BMC Health Services Research

RESEARCH

Open Access

Distinct phenotypes of patients and healthcare resource utilization after hospitalization for COVID-19: an observational study

Claire Marant Micallef[1*], Manon Belhassen[1], Florence Ader[2,3], Valeria Martinez[4,5,6], Eric Van Ganse[7,8], Marjorie Bérard[1], Mélanie Née[1], Mikhail Dziadzko[6,9] and Frédéric Aubrun[8,9]

Abstract

Background Little is known about postdischarge healthcare resource use (HCU) among patients hospitalized for coronavirus disease 2019 (COVID-19). The objective was to identify distinct profiles of patients based on postdischarge cares.

Methods This was a retrospective cohort study using the French National Health System claims database. We followed up all patients hospitalized for COVID-19 between 2020/02/01 and 2020/06/30 for 6 months; the discharge date was the index date. We excluded patients who died during the index stay or within 30 days after discharge. We described patients' HCU over 5 months from day 31 after the index date to the end of follow-up, i.e., the post-COVID-19 period. We described the sociodemographic and clinical characteristics of the participants and 44 selected types of HCU, including medical and emergency room visits, medications, medical and biological tests, oxygen therapy, rehabilitation, rehospitalization, nurse visits, and sick leave. We performed Ward's ascendant hierarchical clustering (AHC) analysis to identify groups of patients with similar post-COVID-19 HCU and described HCU and clinical characteristics by cluster.

Results The study population included 68,822 patients (median age: 64.8 years, 47% women). Eight clusters of patients were identified, each comprising between 1,163 and 35,501 patients. Four clusters were characterized by older patients and high proportions of comorbidities, i.e. cancer (cluster 3), mental disorders (cluster 4), cardiac insufficiency (cluster 5) and respiratory failure (cluster 6). Cluster 8 was characterized by younger patients, often obese and with low mortality. Another cluster was characterized by complex index stays (cluster 7) and a last cluster (cluster 2) by specific medical contacts and therapy. The main cluster (cluster 1, $n=35,501$) was similar to the overall study population. The duration and complexity of the index stay also varied across clusters.

Conclusions Based on HCU data, AHC identified 8 clinically relevant profiles of patients surviving the acute episode of COVID-19 hospitalization. The clusters illustrate the many impacts of COVID on the health status of infected patients and may help anticipate future needs of care in a similar context.

Keywords SARS-C, Healthcare resource use, Cluster analysis

*Correspondence:
Claire Marant Micallef
claire.marant-micallef@pelyon.fr
Full list of author information is available at the end of the article

Background

The coronavirus disease 2019 (COVID-19) pandemic caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) spread to Europe in early 2020, leading to a rapid increase in hospital admissions and major disruptions in the healthcare system. In France, more than 200,000 patients were hospitalized due to COVID-19 in 2020 [1]. Age, male sex, and specific comorbidities (overweight, hypertension, cardiovascular disease, respiratory disease) were early associated with an increased risk of hospitalization [2, 3]. However, until recently, little has been known about the progression of patients following discharge, i.e. which types of cares they will need, at which frequency, and how much time. In particular, knowledge is missing about types of cares needed according to patients' profile.

A range of studies have distinguished several phenotypes of patients following SARS-CoV-2 infection based on self-reported or observed symptoms in the months following acute infection [4–6]. However, large-scale studies investigating post-COVID-19 conditions may be complex, costly, and time-consuming due to the broad range of outcomes to be considered following SARS-CoV-2 infection [7–14]. Exploring the healthcare resource use (HCU) of patients through electronic claims databases may help address these issues. While short-term HCU after acute COVID-19 [15, 16] has been investigated, post-COVID-19 HCU over several months has not yet been comprehensively described, e.g., whether severe COVID-19 impacts postdischarge HCU and whether profiles of HCU can be distinguished.

In this context, it is of major interest to analyze the long-term postdischarge HCU of COVID-19 inpatients to detect and understand specific patient profiles. This will provide key material to frontline healthcare workers and policymakers for the care of patients after massive and unexpected numbers of acute hospitalizations due to a new viral pulmonary disease for which no vaccine is available.

The French Healthcare System Claims Database (SNDS) contains HCU claims of nearly all the French population, including information on hospital discharge diagnoses. This dataset is thus strongly relevant to study HCU of patients hospitalized for COVID-19. Using SNDS, we recently showed that a range of new care was commonly initiated after discharge in patients hospitalized for COVID-19 in France between February and June 2020 [17]. Using claims of the full population of patients hospitalized for COVID-19 in France at the beginning of the pandemic, the aim of the present study was to further analyze these results to identify groups of patients with similar HCU after discharge, including visits, medications, medical procedures, biological tests, oxygen support therapy and rehabilitation.

Methods

Study design and data sources

This retrospective, population-based cohort study was based on SNDS claims database. It contains anonymous individual information on sociodemographic characteristics, all nonhospital reimbursed healthcare expenditures (without corresponding medical diagnoses), and all hospital discharge summaries with related International Classification of Diseases diagnoses (for more than 98% of the French population. The SNDS does not provide information on behavioral or clinical characteristics (tobacco smoking, body mass index, etc.) or laboratory results [18].

