



## **COVID-19 profiles in general practice: a latent class analysis**

Emilie Ferrat, William Mirat, Emmanuelle Boutin, Emilie Maroto, Sophie Brossier, Jean-Denis Hoonakker, Etienne Audureau, Tan-Trung Phan, Sylvie Bastuji-Garin

### **► To cite this version:**

Emilie Ferrat, William Mirat, Emmanuelle Boutin, Emilie Maroto, Sophie Brossier, et al.. COVID-19 profiles in general practice: a latent class analysis. *BMJ Open*, 2024, 14 (6), pp.e080393. <10.1136/bmjopen-2023-080393>. <hal-04604521>

**HAL Id: hal-04604521**

**<https://hal.science/hal-04604521v1>**

Submitted on 7 Jun 2024



**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

# BMJ Open COVID-19 profiles in general practice: a latent class analysis

Emilie Ferrat,<sup>1,2,3</sup> William Mirat ,<sup>2,4</sup> Emmanuelle Boutin,<sup>1,5</sup> Emilie Maroto,<sup>2</sup> Sophie Brossier,<sup>2,6</sup> Jean-Denis Hoonakker,<sup>2,7</sup> Etienne Audureau,<sup>1,5,8</sup> Tan-Trung Phan ,<sup>2,6</sup> Sylvie Bastuji-Garin<sup>1,8</sup>

**To cite:** Ferrat E, Mirat W, Boutin E, *et al.* COVID-19 profiles in general practice: a latent class analysis. *BMJ Open* 2024;**14**:e080393. doi:10.1136/bmjopen-2023-080393

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (<https://doi.org/10.1136/bmjopen-2023-080393>).

Received 29 September 2023  
Accepted 30 April 2024

## ABSTRACT

**Background** General practitioners (GPs) were on the front line of the COVID-19 outbreak. Identifying clinical profiles in COVID-19 might improve patient care and enable closer monitoring of at-risk profiles.

**Objectives** To identify COVID-19 profiles in a population of adult primary care patients, and to determine whether the profiles were associated with negative outcomes and persistent symptoms.

**Design, setting and participants** In a prospective multicentre study, 44 GPs from multiprofessional primary care practices in the Paris area of France recruited 340 consecutive adult patients (median age: 47 years) with a confirmed diagnosis of COVID-19 during the first two waves of the epidemic.

**Method and outcome** A latent class (LC) analysis with 11 indicators (clinical signs and symptoms) was performed. The resulting profiles were characterised by a 3-month composite outcome (COVID-19-related hospital admission and/or death) and persistent symptoms three and 6 months after inclusion.

**Results** We identified six profiles: 'paucisymptomatic' (LC1, 9%), 'anosmia and/or ageusia' (LC2, 12.9%), 'influenza-like syndrome with anosmia and ageusia' (LC3, 15.5%), 'influenza-like syndrome without anosmia or ageusia' (LC4, 24.5%), 'influenza-like syndrome with respiratory impairment' (LC5) and a 'complete form' (LC6, 17.7%). At 3 months, 7.4% of the patients were hospitalised (with higher rates in LC5), and 18% had persistent symptoms (with higher rates in LC5 and LC6). At 6 months, 6.4% of the patients had persistent symptoms, with no differences between LCs.

**Conclusion** Our findings might help GPs to identify patients at risk of persistent COVID-19 symptoms and hospital admission and then set up procedures for closer monitoring.

## STRENGTHS AND LIMITATIONS OF THIS STUDY

- ⇒ The major limitations of our study relate to the population size and the limited number of events used to validate our classification.
- ⇒ The one-step clustering method used here is known to minimise classification errors.
- ⇒ The use of a bootstrap method and several parsimony indices increased the reliability of our six-class solution.
- ⇒ Our study period did not encompass the presence of all the COVID-19 variants, which might have affected the prevalences of the latent classes and the outcomes.
- ⇒ The study was performed before COVID-19 vaccines became available; the prevalence of some latent classes and persistent symptoms might, therefore, have been overestimated.

care and manage complications as early as possible. Despite the scale of the COVID-19 epidemic, we are aware of only one study in which COVID-19 symptoms were self-reported via a smartphone application with a view to identifying patient profiles and the associated need for respiratory support.<sup>2</sup> Hence, there are no literature data on COVID-19 profiles among patients consulting a GP and the corresponding associations with disease progression. Latent class analysis (LCA) is a patient-centred approach specifically designed to reliably identify subgroups of patients when they exist. LCA has been used successfully to investigate, characterise and validate disease subtypes, stratify patients into risk groups and predict treatment responses.<sup>3</sup>

We, therefore, sought to identify COVID-19 profiles (based on combinations of initial clinical symptoms and/or signs) in a population of adult primary care patients, using a hypothesis-free LCA. We then determined whether or not these profiles were associated with negative outcomes (COVID-19-related hospital admissions and deaths) at 3 months and persistent symptoms at 3 and 6 months.

