Supplementary Appendix

Supplementary Table 1. Demographic and clinical characteristics of residents included in the analysis.

	Total (N=8.716)
Demographics	
Age in years, Mean (SD)	85.7 (9.0)
Percentage of residents aged > 80 years, Mean (SD)	80.5 (8.8)
Males	2333 (26.8)
Clinical Characteristics	
Comorbidities	
Number of comorbidities, *n (%)*	
0	2142 (24.6)
1	2528 (29.0)
2	1869 (21.4)
>2	2177 (25.0)
Type of Comorbidity* *n (%)*	
Dementia	3954 (45.4)
COPD or asthma	885 (10.2)
Hypertension	4195 (48.1)
Diabetes mellitus type 1	19 (0.2)
Diabetes mellitus type 2	1704 (19.6)
Chronic kidney disease	1707 (19.6)
Cerebrovascular diseases	231 (2.7)
Cardiovascular diseases	1228 (14.1)
Recipients of specific healthcare programs, *n (%)*	
Complex Chronic Patients (CCPs)	3703 (42.4)
Advanced Chronic Disease (ACDs)	752 (8.6)
Medication	
Polypharmacy (≥5 drugs), *n (%)*	7305 (83.8)
Prescription of drugs of interest†, *n (%)*	
Angiotensin-converting-enzyme inhibitors	2184 (25.1)
Antiplatelets	2602 (29.9)
Anticoagulants	1146 (13.1)
Inhibitors of the sodium-glucose transport protein 2	16 (0.2)
Angiotensin II receptor blockers	1098 (12.6)
Oral Corticoids	220 (2.5)
Clozapine	17 (0.2)

Data are n (%) unless otherwise specified.
*Categories are not mutually exclusive.
†Drugs of interest are those with potential effect on the course of Covid-19 disease.

Supplementary Table 2: Estimated effect of individual and facility characteristics in Covid-19 related deaths at individual level.

	Univariable			Multivariable		
	OR	95% CI	p-value	aOR	95% CI	p-value
Individual Level Factors						
Age	1.015	1.00-1.027	0.0138	1.02	1.01-1.04	0.0002
Male Gender	0.566	0.47-0.69	<0.0001	0.51	0.42-0.63	<0.0001
Number of Comorbidities	1.1	1.02-1.17	0.009	1.08	1.00-1.15	0.036
Nursing Home Characteristics						
% Dependent Residents	0.99	0.95-1.03	0.65			
% CCP Residents	1.01	0.99-1.03	0.53			
% ACD Residents	1.02	0.98-1.07	0.25			
Number of residents	1.01	0.99-1.02	0.36			
Pandemic Preparedness						
SNQ12	1.22	1.01-1.48	0.044	1.18	0.99-1.43	0.062
% of residents who left	1.04	0.92-1.16	0.054	1.07	0.96-1.19	0.2
Geographic Location Characteristics						
Household Income	1	0.94-1.07	0.94			
Population Density	1.03	0.29-3.74	0.96			
Population Incidence of Covid-19	11.34	3.54-36.33	<0.0001	10.88	3.41-34.72	<0.0001

Odds ratio (OR) are shown for the multi-level model fitting an individual participant regression. The model controls for both individual and facility level characteristics.