Study population and periods

The study population consisted of all patients hospitalized for COVID-19. We selected all stays with the ICD-10 codes U07.10 (respiratory form of COVID-19, virus identified), U07.11 (respiratory form of COVID-19, virus not identified), U07.14 (other form of COVID-19, virus identified), U07.15 (other form of COVID-19, virus not identified), as the main diagnosis, starting between 01 February 2020 and 30 June 2020. The discharge date was defined as the index date. As we would like to focus on long-term HCU following hospitalization for COVID-19, patients who died during hospitalization or within 30 days following discharge were excluded. We also excluded those with a pregnancy identified between 01 February 2019 and and 31 December 2020, and those with no HCU in the 12 months before the index date were excluded, as one of the initial goal of the study was to compare HCU in both the pre- and post-COVID periods. Patients were followed up until death or for a maximum of 6 months. The 5-month period starting on day 31 after the index date and ending with follow-up was defined as the post-COVID-19 period, and the same 5-month period 12 months before the post-COVID-19 period was defined as the pre-COVID-19 period [17].

Study variables

Variables related to patients' sociodemographic and clinical characteristics, i.e., age at index date, sex, free-access-to-care status (as a proxy of social deprivation), death during follow-up, and characteristics of the index stay were described as proxies of the severity of COVID-19: length of stay, requirement for oxygen support (high-flow oxygen devices, noninvasive and

invasive mechanical ventilation) during the hospital stay, admission to the intensive care unit (ICU), Simplified Acute Physiology Score (SAPS II) score (severity score mostly based on clinical characteristics [19]), and Charlson Comorbidity Index [20]. Complex index stays were defined by stays of long duration, the need for oxygen support (especially intubation), admission to ICU and high SAPS II score.

As detailed in the previous article, seven categories of HCU were identified and described over the pre- and post-COVID-19 periods [17]: medical visits, medications, medical procedures, biological tests, oxygen support therapy, rehabilitation, and others (rehospitalization, nurse visits, sick leave and emergency room visits).

Statistical analyses

Descriptive statistics of patients' sociodemographics, clinical characteristics and HCU were provided for the pre- and post-COVID-19 periods in the entire population, i.e. in all patients included; for the number and proportion of patients with at least one care session (i.e. patients for whom at least one visit/care/medication/procedure/hospitalization was identified) over the two periods. The mean, standard deviation, median and quartile of the monthly number of cares were provided [17].

Principal Component Analysis (PCA) was performed to minimize the initially identified 60 variables to be used for the clustering analysis, using the monthly HCU over the post-COVID-19 period. Using the K-means classification method, patients with outlying values of HCU were excluded [21, 22]. Finally, a clustering analysis was performed using Ward's ascendant hierarchical clustering (AHC) method [23] to obtain similar groups of patients based on their HCU. The optimal number of clusters was determined using the CCC criteria, pseudo-F test and pseudo-R-squared values [24].

For each identified cluster, sociodemographic and clinical characteristics were described, as was the proportion of patients with at least one care episode in each HCU category during the post-COVID-19 period. Sociodemographic and clinical characteristics were compared across clusters using chi-square tests and ANOVA.

All analyses were performed using SAS (SAS Institute, North Carolina, USA), version 9.4.

Ethics and consent to participate

This study was approved by the ethics committee of the French Health Data Hub (no. 4653731, 2021/07/08). It was conducted using anonymized data and approved by the ethics committee of the National Informatics and Liberty Committee (CNIL, no. 921290, 2021/07/19). This study was registered on clinicaltrials.gov under the number NCT05073328.

Results

Flowchart and cluster identification

In total, 90,025 patients were hospitalized with a main diagnosis of COVID-19 between February and June 2020 in France; 15,272 (17.0%) died during the hospital stay, and 2,346 (2.6%) within the first 30 days following discharge [17]. After excluding relevant patients, the final study population included 68,822 patients (Fig. 1): the median age was 64.8 years, 47% were women, and the median follow-up duration was 179.3 days (Table 1) [17].

The PCA identified 14 dimensions of HCU variables (Supplemental material S1). Following the K-means classification, 1,770 patients with outlying values of HCU were excluded (clustering criteria presented in Figure S2), resulting in a population of 67,052 patients available for the AHC method. This resulted in eight groups of homogeneous patients identified during the post-COVID-19 period, comprising between 1,163 and 35,501 patients (Table 1). Figure 2 shows the sociodemographic, clinical and index stay characteristics of patients by cluster (Fig. 2a) and for each category of HCU (visits: Fig. 2b; medications: Fig. 2c; and other HCU: Fig. 2d).

Patient age, sex distribution and percentage of deaths during the follow-up period were significantly different across the clusters; as were the duration of index stay, the proportion of patients who received oxygen therapy during the index stay, admitted to the ICU, SAPS-II and Charlson Comorbidity index scores (Table 1).

The two main clusters: cluster 1 and cluster 4

The main cluster (cluster 1) represented half of the study population ($n=35{,}501$, 52.9%). It included younger patients (median age: 62.0, Q1-Q3: 48,0–76,0), with few comorbidities and lower HCU than the other clusters. They mostly had short and noncomplex index stays (few oxygen support therapy requirements or ICU admissions). As illustrated in Fig. 2, for most HCU categories, few patients in cluster 1 consumed cares in post-COVID period compared to other clusters.