## INTRODUCTION

COVID-19 emerged in China in late 2019. In France, around 38.5 million people have been infected to date, and around 161 000 of these died.<sup>1</sup> General practitioners (GPs) were on the front line of the COVID-19 epidemic; they managed the less severe cases and referred the more serious ones to a hospital. Identifying patient profiles in COVID-19 and especially those at a greater risk of a poor outcome might help to improve initial

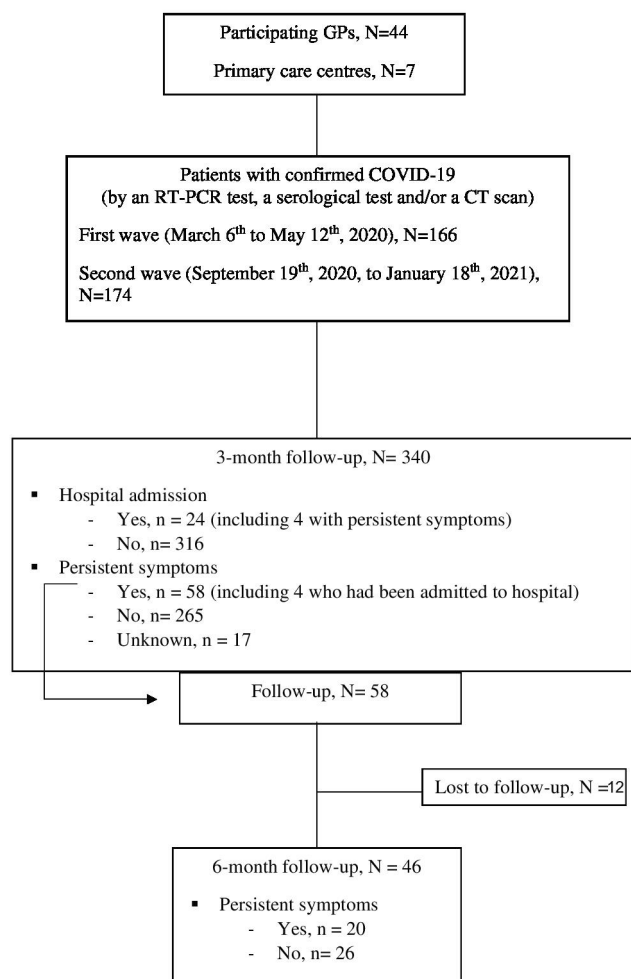


© Author(s) (or their employer(s)) 2024. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

## Correspondence to

Dr William Mirat;  
[william.mirat@gmail.com](mailto:william.mirat@gmail.com)



**Figure 1** Study flow chart. GP, general practitioner.

## METHOD

### Setting, design and participants

From 6 March 2020 to 12 May 2020 and from 19 September 2020 to 18 January 2021 (ie, during the first two waves of COVID-19 in France), we conducted a multicentre, prospective study in four counties in the Paris area (France): Val-de-Marne, Seine-et-Marne, Essonne and Seine-Saint-Denis. 44 GPs were recruited from multiprofessional primary care practices affiliated with the Faculty of Health at Université Paris-Est Créteil (France). During the first wave, the GPs recruited consecutive adult patients consulting for suspected COVID-19.<sup>4</sup> During the second wave, only patients with a confirmed diagnosis of COVID-19 (ie, a positive RT-PCR test and/or a positive serology test, and/or a chest CT result suggestive of COVID-19, according to the French national guidelines)<sup>5–7</sup> were included. The exclusion criteria were age under 18 and residence in an institution. In the present analysis, we considered all patients with a confirmed diagnosis (figure 1). Patients were followed up for 3 months, and those with persistent symptoms at 3 months were followed up for 6 months.

### Data collection

Data were extracted from the GPs' electronic medical records in November 2021. The extracted variables

included demographic characteristics, comorbidities and the symptoms and signs of COVID-19 documented for the LCA (documented fever, chills, body aches, cough, sputum, respiratory discomfort, dyspnoea (on effort or at rest), chest pain, rhinorrhoea, odynophagia, ageusia, anosmia, headache, abdominal pain, diarrhoea, nausea, vomiting, asthenia, poor general condition, lung auscultation findings, blood pressure and heart rate).