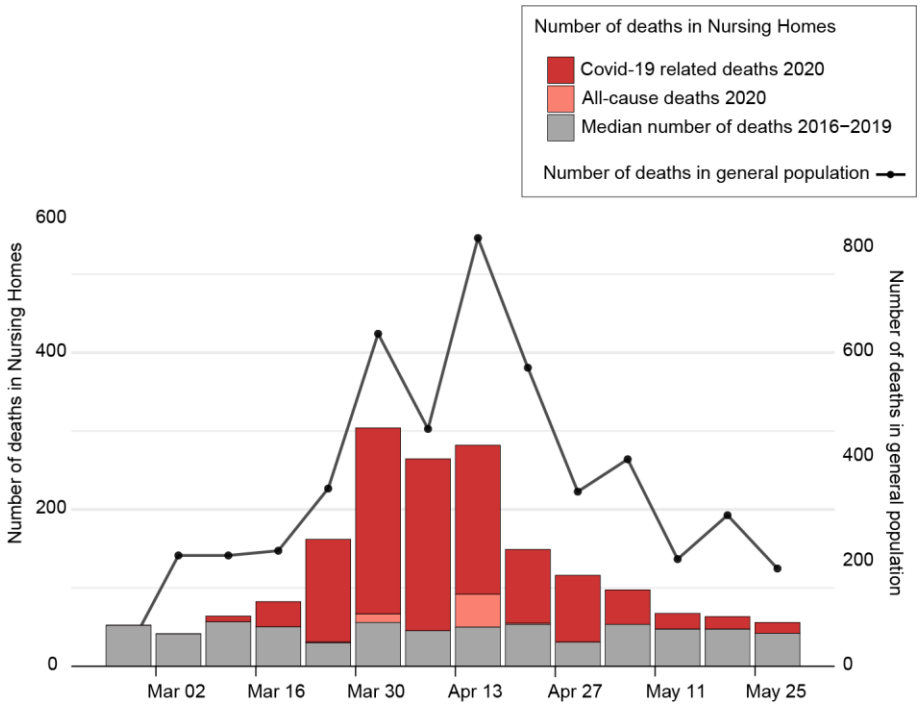
Supplementary Table 3. STROBE Statement—Checklist of items that should be included in reports of cohort studies

	Item No	Recommendation
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported
Objectives	3	State specific objectives, including any prespecified hypotheses
Methods		
Study design	4	Present key elements of study design early in the paper
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group
Bias	9	Describe any efforts to address potential sources of bias
Study size	10	Explain how the study size was arrived at
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding
		(*b*) Describe any methods used to examine subgroups and interactions
		(*c*) Explain how missing data were addressed
		(*d*) If applicable, explain how loss to follow-up was addressed
		(*e*) Describe any sensitivity analyses
Results		
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed
		(b) Give reasons for non-participation at each stage
		(c) Consider use of a flow diagram
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders
		(b) Indicate number of participants with missing data for each variable of interest
		(c) Summarise follow-up time (eg, average and total amount)
Outcome data	15*	Report numbers of outcome events or summary measures over time
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included
		(*b*) Report category boundaries when continuous variables were categorized

		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses
Discussion		
Key results	18	Summarise key results with reference to study objectives
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence
Generalisability	21	Discuss the generalisability (external validity) of the study results
Other information		
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based