In terms of the number of patients, cluster 1 was followed by cluster 4, which included close to a quarter of the population ($n=15{,}4766$, 23.1%). Cluster 4 was older (median age: 75.0 years, Q1-Q3: 61,0–86), had a high comorbidity rate (mean Charlson Comorbidity Index: 5.0), and a mortality rate of 6.7% during the follow-up. A large proportion of cluster 4 patients used antidepressants (37.1%), anxiolytics (40.0%), and hypnotics (19.2%),

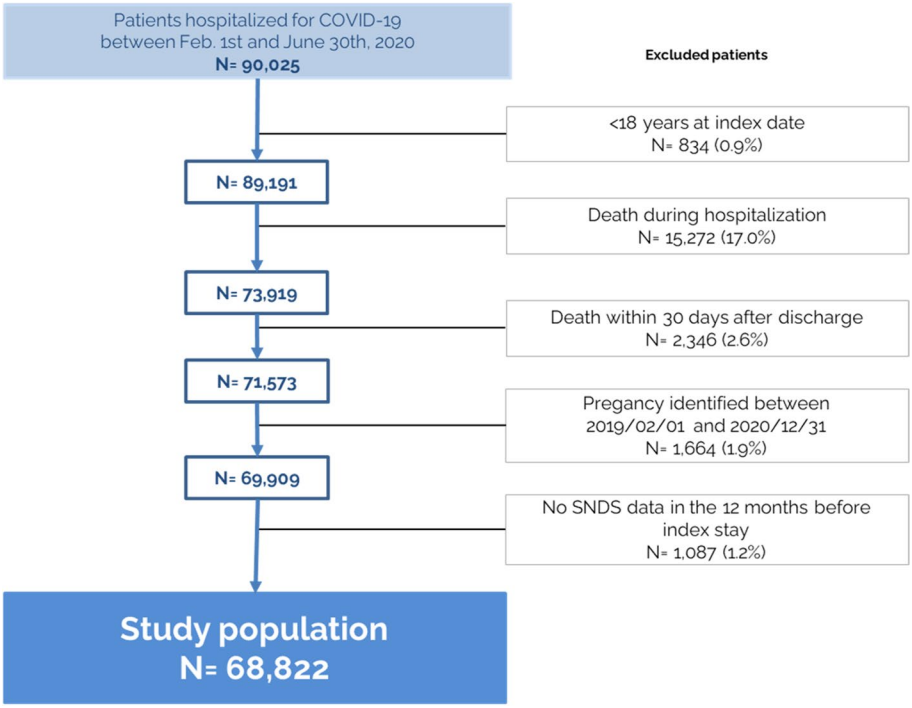

Fig. 1 Study population selection

in addition to analgesics (73.3%), betablockers (38.1%) and anticoagulants (35.1%) (Table 2 and Fig. 2). A high percentage visited hospital practitioners (HPs) (42.1%) and mental health physicians (4.2%). Finally, 6.1% of the patients received oxygen therapy after discharge (Table 2).

Other clusters

Cluster 2 ($n=3,431$) included 55% females with a very low mortality rate (0.6%). The median age was 63.0 years (Q1-Q3: 51,0–74,0). A high percentage of them used analgesiccs (71.6%) and nonsteroidal anti-inflammatory drugs (NSAIDs) during the post-COVID-19 period (40.5%) (Table 2 and Fig. 2) but in similar proportions to those observed during the pre-COVID-19 period *(data not shown)*; 30.1% visited dermatologists, 20.2% visited rheumatologists, and 34.6% underwent radiology (lung, hips, shoulders or knees).

In Cluster 5 ($n=3,221$), one-third of the patients had cardiac insufficiency/failure/chronic heart conditions (36.1%); accordingly, a high proportion of patients required visits to cardiologists (60.2%) and HPs (40.6%). The index stay was complex: 29% of the patients were admitted to the ICU, 13% of whom required invasive mechanical ventilation.

Cluster 6 ($n=3,214$) included 48.3% of patients with chronic respiratory diseases, and 22% visited pneumologist at least once (versus 4% over the pre-COVID-19 period). More than three-quarters (76.6%) used inhaled corticosteroids, 62.6% used short-acting bronchodilators, 42.4% had respiratory function exploration, 28.7% required an imaging evaluation, and 13% had oxygen therapy during the follow-up period.

Cluster 8 ($n=2,805$) included 28.5% of patients with obesity (median age: 60 years, Q1-Q3: 52,0–70,0); however, the prevalence of other comorbidities was low, as were HCU rates in the pre-COVID-19 period (*data not shown*). The index hospital stay was complex (median length: 17 days; 55.1% were admitted to the ICU, among whom 43.4% required invasive mechanical ventilation, mean SAPS-II score: 39.7), and 31.5% were transferred to rehabilitation care following the index hospitalization. Almost half of them (49.1%) underwent physiotherapist interventions during the post-COVID-19 period, and 6.5% were reimbursed for occupational therapy (Table 2 and Fig. 2). Finally, two-thirds (67.2%) were reimbursed for analgesics during the post-COVID-19 period.

Cluster 7 ($n=2,241$) was characterized by the severity of the index stay: 26.3% of patients were admitted to the ICU, among whom 15.4% required invasive mechanical ventilation; their median SAPS-II score was 35. Over the

Table 1 Number of patients and main characteristics of patients by cluster and in the total study population

	Total cohort	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7	Cluster 8	*p* value*	*Outliers*
No	68,822	35,501	3,431	1,163	15,476	3,221	3,214	2,241	2,805		*1,770*
Median age	64.8	62.0	63.0	75.0	75.0	67.0	70.0	68.0	60.0	< 0.0001	*66.0*
% of males	53.0	54.8	45.0	59.7	45.5	64.3	51.1	58.5	62.1	< 0.0001	*53.0*
% deaths over follow-up	4.5	4.1	0.6	17.6	6.6	0.5	2.3	0.7	1.1	< 0.0001	*16.0*
Median follow-up period (in days)	183.0	183.0	183.0	183.0	183.0	183.0	183.0	183.0	183.0		*183.0*
Median index stay duration (in days)	9.0	7.0	7.0	12.0	11.0	9.0	9.0	9.0	17.0	< 0.0001	*10.0*
Oxygen therapy during index stay (% of patients)	16.1	13.7	17.5	15.9	14.9	19.3	17.9	20.9	38.3	< 0.0001	*22.4*
Mean SAPS II score	33.8 (15.0)	30.7 (14.3)	32.5 (14.1)	39.6 (15.2)	35.3 (14.7)	34.7 (15.1)	31.6 (13.2)	37.6 (15.4)	39.7 (15.7)	< 0.0001	*37.9*
Mean Charlson Comorbidity Index (SD)	3.6 (3.2)	2.9 (2.9)	2.7 (2.4)	6.1 (3.7)	5.0 (3.4)	3.4 (2.5)	4.5 (3.0)	3.4 (2.8)	3.0 (2.7)	< 0.0001	*4.4*
% ICU admissions	20.3	16.6	22.2	21.8	18.1	29.0	19.6	26.3	55.1	< 0.0001	*29.8*
% obesity	15.3	12.6	16.7	19.1	15.9	18.2	20.1	15.8	18.5	< 0.0001	*22.4*