Patients were followed up as usual by their GP, and all other consultations with healthcare professionals were recorded. Three months after inclusion, the GP phoned or visited the patient and collected data on persistent COVID-19 symptoms, related deaths and hospital admissions. Persistent COVID-19 symptoms were identified according to the GP's usual clinical practice. We asked the GPs three questions: 'Do you consider that the patient has recovered from COVID-19? If not, which symptoms persisted? Do you attribute these symptoms to the initial disease?' Persistent symptoms (if any) were not rated on a scale or using a questionnaire. Patients with persistent symptoms at 3 months were contacted again at 6 months, and the same variables were recorded.

### Illustrative variables and outcomes

To characterise the COVID-19 profiles identified in an LCA, we considered the comorbidities at baseline. To investigate the prognostic value of these profiles, we considered the following two outcomes: (1) a 3-month composite outcome that included COVID-19-related hospital admissions and deaths (the relatedness to COVID-19 was judged by consulting the hospital's records) and (2) the persistence of COVID-19-related signs or symptoms 3 months and 6 months after inclusion. Lastly, we noted about whether a patient had been referred to hospital by the GP in the month following the first consultation.

### Statistical analysis

Quantitative and qualitative variables were described, respectively, as the median (IQR) and the number (%). The prevalence of persistent symptoms was calculated at 3 months (as a proportion of the whole study population) and at 6 months (for patients who had symptoms at three and 6 months, as a proportion of the whole study population less those lost to follow-up).

### Indicators used to determine COVID-19 health profiles

We first considered all COVID-19 signs and symptoms and the lung auscultation findings as indicators. Given that some indicators were highly correlated, the investigators (EF, SB-G, EA and EM) selected the indicators that they considered to be most relevant. Very poorly documented variables (such as tachycardia and blood pressure) were not considered, and highly correlated variables were grouped together in relevant health domains; for example, abdominal pain, diarrhoea and nausea and/or vomiting were grouped together.

The investigators reached a consensus on 11 indicators, which were then used in the LCA (table 1). To characterise the LCs and predict class membership, we considered the following three active covariates: age, sex and the presence of at least one comorbidity.

### Latent class analysis

Based on the selected indicators, we iteratively fitted models comprising up to eight LCs. We then selected the model with the best statistical properties by using a variety of parsimony indices (table 2), since no single approach is universally accepted.<sup>8 9</sup> Our one-step approach involved the simultaneous estimation of (1) the LC model of interest and (2) a logistic regression model in which the LCs were related to the active covariates listed above. The LCA was performed using Latent Gold software (V.5.0, Statistical Innovations, Belmont, Massachusetts, USA). We performed a sensitivity analysis with imputation for missing lung auscultation data, using Latent Gold's multiple imputation procedure. In a second sensitivity analysis, the LC data were stratified on the wave of COVID-19.

### Characterisation of the patient profiles

To characterise the identified profiles, we used posterior probabilities to assign patients to their most likely LC. The prevalence of comorbidities and outcomes was compared across LCs using the  $\chi^2$ , Fisher's exact test or the Kruskal-Wallis test, as appropriate. Post hoc pairwise comparisons were performed when the p value was  $\leq 0.05$ .

All tests were two sided, and the threshold for statistical significance was set to  $p \leq 0.05$ . P values from multiple pairwise comparisons were corrected using the false discovery rate method.<sup>10</sup> Statistical analyses were performed with Stata software (version 14.2, StataCorp, College Station, TX).

### Patient and public involvement

All patients received a study information sheet and gave their verbal consent to participation. Patients were not involved in the study design, conduct or reporting or plans for dissemination.

## RESULTS

### Study population

During the study period, 340 COVID-19 patients were included (figure 1). The median (IQR) age was 47 years (35–57), 202 were female (59.4%) and 163 (47.9%) had at least one comorbidity (table 1). Of the 340 patients, 24 (7.4%) were hospitalised in the first 3 months of follow-up, and 58 (out of 323 with data, 18%) still had persistent symptoms at 3 months. The most frequent symptoms were asthenia (6.8%), anosmia (5.6%) and dyspnoea (5%). At 6 months,

20 (6.4%) of the 311 patients still had persistent symptoms.

### Determination of COVID-19 profiles

A six-class solution showed the best fit with a non-significant bootstrap p value, the lowest sample size-adjusted Bayesian criterion, a significant improvement in fit compared with the five-class solution, and no improvement in fit for a seven-class solution, using the bootstrapped likelihood ratio test (table 2). The classification quality was good (entropy=0.83).