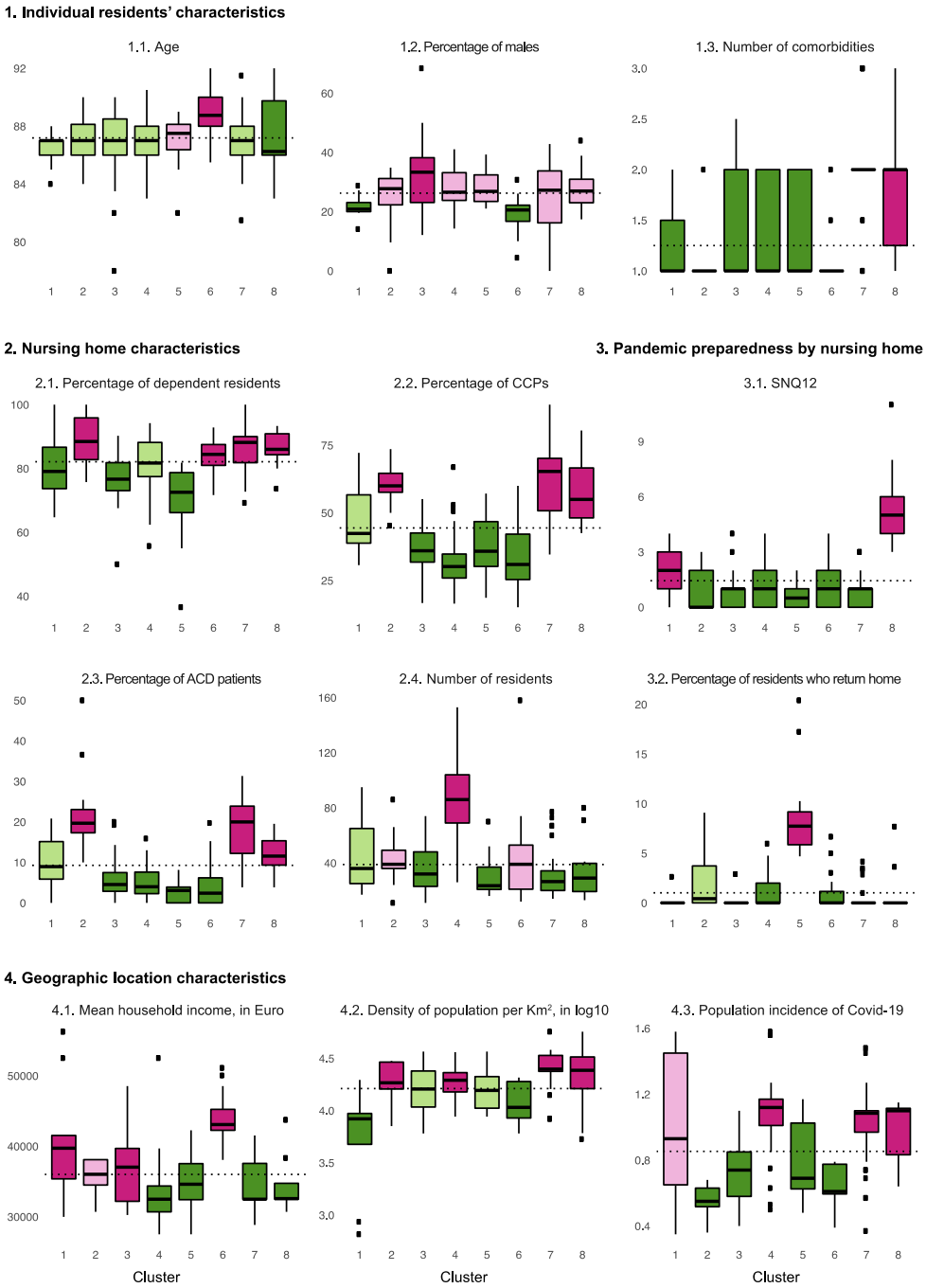
*Give information separately for exposed and unexposed groups.

Supplementary Figure 1. Excess mortality of 2020 relative to the average of the past four years (2016-2019).



Supplementary Figure 1. Excess mortality of 2020 relative to the average of the past four years (2016-2019). Bars show the number of weekly deaths reported in 2020 in all nursing homes included in the analysis over the study period. Deaths reported in 2020 have been classified as Covid-19 confirmed and unconfirmed, which include deaths of individuals with suspected Covid-19 diagnosis. The median number of deaths for the same weeks in the previous 4 years (2016-2019) is shown in grey. The continuous line shows the death toll attributed to Covid-19 in the general population of the catchment area.

Supplementary Figure 2. Characteristics among nursing home clusters.



Supplementary Figure 2. Characteristics among nursing home clusters. For each cluster (x-axis), bar plots show the mean (SD) of a given characteristic. The median of the entire sample is shown with a dashed line. 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.

Supplementary Methods 1. SNQ12 Questionnaire

1) Health staff wears personal protective equipment (PPE), including face masks and gloves for low-risk exposures.

2) The onset of Covid-19-related symptoms is notified to occupational health services

3) All staff members regularly wash their hands with soap and/or alcohol-based antiseptic solution before and after any interaction with Covid-19 patients or their contacts.

4) Resident's relatives and non-healthcare professionals of the nursing home self-monitor the onset of symptoms associated with an acute respiratory infection (e.g., fever, sore throat, cough, and shortness of breath) and notify them to healthcare services.

5) The nursing home has a laundry procedure to wash bedding, towels, etc. of ill residents with regular soaps or detergents at 60-90 ºC and dry them thoroughly.

6) The patient’s wastes are disposed using adequate PPEs.

7) The patient's wastes are disposed of together with all domestic waste in a plastic bag, which is closed before leaving the room.