* chi-square tests and ANOVA tests result

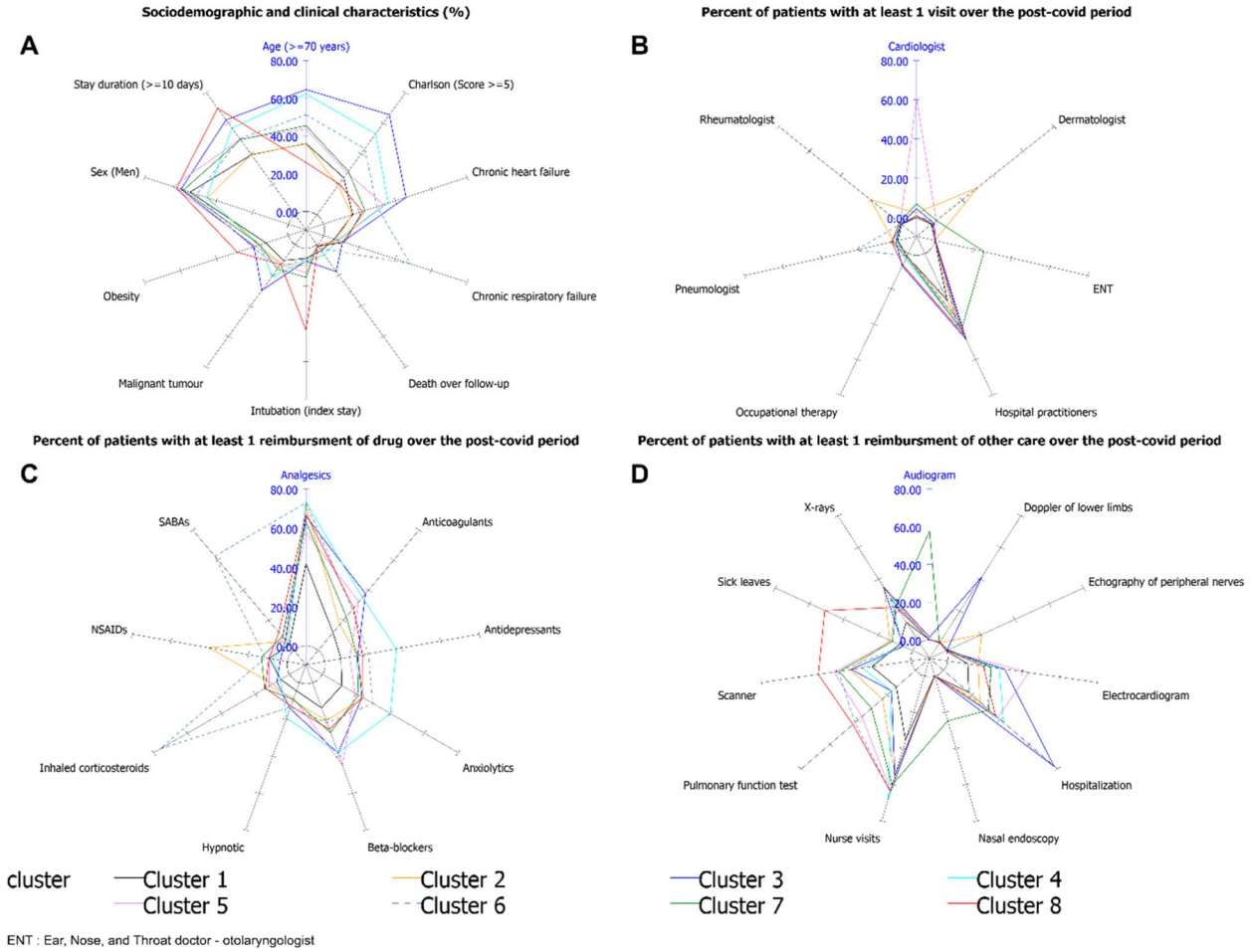

Fig. 2 Description of clusters 1 to 8 according to sociodemographic and clinical characteristics, visits, drug consumption, and other types of healthcare consumption

follow-up, one-quarter visited otolaryngologists (25,6%), were reimbursed for nasal endoscopy (24.8%), and more than half were reimbursed for audiograms (57.6%). However, they also had care related to their ears/noses/throats during the pre-COVID-19 period (e.g., 4.5% had nasal endoscopy, and 6.1% had audiogram; *data not shown*).

Cluster 3 ($n=1,163$) included elderly (median: 75.0 years, Q1-Q3: 64,0–83,0) patients, mostly men (59.7%). They had the highest death rate over the follow-up (17.6%) and the highest comorbidity rate (mean Charlson Comorbidity Index: 6.1), including 29.9% of patients with neoplasms. Consistently, 8.5% used strong opioids. The rate of rehospitalization after discharge was 77.7%, and 55.2% of the patients required regular nurse visits. After their index hospitalization, 18.2% of the patients had been hospitalized in rehabilitation care units, and 18.8% at home.