The conditional probabilities of indicators and covariates for each of the six LCs (ie, the probability of each indicator being present in class members) are summarised in table 1. Of the 340 patients, 30 (9.0%) belonged to LC1, 43 (12.9%) in LC2, 52 (15.5%) in LC3, 91 (24.5%) in LC4, 65 (20.4%) in LC5 and 59 (17.7%) in LC6. We named the LCs as follows: LC1='paucisymptomatic'; LC2='anosmia and/or ageusia'; LC3='influenza-like syndrome with anosmia and ageusia'; LC4='influenza-like syndrome without anosmia or ageusia'; LC5='influenza-like syndrome with respiratory impairment' and LC6='complete form'. LC1 and LC2 were also characterised by the lowest probabilities of asthenia and poor general condition (0.12 and 0.01, respectively), and LC5 was characterised by the highest probability of abnormal lung auscultation findings (0.37). Among the active covariates, age and presence of at least one comorbidity were significantly associated with class membership ( $p=0.017$  and  $p=0.04$ , respectively). More than half of the members of LC1, LC5 and LC6 had at least one comorbidity. The members of LC1, LC4, LC5 and LC6 were significantly older (with a median age ranging from 49 (39–59) to 51 (40–61)) than members of LC2 and LC3 (median age 40 (28–49) and 37.5 (32–50), respectively) (tables 3 and 4). The sensitivity analysis with imputation of missing data for lung auscultation ( $n=91$ ) gave similar results (online supplemental tables S1 and S2). After stratification on the wave of COVID-19, all LCs were present. LC4, LC5 and LC6 were the most prevalent in wave 1, whereas LC4 was the most prevalent in wave 2 (online supplemental table S3).

### Characteristics of the clinical profiles

The prevalence of hypertension (but not that of other comorbidities) significantly differed from one LC to another ( $p=0.006$ ); the patients in LC1, LC4 and LC5 were more likely to have hypertension than those in LC2 and LC3. At the 3-month follow-up time point, none of the patients had died but 24 had been admitted to hospital. The LCs differed with regard to the frequency of hospital admission ( $p=0.016$ ), with a higher rate in LC5 than in LC2 and LC6. Two-thirds of the referrals to hospital by a GP occurred in the month after the first consultation (tables 3 and 4).



**Table 1** Indicators estimates for the six-class solution among the 340 COVID-19 patients

Frequency of indicators (%)	LC1 (n=30)	LC2 (n=43)	LC3 (n=52)	LC4 (n=91)	LC5 (n=65)	LC6 (n=59)	P value
	‘Paucisymptomatic’	‘Anosmia and/or ageusia’	‘Influenza-like syndrome with anosmia and ageusia’	‘Influenza-like syndrome without anosmia or ageusia’	‘Influenza-like syndrome with respiratory impairment’	‘Complete form’	
Cluster size (%)	9.0	12.9	15.5	24.5	20.4	17.7	
Conditional probabilities							
Indicators: clinical signs and symptoms							
Documented fever and/or chills and/or body aches	0.133	0.441	<b>0.769</b>	<b>0.855</b>	<b>0.9</b>	<b>0.936</b>	<0.001
Cough and/or sputum	0.408	0.702	0.63	0.594	<b>0.836</b>	<b>0.999</b>	0.004
Respiratory discomfort and/or dyspnoea (on effort or at rest)	0.004	0.265	0.343	0.004	<b>0.893</b>	<b>0.815</b>	<0.001
Chest pain	0.156	0.128	0.343	0.001	<b>0.397</b>	<b>0.522</b>	<0.001
Abnormal lung auscultation*	0.01	0.066	0.107	0.082	<b>0.372</b>	0.168	0.046
Rhinorrhoea and/or odynophagia	0.269	0.587	0.488	0.565	<b>0.594</b>	<b>0.716</b>	0.001
Ageusia	0.003	<b>0.643</b>	<b>0.993</b>	0.108	0.005	<b>0.921</b>	<0.001
Anosmia	0.004	<b>0.851</b>	<b>0.996</b>	0.028	0.07	<b>0.851</b>	<0.001
Headache	0.006	0.222	<b>0.543</b>	0.493	0.498	<b>0.861</b>	
Abdominal pain, diarrhoea, nausea and/or vomiting	0.002	0.192	0.224	0.368	<b>0.42</b>	<b>0.531</b>	0.026
Asthenia and/or poor general condition	0.126	0.01	<b>0.926</b>	<b>0.724</b>	<b>0.854</b>	<b>0.771</b>	0.001
Active covariables							
Age (years)							
18–49	0.433	0.771	0.699	0.451	0.44	0.56	0.017
50–64	0.476	0.206	0.249	0.37	0.356	0.344	
65–74	0.001	0.022	0.062	0.098	0.112	0.063	
≥75	0.09	0.001	0	0.082	0.092	0.033	
Female sex	0.578	0.582	0.582	0.531	0.572	0.766	0.40
Presence of at least one comorbidity	<b>0.553</b>	0.419	0.419	0.434	<b>0.675</b>	0.503	0.040
Bolded values refer to the higher probabilities that allowed to characterise LCs.							
*Missing data on chest auscultation (n=91).							
LC, latent class.							