8) Personal protective equipment (PPE) used to care for Covid-19 patients is disposed of using a double bag.

9) All areas and surfaces that have been in contact with a Covid-19 case are exhaustively cleaned.

10) The disinfectant solution regularly used contains bleach or a solution of sodium hypochlorite that contains 1,000 ppm of active chlorine (e.g., newly prepared 1:50 water solution of bleach that contains 40-50 g/L of sodium hypochlorite).

11) Cleaning staff is protected with a liquid-resistant coat, facial mask, and gloves.

12) An adequate supply of personal protective equipment is foreseen.

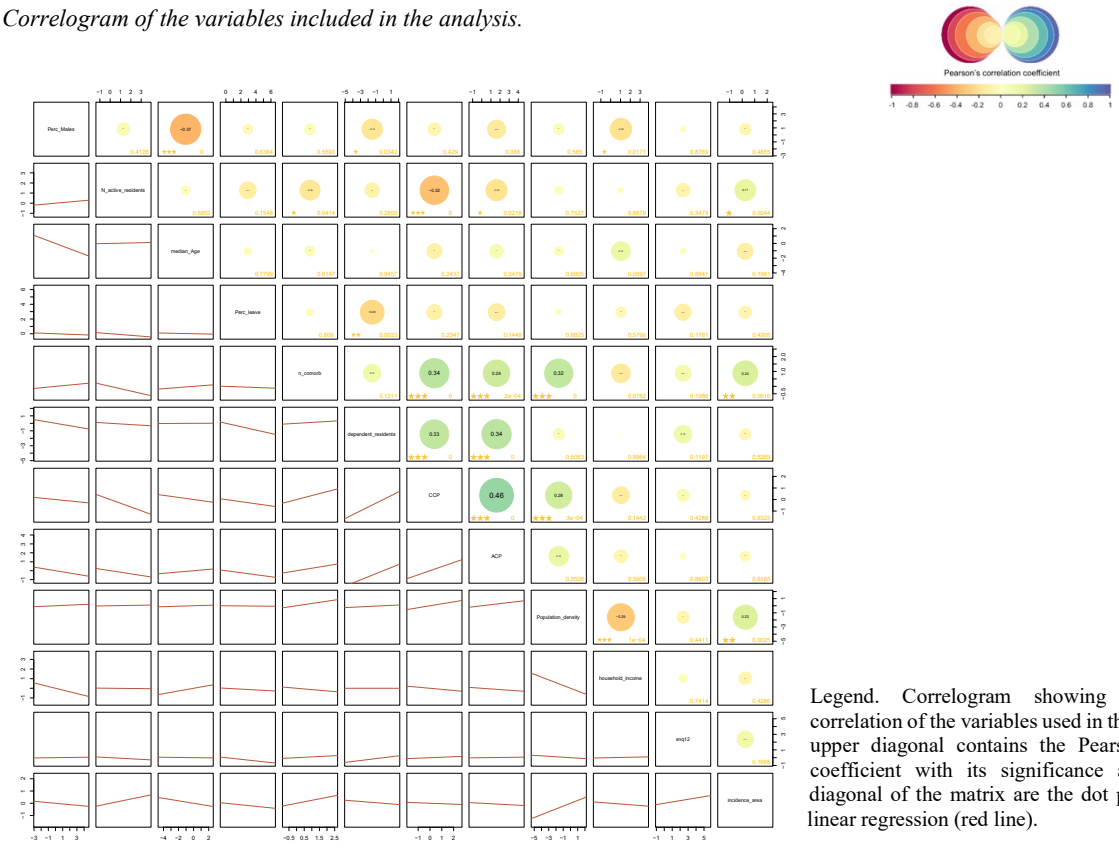
Supplementary Methods 2.

Pre-processing: The data were normalized (centred and scaled) before performing the algorithm.

