pandemic preparedness by nursing homes were characterized according to their capacity for implementing preventative measures within the facility on the basis of an ad-hoc score (SNQ12 score) and organizational variables related to the percentage of residents who returned home to live with their relatives as a protective measure against Covid-19. The capacity of the nursing home for pandemic preparedness and response was assessed using a non-validated *ad hoc* set of 12 essential items that yields a score, called SNQ12 (*sine qua non* conditions for implementing the measures, Supplementary Appendix)[23]. The score indicates the number of unmet requirements, which ranges from 0 (all requirements are met) to 12 (all requirements are unmet). The requirements are related to three areas: 1) personal protective equipment (PPEs), including a waterproof gown, face mask (i.e., either filtering facepiece 2 or surgical), gloves, and protective glass (adequate supply, routine use, and use for waste management and cleaning/disinfection), 2) surveillance and communication (routine monitoring of symptom onset by non-healthcare professionals and communication of symptoms to occupational health services), and 3) cleaning and waste management (regular hand washing before and after contact with Covid-19 patients or their contacts, adequate laundry procedures, cleaning and disinfection of surfaces, use of an adequate disinfectant, adequate disposal of used PPEs). The SNQ12 score was developed by the Catalan Ministry of Health and distributed to medical directors of the nursing homes along with the request of returning the result within the next 14 days. The number of residents returning home

was recorded by LTCF with the Covid-19 App[24]. These data were also included in the electronic health records.

The geographic location of the nursing home was assessed and characterized using the mean household income and density of population in the municipality, and the Covid-19 14-day incidence rate per 100,000 in the post code district (lowest administrative division) where the nursing home is located.

Excess mortality directly or indirectly due to Covid-19 was classified as Covid-19-related when individuals had a positive rt-PCR or a clinical suspicion of Covid-19. Clinical suspicion of Covid-19 was defined based on the national guidelines available at the time as individuals with clinical features of acute respiratory disease of sudden onset and any severity, primarily characterized by fever, cough, and shortness of breath. Other symptoms such as odynophagia, anosmia, dysgeusia, muscular pain, diarrhoea, chest pain, or headache could also be considered suggestive of SARS-CoV-2 at the physician's discretion.

Statistical Analysis

Continuous and categorical variables were presented as the mean and standard deviation (SD) (or median and interquartile range [IQR], defined by 25^{th} and 75^{th} percentiles) and number (%), respectively. The excess deaths were defined as the difference between deaths reported in 2020 and the median of 2016-2019 for the same months of the year; the Covid-19 contribution to the excess deaths was computed by the difference between confirmed or suspected Covid-19 deaths and all-cause mortality. In our primary analysis to determine the risk factors associated with mortality, we used univariate and multivariable zero inflated Poisson models at facility level. Independent variables for the multivariable model were chosen using an Akaike Information

Criteria (AIC)-based backward stepwise procedure. Continuous variables were fitted assuming a linear relationship. Results were presented as the incidence rate ratio (IRR) and the 95% confidence interval (CI). In a supplementary analysis we fitted a model to individual level data using a random-effects logistic regression with clustering at the nursing home level. In a secondary analysis, we grouped the nursing homes according to their characteristics using the unsupervised clustering algorithm phenograph[25,26]. Data was centred and scaled prior to running the algorithm. To validate the results, we used different values of the k parameter (number of nearest neighbours) and chose one based on the mean shortest distance and the stability in the number of clusters (supplementary results 1). The resulting clusters were described in a heatmap that represents the intensity of each characteristic based on the difference (below or above) between the average of the given cluster and that of the overall sample. To facilitate interpretation, we used random forest classifier to determine the weight of each variable in each cluster based on the Gini measure [27]. Mean mortality rate of each cluster was calculated and differences between groups were determined by one-way ANOVA. The significance threshold was set at a two-sided alpha value of 0.05. All analyses and plots were performed using R version 3.6[28].

Data availability

All data generated or analysed during this study are included in this published article (and its supplementary information files).

Code availability