**Table 2** Goodness-of-fit indices for latent class models comprising one to eight classes (n=340)

Number of classes	P value* L2	BIC (LL)	SABIC (LL)	AIC3 (LL)	BLRT p value	Entropy
1	0.24	4683	4648	4652	--	1.00
2	0.30	4515	4432	4441	<0.001	0.85
3	0.41	4471	4341	4355	<0.001	0.78
4	0.33	4490	4313	4332	<0.001	0.76
5	0.24	4518	4292	4317	<0.001	0.83
6†	0.17	4564	4291	4320	<0.001	0.83
7	0.11	4620	4299	4334	0.14	0.82
8	0.09	4677	4309	4349	0.13	0.83

\*Obtained by bootstrapping (n=500).

†Model selected with the largest number of matching goodness-of-fit indices, including the BLRT. The model's classification error was p=0.10.

AIC3, Akaike information criterion; BIC, Bayesian information criterion; BLRT, bootstrapped likelihood ratio test; L2, likelihood ratio; LL, log likelihood; SABIC, sample size-adjusted Bayesian criterion.

The LCs also differed with regard to the prevalence of persistent symptoms (p=0.002). Patients in LC5 and LC6 were more likely to have persistent symptoms at 3 months than those in LC1, LC2 and LC4. The most prevalent persistent symptoms were asthenia and anosmia in LC6 and asthenia and dyspnoea in LC5. At 6 months, there were no significant differences in the prevalence of persistent symptoms between the LCs (p=0.096).

## DISCUSSION

### Summary

Our data-driven approach of a population COVID-19 patients managed by GPs identified six profiles, namely, 'paucisymptomatic' (LC1, 9% of the participants), 'anosmia and/or ageusia' (LC2, 12.9%), 'influenza-like syndrome with anosmia and ageusia' (LC3, 15.5%), 'influenza-like syndrome without anosmia or ageusia' (LC4, 24.5%), 'influenza-like syndrome with respiratory impairment' (LC5) and a 'complete form' (LC6, 17.7%). Age and the presence of at least one comorbidity were associated with class membership. At 3 months, 7.4% of the patients had been admitted to hospital (with a higher incidence in LC5 than in LC2 and LC6), and 18% had persistent symptoms. Persistent symptoms at 3 months were more prevalent in LC5 and LC6 than in LC1, LC2 and LC4. At 6 months, 20 patients (6.4%) still had persistent symptoms; the LCs did not differ significantly in this respect.

### Strengths and limitations

To the best of our knowledge, this study is the first to have defined clusters of COVID-19 patients followed up in general practice. The one-step clustering method used here is known to minimise classification errors, and the use of a bootstrap method and several parsimony indices have increased the reliability of our six-class solution.<sup>9</sup> The major limitation of our study relates to the population

size and the limited number of events used to validate our classification. Second, our study period did not encompass the presence of all the COVID-19 variants (such as Omicron, found in subsequent waves). Although the COVID-19 variants had different prevalences,<sup>11</sup> the signs and symptoms of disease were always the same; hence, the presence of other COVID-19 variants is unlikely to have affected the nature and number of LCs obtained in our analysis. After stratification on the wave of COVID-19, only the frequency of the LCs differed. Although vaccination may prevent long-term symptoms,<sup>12</sup> the literature data on the nature and frequency of long-term symptoms associated with the different variants are contradictory.<sup>13–15</sup> Long-term symptoms have also been reported in people carrying the Omicron variant. Third, the hospitalisation dates were not available. However, we did have information on the date of referral to the hospital by the patient's GP. Lastly, our study was limited to the greater Paris region and so might not be representative of the French population as a whole.

### Comparison with the literature

In line with a previous longitudinal study in which symptoms were self-reported via a smartphone application,<sup>2</sup> we identified six clinical profiles. However, the two studies differed substantially with regard to the clinical profiles. These disparities might be due to differences in methodology (consultation with a GP vs self-reporting of symptoms via an application; an LCA vs unsupervised 5-day time series clustering and the indicators used), and especially in the characteristics of the study population (patients consulting their GP vs patients able and willing to record their symptoms at least three times over 4 days or more). Sudre *et al* reported on more severe cases because about 20% of their study population had a hospital visit and 6% needed respiratory support (vs respectively, 7.4% and 0.9% in our study). In summary, Sudre *et al* identified