Clustering Algorithm: The *Phenograph* unsupervised clustering algorithm was used. *PhenoGraph* takes as input a matrix of measurements and divides them into subpopulations by clustering according to their similarity. *PhenoGraph* builds this clusters graph in two steps. First, it finds the k nearest neighbours for each data point (care facility), using Euclidean distances, resulting in N sets of k-neighbourhoods. Second, it operates on these sets to build a weighted graph such that the weight between nodes is balanced with the number of neighbours they share. The Louvain community detection method (Blondel et al., 2008) is then used to find a partition of the graph that maximizes nodularity and results in the optimal number of differentiated clusters.

Correlation between variables: We used all potential risk factors included in the mortality model. Before fitting the model, we characterized pairwise correlation between variables to control for multi-colinearity. Overall, the median correlation between variables was 0.086 [q1-q3, 0.039 – 0.164]. The maximum value of Pearson's correlation coefficient, 0.456, was observed between CCP and ACP variables (the correlogram below shows each pairwise correlation).

Correlogram of the variables included in the analysis.



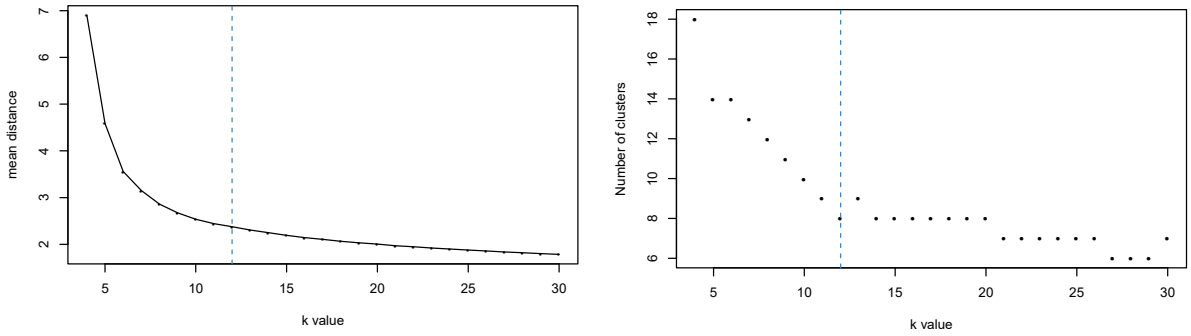
Legend. Correlogram showing the pairwise correlation of the variables used in the analysis. The upper diagonal contains the Pearson correlation coefficient with its significance and the lower diagonal of the matrix are the dot plots with their linear regression (red line).

After clustering: To search for an optimal value, we studied k-NN values from 4 to 30 and measured the number of clusters, nodularity and the mean of shortest distances between nodes.

Cluster dimension: If we look at the graph of the number of clusters vs parameter *k*, it seems that the number of clusters can be reduced from 8 to 6-7. Moreover, the LTC facilities distribution in 6-7 clusters was consistent in composition to the distribution with 8 clusters, except for LTC facilities in clusters 1 and 8 that tend to merge with nursing homes in clusters 3 and 7, respectively. Looking at heatmap, 1-3 and 7-8 profiles are very similar so we understand that these facilities could be combined in a unique cluster. However, since we are making a description of

the profiles based on mortality, we consider important to differentiate between those with differences in SNQ12 and Covid-19 incidence, variables that were shown to be important in the risk model. Importantly, no significant changes in the interpretation using a smaller number of clusters are expected. Thus, we decided to proceed with 8 clusters given by the algorithm under k =12.

k-NN value against mean distance and number of clusters. Vertical line at k=12



In addition, a further evaluation of the clusters was done in terms of robustness of the method by correlating the Phenograph clusters with a standard k-means method. While some variability between methods is expected, the contingency table shows robustness in cluster member composition, since virtually all clusters are composed by the same LTC facilities (accuracy=89.9, Chi-square = 870).

Contingency table of the clusters obtained using k-means against Phenograph clusters.

	Phenograph clusters							
k-means clusters	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1	8	0	0	0	0	0	0	0
2	0	16	0	0	1	0	0	1
3	3	0	23	0	2	0	0	0
4	0	0	1	33	1	0	0	2
5	0	0	0	0	8	0	0	0
6	1	0	0	0	0	18	0	0
7	1	0	1	0	0	0	36	4
8	0	0	0	0	0	0	0	7