Discussion

This is the first study to identify distinct clusters of patients after hospital discharge for COVID-19 in the first months of the pandemic based on posthospitalization HCU. Eight clusters were identified with specific patient characteristics.

Cluster 1 included half of the patients, with no specific pattern of comorbidities or HCU after discharge; this may be interpreted as a type of HCU similar to that of the general population. Clusters 3, 4 and 5 were composed of older patients with major comorbidities: 29.9% had cancer, 28.0% mental disorders, and 33.6% cardiac insufficiency respectively. In cluster 6, nearly half of the patients had major respiratory insufficiency, and respiratory-related HCU was high during the post-COVID-19 period (higher than that in the pre-COVID-19 period), suggesting worsening previous respiratory diseases. Patients in cluster 7 had a very complex index stay with a particularly

Table 2 Number and proportion of patients with at least one reimbursement of care by cluster (post-COVID-19 period)

	Total cohort (*N* = 68,822)		Cluster 1 (*N* = 35,501)		Cluster 2 (*N* = 3,431)		Cluster 3 (*N* = 1,163)		Cluster 4 (*N* = 15,476)		Cluster 5 (*N* = 3,221)		Cluster 6 (*N* = 3,214)		Cluster 7 (*N* = 2,241)		Cluster 8 (*N* = 2,805)	
	No	%	No	%	No	%	No	%	No	%	No	%	No	%	No	%	No	%
Visits																		
General practitioner	51,014	74.1	22,580	63.6	2,985	87.0	777	66.8	13,121	84.8	2,781	86.3	2,808	87.4	1 967	87.8	2 509	89.5
Otolaryngologist	677	1.0	0	0.0	0	0.0	< 11	-	< 11	-	< 11	-	0	0.0	574	25.6	0	0.0
Mental health physician	1,507	2.2	394	1.1	80	2.3	< 11	-	643	4.2	48	1.5	41	1.3	47	2.1	60	2.1
Rehabilitation procedures (all)	5,805	8.4	757	2.1	224	6.5	131	11.3	839	5.4	1,516	47.1	102	3.2	288	12.9	1 488	53.1
Physiotherapist	2,841	4.1	253	0.7	142	4.1	20	1.7	362	2.3	203	6.3	57	1.8	179	8.0	1 378	49.1
Outpatient oxygen therapy	2,419	3.5	474	1.3	66	1.9	45	3.9	946	6.1	96	3.0	422	13.1	64	2.9	114	4.1
Analgesics	38,407	55.8	14,845	41.8	2,458	71.6	768	66.0	11,347	73.3	1,867	58.0	2,344	72.9	1 418	63.3	1 885	67.2
NSAIDs	7,412	10.8	3,273	9.2	1,390	40.5	41	3.5	1,103	7.1	303	9.4	363	11.3	300	13.4	265	9.5
Antidepressants	12,376	18.0	3,021	8.5	583	17.0	199	17.1	5,746	37.1	422	13.1	728	22.7	376	16.8	547	19.5
Anxiolytics	14,916	21.7	4,064	11.5	839	24.5	272	23.4	6,188	40.0	616	19.1	955	29.7	476	21.2	658	23.5
Hypnotics	6,759	9.8	1,551	4.4	335	9.8	162	13.9	2,965	19.2	268	8.3	424	13.2	203	9.1	355	12.7
Nurse visits	33,946	49.3	12,507	35.2	1,883	54.9	642	55.2	10,492	67.8	2,043	63.4	2,045	63.6	1 348	60.2	1 788	63.7
Hospitalizations																		
SSR hospitalizations	3,712	5.4	1,234	3.5	94	2.7	212	18.2	1,119	7.2	164	5.1	136	4.2	124	5.5	374	13.3
At-home hospitalizations	645	0.9	101	0.3	< 11	-	219	18.8	203	1.3	< 11	-	< 11	-	< 11	-	15	0.5

NSAID Nonsteroidal anti-inflammatory drug, *SSR* Hospitalization in rehabilitation care units (Soins de Suite et de Réadaptation)

high post-COVID-19 HCU for ear-nose-throat-related care: that may be due to the need for voice or deglutition reeducation following intubation. Finally, patients in cluster 8 were younger than the overall study population and had long and complex index stays, a high proportion of patients who underwent rehabilitation care, and a marked increase in post-COVID HCU compared to pre-COVID-19 HCU. It was more complex to characterize patients in cluster 2: they had high pain-related and NSAID consumption both pre- and post-COVID-19 but no particular comorbidity profile.

As a rule, the impact of previous comorbidities should not be ignored in the interpretation of these results, as it explains a large part of the HCU observed over the post-COVID-19 period. Unfortunately, we could not differentiate nor measure the magnitude of COVID itself, comorbidities, and COVID-related decompensation of comorbidities. Overall, we observed that patients who were already highly comorbid before COVID-19 hospitalization (i.e., cluster 3 with a high rate of cancer) had a high mortality rate over the post-COVID-19 period, most likely due to their underlying comorbidities. We may assume that this high mortality is not fully attributable to COVID-19 and was not avoidable. Consequently, posthospitalization care for these patients should focus more on supportive care than on long-term or rehabilitation hospitalization. In clusters with high rates of chronic diseases other than neoplasms (e.g., cluster 5 or cluster 6), we observed an expected high HCU in the post-COVID-19 period. Reinforced monitoring should be planned for these patients during pandemic periods, as worsening of their underlying disease is predictable in the context of hospitalization for a viral infection. Finally, cluster 8, with a high percentage of obese patients (28.5%), presented a low mortality rate despite the severity of their index stay, but a marked increase in post-COVID HCU compared to pre-COVID-19 HCU, likely reflecting a severe infection. These results are in line with the findings of a recent US study showing that obesity is associated with higher levels of HCU in patients hospitalized for COVID-19 [25]. Based on data on influenza, the association between obesity and viral respiratory infection severity is indeed well established [26]. Consequently, clinicians tend to monitor early and closely obese patients in the context of hospitalization for viral infection to avoid death, which is mostly avoidable in this population (young, with few comorbidities). Finally, while our study did not specifically analyze HCU related to post-ICU care, it is important to note that the increased healthcare burden is well-recognized among ICU survivors, particularly following complex and prolonged stays [27].