**Table 3** Global comparison of illustrative variables and 3-month and 6-month outcomes in the six-latent classes (LCs) identified (using posterior probabilities classifications; n=340)

	LC1 (n=30)	LC2 (n=43)	LC3 (n=52)	LC4 (n=91)	LC5 (n=65)	LC6 (n=59)
	‘Paucisymptomatic’ ‘Anosmia and/or ageusia’					
	‘Influenza-like syndrome with anosmia and ageusia’			‘Influenza-like syndrome with respiratory impairment’ ‘Complete form’		
Age, years, median (IQR)	51 (40–57)	40 (28–49)	37.5 (32–50)	51 (37–60)	51 (40–61)	49 (39–59)
<0.001						
Cross-sectional analysis						
Comorbidity						
Cardiac and cardiovascular comorbidities (n=340)	11 (36.7)	4 (9.3)	4 (7.7)	25 (27.5)	22 (33.9)	12 (20.3)
0.001						
Heart diseases	4 (13.3)	1 (2.3)	3 (5.8)	7 (7.7)	8 (12.3)	3 (5.1)
0.330						
Vascular disease	2 (6.7)	0 (0)	0 (0)	2 (2.2)	3 (4.6)	1 (1.7)
0.280						
Hypertension	9 (30.0)	3 (7.0)	3 (5.8)	22 (24.2)	16 (24.6)	9 (15.3)
0.006						
Diabetes (n=340)	4 (13.3)	1 (2.3)	2 (3.9)	5 (5.5)	4 (6.2)	2 (3.4)
0.461						
Cancer and/or immunodeficiency (n=340)	1 (3.3)	1 (2.3)	1 (1.9)	5 (5.5)	3 (4.6)	4 (6.8)
0.861						
Respiratory disease (n=340)	1 (3.3)	5 (11.6)	3 (5.8)	8 (8.8)	10 (15.4)	6 (10.2)
0.463						
Obesity (BMI>30) (n=340)	6 (20.0)	4 (9.3)	4 (7.7)	5 (5.5)	8 (12.3)	8 (13.6)
0.229						
3 month follow-up						
Composite criteria (n=340)*†	2 (6.7)	0 (0)	5 (9.6)	6 (6.6)	10 (15.4)	1 (1.7)
0.016						
Hospital admissions (n=24)‡						
0.951						
Referral to hospital by the GP in the month after the first consultation	2 (100)	0 (0)	3 (60.0)	4 (66.7)	6 (60.0)	1 (100)
Hospital admission without referral by the GP						
0 (0)						
Persistent symptoms (n=323)	1 (3.6)	4 (9.3)	9 (18.4)	10 (11.6)	16 (26.2)	18 (32.1)
0.002						
Asthenia	0 (0.0)	1 (2.3)	2 (4.1)	3 (3.5)	7 (11.5)	9 (16.1)
Cough	1 (3.6)	0 (0.0)	0 (0.0)	2 (2.3)	2 (3.3)	2 (3.6)
Respiratory discomfort	0 (0.0)	0 (0.0)	0 (0.0)	1 (1.2)	3 (4.9)	1 (1.8)
Dyspnoea	0 (0.0)	1 (2.3)	4 (8.2)	1 (1.2)	6 (9.8)	4 (7.1)
Chest pain	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	2 (3.3)	2 (3.6)
Anosmia	0 (0.0)	3 (7.0)	4 (8.2)	4 (4.7)	2 (3.3)	5 (8.9)
Ageusia	0 (0.0)	1 (2.3)	3 (6.1)	2 (2.3)	0 (0.0)	2 (3.6)
Pain (aches, myalgia)	0 (0.0)	0 (0.0)	0 (0.0)	2 (2.3)	0 (0.0)	2 (3.6)
Headache	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (1.6)	0 (0.0)
Other§	0 (0.0)	0 (0.0)	0 (0.0)	2 (2.3)	1 (1.6)	5 (8.9)

Continued

Table 3 Continued

	LC1 (n=30)	LC2 (n=43)	LC3 (n=52)	LC4 (n=91)	LC5 (n=65)	LC6 (n=59)
		'Anosmia and/or ageusia'	'Influenza-like syndrome with anosmia and ageusia'	'Influenza-like syndrome without anosmia or ageusia'	'Influenza-like syndrome with respiratory impairment'	'Complete form'
6-month follow-up	(n=27)	(n=41)	(n=47)	(n=86)	n=61	n=49
Persistent symptoms	0 (0.0)	1 (2.4))	3 (6.4)	3 (3.5)	8 (13.1)	5 (10.2)
						0.096

Data are quoted as the number (%).

P value in a  $\chi^2$  test, Fisher's test or the Kruskal-Wallis test.