Other studies have investigated HCU following COVID-19: a US study based on the Veterans Health Administration showed that not hospitalized patients surviving the first 30 days after COVID-19 diagnosis were more likely to use medication, have laboratory abnormalities and incident diagnoses than no COVID-19 patients [28]. Other studies based on claims data explored COVID-19 inpatients' outpatient or emergency room visits and readmissions at 30 days after discharge [15], or medications use and hospitalizations in unvaccinated patients diagnosed with COVID-19, depending on whether they were hospitalized [16]. These results are not directly comparable to ours because of the different study populations, follow-up times and outcomes. Another UK study estimated HCU of more than 1,5 millions of patients in the four weeks following COVID-19 diagnosis between 2020 and 2022: they found that HCU were greater in those at higher risk of severe COVID-19, particularly in older patients [29], supporting our results.

Other studies have used clustering analyses to explore the clinical consequences of COVID-19: Fisher et al. identified clinically relevant subgroups of patients using hierarchical clustering based on symptoms and quality of life indices [6, 30]. Similarly, Zhang et al. identified four subphenotypes of postacute symptoms of SARS-CoV-2 infection over a follow-up period similar to ours: i) cardiac/renal; ii) respiratory/sleep/anxiety; iii) musculoskeletal/nervous system; iv) digestive/respiratory system sequelae [11]. However, our study differs from those that investigated postacute syndrome conditions [30, 31], as no patient-reported symptom defining this is available in claims database[32, 33].

Our study was based on an exhaustive cohort of patients hospitalized for COVID-19 from the start of the pandemic in France to the end of the first wave. This very large study population provided high power for identifying patient profiles based on robust data on all relevant HCU following SARS-CoV-2. This is the first study using such a wide range of HCU, thus providing a large picture of postdischarge HCU among COVID-19 inpatients. The WARD clustering analysis performed on HCU over the 2 to 6 months following discharge was able to identify patient profiles consistent with clinical situations, as demonstrated by their comorbidity and HCU patterns. This is of significant interest, as there was no a priori assumption about the number of clusters. Finally, because the study was performed over a period over which no SARS-CoV-2 vaccine was available, the results are of particular interest for predicting what might occur in the overall population in the case of a future pandemic involving an unknown respiratory virus. Indeed, the study provide key elements to clinicians to adapt their management to the patients' profile and predict their

needs in terms of healthcare resources in the presence of an unknown respiratory pathogen.

However, our study included only patients who were hospitalized for SARS-CoV-2. Consequently, the results are representative only of patients with severe forms of the disease. In addition, we included in the clustering analysis only patients who survived until discharge and at least 30 days after. We can thus assume that we included the less fragile patients, which may explain the low death rates observed in some clusters. Moreover, this would be of interest to study HCU of the deceased patients, whose HCU should also be non-negligible. Here, the mean duration of the stay in the 15,272 patients who deceased during their index stay was 11.5 days (SD = 11.1) and 31.4% of them needed ICU. However, in the SNDS, detailed cares consumed during a hospitalization are not available: only the cost of the hospitalization itself is available. In addition, as a claims database, little clinical information is available in the SNDS: only diagnoses related to hospitalizations were available, but no medication-related diagnoses, symptoms or PCR test results. However, in early 200, PCR tests were rarely performed in the general population in France. The data cover the first wave of COVID-19 and the following months; thus, the profiles of hospitalized patients during subsequent waves may be very different from those hospitalized during the first wave [34], and the use of vaccines and variants that occurred later in the pandemic course impacted COVID-19 patient outcomes. A replication of the same analysis over a later period of the pandemic would thus be of interest. However, these early data remain highly useful for understanding the consequences of a new pandemic on HCU in its early phase and identifying patient that may have similar postdischarge profiles.

Conclusion

This study presents robust and little explored data on the healthcare consumption of patients who were hospitalized for COVID-19 and who survived at least 30 days following discharge. We distinguished eight clinically relevant profiles of patients based on their postdischarge HCU. These results provide valuable insight for outpatient monitoring during new viral respiratory pandemics. The results could be replicated over subsequent periods of the pandemic to account for different variants and vaccination. Our study contributes to a better understanding of the impact of a new contagious respiratory condition on HCU following the discharge of hospitalized patients, which is crucial for allocating needed resources in such an urgent situation.

Abbreviations

AHC	Ascendant Hierarchical Clustering
HCU	Healthcare Resource Use
HP	Hospital Practitioner
ICD	International Classification of Diseases
ICU	Intensive Care Unit
NSAID	Non Steroid Anti-Inflammatory Drugs
PCA	Principal Component Analysis
SAPS	Simplified Acute Physiology Score
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
SNDS	Système National des Données de Santé

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12913-025-12308-5.

Supplementary Material 1

Acknowledgements
The authors thank the French NHS (Caisse Nationale de l'Assurance Maladie) for providing access to its claims data. We also thank Dr. André Lecoannet for his valuable advice about the codes used to identify the study population.

Authors' contributions
Claire Marant Micallef: CMM; Manon Belhassen: MB; Florence Ader: FA; Valeria Martinez: VM; Eric Van Ganse: EVG; Marjorie Bérard: MBer; Mikhail Dziadzko: MD; Frédéric Aubrun: FAu; Mélanie Née: MN Concept and design: FAu, EVG, MD, MB Acquisition of data: MBer, CMM Analysis and interpretation of data: MB, CMM, EVG, MD, FAu, VM, FA Drafting of the manuscript: CMM Critical revision of the paper for important intellectual content: MD, MB, EVG, FA, VM, FAu Statistical analysis: MBer, MN Obtaining funding: MB, CMM, FAu, MD.

Funding
This study was funded by ANR Covid Résilience, Fondation de France and Apicil Fundation.

Data availability
The dataset analyzed (i.e. the SNDS database) during the current study is not publicly available due data privacy. Data access is only granted to trained personnel upon request specific for the project and after approval from the Health Data Hub scientific committee and CNIL.

Declarations

Ethics approval and consent to participate
This study was approved by the ethic committee of the French Health Data Hub (no. 4653731, 2021/07/08). It was conducted using anonymized data and approved by the ethics committee of the National Informatics and Liberty Committee (CNIL, no. 921290, 2021/07/19). This study was registered on clinicaltrials.gov under the number NCT05073328.
The data used were secondary data from the French healthcare claims database (SNDS). This is an anonymized database; thus we did not have access to patients to request their consent to participate. However, the persons concerned wishing to assert their rights should make a request to their health insurance to attachment (CPAM), in accordance with article 111 of the implementing decree of the Data Protection Act (https://www.snds.gouv.fr/SNDS/Protection-de-la- donnee).

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]PELyon, Pharmacoépidémiologie Lyon, 210 Avenue Jean Jaurès, Lyon 69007, France. [2]Infectious and Tropical Diseases Department, Hospices Civils de Lyon, Hôpital de La Croix-Rousse, 103 Gd Rue de La Croix-Rousse, Lyon 69004, France. [3]Université Claude Bernard Lyon 1, CIRI, INSERM U1111, CNRS UMR5308, ENS Lyon, Lyon F-69372, France. [4]Service d'Anésthésie Douleur,

Raymond Poincaré Hospital, Assistance Publique Hôpitaux de Paris, 104 Boulevard Raymond Poincaré, Garches 92380, France. [5]INSERM, U-987, Hôpital Ambroise Paré, Centre d'Évaluation Et de Traitement de La Douleur, 9 Av. Charles de Gaulle, 92100, Boulogne Billancourt F-92100, France. [6]Université Versailles Saint-Quentin, 55 Avenue de Paris, Versailles 78035, France. [7]Hospices Civils de Lyon, Croix Rousse University Hospital, Respiratory Medicine, 103 Gd Rue de La Croix-Rousse, Lyon 69004, France. [8]Laboratoire RESHAPE, Université Claude Bernard Lyon 1, INSERM UMR 1290, 8 Av. Rockefeller, Lyon 69008, France. [9]Département d'Anesthésie-Réanimation, Hospices Civils de Lyon, Hôpital de La Croix Rousse, Douleur, 103 Gd Rue de La Croix-Rousse, Lyon 69004, France.

Received: 21 May 2024 Accepted: 20 January 2025
Published online: 27 January 2025