The p value was calculated using the false-discovery rate method for multiple comparisons.

\*The composite outcome included COVID-19-related death and/or hospital admission. No COVID-19-related deaths were recorded.

†Three patients requiring respiratory support were admitted to intensive care units (one in each of the following classes: LC3, LC4 and LC5).

‡All hospitalisations required by GPs occurred in the month following the first visit.

§Included: difficulty concentrating/memory impairment (n=3) deep vein thrombosis (n=1), alopecia (n=1), palpitations (n=1) and feeling feverish (n=2).

BMI, body mass index; GPs, general practitioners.

two mild forms (clusters 1 and 2, characterised by upper respiratory tract symptoms) and four clusters (3–6) with higher proportions of patients requiring respiratory support (ranging from 8.6% to 19.8%). Clusters 3 was chiefly characterised by gastrointestinal symptoms, and the patients in clusters 5 and 6 (which could have been called ‘complete forms’) had the highest number of symptoms and the highest hospital admission rates (27.2 and 45.5%, respectively). In our study, the highest hospital admission rate was observed for LC5 (chiefly characterised by the highest prevalence of abnormal lung auscultation findings), rather than for the ‘complete form’ (LC6). It is well known that respiratory impairment with abnormal lung auscultation findings is associated with more severe COVID-19.<sup>16</sup> This observation underlines the relevance of the clinical examination performed in our study. Moreover, most of hospital admissions in our study occurred early in the course of disease and were, therefore, more likely to be related to the severity of COVID-19 than to long-term symptoms.

The form with ‘anosmia and/or ageusia’ (LC2) was characterised by the lowest prevalence of asthenia and was similar to Sudre *et al*’s cluster 1<sup>2</sup>. Apart from LC6 (the complete form), the patients in the classes with the highest probabilities of anosmia and/or ageusia (LC2–3) were younger than those in all the other classes.<sup>17</sup> It should be noted that in another LCA-based study of COVID-19 symptoms, the study population comprised both COVID-19 patients and non-COVID-19 patients.<sup>18</sup>

The demographic characteristics of our COVID-19 patients consulting a GP were similar to those reported in the literature—especially with regard to age and the most frequent comorbidities (hypertension and diabetes).<sup>19</sup> Likewise, the prevalences of COVID-19 symptoms in our study (including asthenia, fever, cough, myalgia and headache) were similar to those in the literature.<sup>19–22</sup> Anosmia and ageusia were also common and are considered to be specific for COVID-19.<sup>17–22</sup> In line with the literature data, our patients' course of disease was generally mild and did not often require hospital admission.<sup>5–19–23</sup> Various studies have shown that hospital admission and deaths are associated with older age,<sup>24</sup> male sex<sup>25</sup> and the presence of comorbidities.<sup>19–26</sup>

At the 3-month time point, we observed persistent COVID-19 symptoms in 18% of the patients. This result is in line with the values (from 10% to 30%) observed in other French studies.<sup>24–27</sup> The main risk factor for developing persistent symptoms was membership of LC5 ('influenza-like with respiratory impairment') or LC6 ('complete form'). These findings are also consistent with literature data showing that the presence of more than six initial symptoms<sup>28</sup> and hospital admission (most common in LC5)<sup>28–30</sup> are risk factors for persistent symptoms. As found in the literature, the most frequent symptoms observed at 3 and 6 months in this study were asthenia, anosmia and dyspnoea.<sup>31–33</sup>



Table 4 Pairwise comparison of illustrative variables and 3-month and 6-month outcomes in the six-latent classes (LCs) identified (using posterior probabilities classifications; n=340)													
	P value for pairwise comparisons*												
	LC1		LC2		LC3		LC4		LC5		LC6		
	vs		vs		vs		vs		vs		vs		
	LC2	LC3	LC4	LC5	LC6	LC3	LC4	LC5	LC6	LC4	LC5	LC6	LC6
Age, years, median (IQR)	0.003	0.002	0.962	0.773	0.778	0.890	<0.001	<0.001	<0.001	<0.001	0.001	0.592	0.675
Cross-sectional analysis													0.366
Comorbidity													
Cardiac and cardiovascular comorbidities (n=340)	0.015	0.008	0.424	0.788	0.160	0.788	0.043	0.015	0.195	0.015	0.008	0.124	0.424
Heart diseases													0.160
Vascular disease													
Hypertension	0.053	0.035	0.658	0.669	0.203	1	0.053	0.053	0.273	0.035	0.035	0.203	0.273
Diabetes (n=340)													0.273
Cancer and/or immunodeficiency (n=340)													
Respiratory disease (n=340)													
Obesity (BMI>30) (n=340)													
3 month follow-up													
Composite criteria (n=340)†‡	0.377	1	1	0.483	0.437	0.305	0.377	0.053	1	0.661	0.483	0.321	0.437
Persistent symptoms (n=323)	0.688	0.178	0.394	0.045	0.023	0.355	0.773	0.078	0.035	0.394	0.410	0.201	0.066
													0.556
LC1: 'paucisymptomatic'; LC2: 'anosmia and/or ageusia'; LC3: 'influenza-like syndrome with anosmia and ageusia'; LC4: 'influenza-like syndrome without anosmia or ageusia'; LC5: 'influenza-like syndrome with respiratory impairment'; LC6: 'complete form'. Data are quoted as the number (%). P value in a $\chi^2$ test, Fisher's test or the Kruskal-Wallis test. *The p value was calculated using the false-discovery rate method for multiple comparisons. †The composite outcome included COVID-19-related death and/or hospital admission. No COVID-19-related deaths were recorded. ‡Three patients requiring respiratory support were admitted to intensive care units (one in each of the following classes: LC3, LC4 and LC5). BMI, body mass index.													