References

1. ATIH. Analyse de l'activité hospitalière 2020. ATIH; 2021.
2. Yang J, Zheng Y, Gou X, Pu K, Chen Z, Guo Q, et al. Prevalence of comorbidities and its effects in patients infected with SARS-CoV-2: a systematic review and meta-analysis. Int J Infect Dis. 2020;94:91–5.
3. EPI-PHARE. Maladies chroniques, états de santé et risque d'hospitalisation et de décès hospitalier pour COVID-19 lors de la première vague de l'épidémie en France: Étude de cohorte de 66 millions de personnes. GIS ANSM-CNAM; 2021 9 février 2021.
4. Machine learning identifies long COVID patterns from electronic health records. Nat Med. 2023;29(1):47–8.
5. Al-Aly Z, Xie Y, Bowe B. High-dimensional characterization of post-acute sequelae of COVID-19. Nature. 2021;594(7862):259–64.
6. Fischer A, Badier N, Zhang L, Elbeji A, Wilmes P, Oustric P, et al. Long COVID Classification: Findings from a Clustering Analysis in the Predi-COVID Cohort Study. Int J Environ Res Public Health. 2022;19(23):16018.
7. Ayoubkhani D, Khunti K, Nafilyan V, Maddox T, Humberstone B, Diamond I, et al. Post-covid syndrome in individuals admitted to hospital with covid-19: retrospective cohort study. BMJ (Clinical research ed). 2021;372:n693.
8. Carfì A, Bernabei R, Landi F. Gemelli Against C-P-ACSG. Persistent Symptoms in Patients After Acute COVID-19. Jama. 2020;324(6):603–5.
9. Lopez-Leon S, Wegman-Ostrosky T, Perelman C, Sepulveda R, Rebolledo PA, Cuapio A, et al. More than 50 long-term effects of COVID-19: a systematic review and meta-analysis. Sci Rep. 2021;11(1):16144.
10. Perrin R, Riste L, Hann M, Walther A, Mukherjee A, Heald A. Into the looking glass: Post-viral syndrome post COVID-19. Med Hypotheses. 2020;144:110055.
11. Zhang H, Zang C, Xu Z, Zhang Y, Xu J, Bian J, et al. Data-driven identification of post-acute SARS-CoV-2 infection subphenotypes. Nat Med. 2023;29(1):226–35.
12. George PM, Barratt SL, Condliffe R, Desai SR, Devaraj A, Forrest I, et al. Respiratory follow-up of patients with COVID-19 pneumonia. Thorax. 2020;75(11):1009–16.
13. Mizrahi B, Sudry T, Flaks-Manov N, Yehezkelli Y, Kalkstein N, Akiva P, et al. Long covid outcomes at one year after mild SARS-CoV-2 infection: nationwide cohort study. BMJ (Clinical research ed). 2023;380:e072529.
14. Davis HE, McCorkell L, Vogel JM, Topol EJ. Long COVID: major findings, mechanisms and recommendations. Nature Rev Microbiol. 2023;21(3):133–46.
15. Moon RC, Brown H, Rosenthal N. Healthcare Resource Utilization of Patients With COVID-19 Visiting US Hospitals. Value Health. 2022;25(5):751–60.
16. Scott A, Chambers R, Reimbaeva M, Atwell J, Baillon-Plot N, Draica F, et al. Real-world retrospective analysis of patient characteristics, healthcare resource utilization, costs, and treatment patterns among unvaccinated adults with COVID-19 diagnosed in outpatient settings in the United States. J Med Econ. 2022;25(1):287–98.
17. Dziadzko M, Belhassen M, Van Ganse E, Heritier F, Berard M, Marant-Micallef C, et al. Health Care Resource Use and Total Mortality After Hospital Admission for Severe COVID-19 Infections During the Initial Pandemic Wave in France: Descriptive Study. JMIR Public Health Surveill. 2024;10:e56398.
18. Tuppin P, Rudant J, Constantinou P, Gastaldi-Menager C, Rachas A, de Roquefeuil L, et al. Value of a national administrative database to guide public decisions: From the systeme national d'information interregimes de l'Assurance Maladie (SNIIRAM) to the systeme national des donnees de sante (SNDS) in France. Rev Epidemiol Sante Publique. 2017;65(Suppl 4):S149–67.
19. Le Gall JR, Lemeshow S, Saulnier F. A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. JAMA. 1993;270(24):2957–63.
20. Bannay A, Chaignot C, Blotiere PO, Basson M, Weill A, Ricordeau P, et al. The Best Use of the Charlson Comorbidity Index With Electronic Health Care Database to Predict Mortality. Med Care. 2016;54(2):188–94.
21. Institute S. The FASTCLUS Procedure: example outliers2023. Available from: https://documentation.sas.com/doc/en/pgmsascdc/9.4_3.4/statug/statug_fastclus_examples02.htm.
22. MacQueen JB. Some Methods for Classification and Analysis of Multivariate Observations. Proc Fifth Berkeley Symp Math Stat Probability. 1967;1:281–197.
23. Ward JH. Hierarchical Grouping to Optimize an Objective Function. J Am Stat Assoc. 1963;58:236–44.
24. Sarle WS. Cubic Clustering Criterion. Technical Report A-1081983. Available from: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://support.sas.com/documentation/onlinedoc/v82/techreport_a108.pdf.
25. Kim N, Aly A, Craver C, Garvey WT. Burden of illness associated with overweight and obesity in patients hospitalized with COVID-19 in the United States: analysis of the premier healthcare database from April 1, 2020 to October 31, 2020. Journal of medical economics. 2023:1-29.
26. Rafia A, Afak YS, Saeed A, Shuja M, Sameer M. Why Are Obese People Predisposed to Severe Disease in Viral Respiratory Infections? Obesities. 2023;3(1).
27. Hirshberg EL, Wilson EL, Stanfield V, Kuttler KG, Majercik S, Beesley SJ, et al. Impact of Critical Illness on Resource Utilization: A Comparison of Use in the Year Before and After ICU Admission. Crit Care Med. 2019;47(11):1497–504.
28. Al-Aly Z, Xie Y, Bowe B. High-dimensional characterization of post-acute sequelae of COVID-19. Nature. 2021;594(7862):259–64.
29. Yang J, Andersen KM, Rai KK, Tritton T, Mugwagwa T, Reimbaeva M, et al. Healthcare resource utilisation and costs of hospitalisation and primary care among adults with COVID-19 in England: a population-based cohort study. BMJ Open. 2023;13(12):e075495.
30. Wulf Hanson S, Abbafati C, Aerts JG, Al-Aly Z, Ashbaugh C, Ballouz T, et al. Estimated Global Proportions of Individuals With Persistent Fatigue, Cognitive, and Respiratory Symptom Clusters Following Symptomatic COVID-19 in 2020 and 2021. JAMA. 2022;328(16):1604–15.
31. O'Mahoney LL, Routen A, Gillies C, Ekezie W, Welford A, Zhang A, et al. The prevalence and long-term health effects of Long Covid among hospitalised and non-hospitalised populations: A systematic review and meta-analysis. EClinicalMedicine. 2023;55:101762.
32. World Health Organization. Clinical case definition of post COVID-19 condition by a Delphi consensus, 6 October 2021 2021 [Available from: https://www.who.int/publications/i/item/WHO-2019-nCoV-Post_COVID-19_condition-Clinical_case_definition-2021.1.
33. Thaweethai T, Jolley SE, Karlson EW, Levitan EB, Levy B, McComsey GA, et al. Development of a Definition of Postacute Sequelae of SARS-CoV-2 Infection. JAMA. 2023;329(22):1934–46.
34. Stepanova M, Lam B, Younossi E, Felix S, Ziayee M, Price J, et al. The impact of variants and vaccination on the mortality and resource utilization of hospitalized patients with COVID-19. BMC Infect Dis. 2022;22(1):702.

Publisher's Note