## Applicability of the findings

Some studies have shown that loss of smell in COVID-19 patients was less common during the Omicron wave than during the Delta wave.<sup>11</sup> In contrast, sore throat was more common during the Omicron wave than during the Delta wave.<sup>11</sup> Furthermore, the hospital admission rate was lower during the Omicron wave than during the Delta wave.<sup>11</sup> However, if the prevalence of signs and symptoms could have differed across variants, signs and symptoms remain similar. We, therefore, believe that these potential differences had little effect on the nature of the LCs but might have influenced their prevalence. The literature data on the long-term symptoms associated with the different variants are contradictory, although vaccination might prevent these symptoms.<sup>12–15</sup> One study showed that persistent symptoms after SARS-CoV-2 infection were more common before the Delta wave than during the Delta and Omicron waves.<sup>13</sup> However, the fact that these differences were no longer statistically significant after adjustment for vaccination status suggested that COVID-19 vaccines reduced the risk of long-term symptoms.<sup>13</sup> A review found that compared with previous variants of SARS-CoV-2, Omicron infections were associated with fewer long-COVID symptoms; however, the small number of studies and the lack of controls for potentially confounding variables (eg, reinfections and vaccination status) in some studies limited the results' generalisability.<sup>14</sup> It appears that individuals infected with the wild-type variant were more likely to develop long-COVID symptoms. In contrast, the results of another review suggested that there are no significant intervariant differences in long COVID-19 other than for certain general symptoms (with the Alpha and Omicron variants) and difficulty sleeping (for the wild-type variant).<sup>15</sup> These findings emphasise the need to identify patients with an elevated risk of developing long-term symptoms (eg, non-vaccinated patients and/or patients with previous variants).

## CONCLUSION

By using a data-driven approach to analyse COVID-19 signs and symptoms, we identified six clinical profiles among patients managed by their GP. Our results highlighted associations with hospital admission and the persistence of symptoms at 3 months. Since most studies of the presentation and clinical course of COVID-19 have been hospital-based, it is important to provide primary care-specific data that might help GPs to optimise patient management. GPs diagnose the majority of patients with COVID-19 and thus have an essential role in combating the ongoing pandemic.<sup>34 35</sup> Our results might help GPs to (1) identify at-risk profiles for hospital admission and persistent symptoms, (2) set up procedures for closer follow-up, (3) anticipate possible worsening<sup>35</sup> and (4) manage complications as early as possible. The higher prevalence of persistent symptoms in some COVID-19 profiles suggests that the corresponding patients should

be followed up by their GPs, who are well placed to take account of the disease's impact on quality of life and overall health via a patient-centred approach.<sup>36</sup> Thus, our findings may help GPs to improve the follow-up of COVID-19 patients in primary care.

## Author affiliations

<sup>1</sup>Univ Paris Est Creteil, INSERM, IMRB, CepiA Team, Creteil, France

<sup>2</sup>Department of General Practice, Univ Paris Est Creteil, Creteil, France

<sup>3</sup>Maison de Sante Pluriprofessionnelle Universitaire (MSPU) Saint-Maur-Des-Fosses, MSPU, Saint-Maur-Des-Fosses, France

<sup>4</sup>Maison de Sante Pluriprofessionnelle Universitaire (MSPU) Torcy, MSPU, Torcy, France

<sup>5</sup>AP-HP, Hôpital Henri-Mondor, Clinical Research Unit (URC Mondor), AP-HP, Creteil, France

<sup>6</sup>Maison de Sante Pluriprofessionnelle Universitaire (MSPU) Fontainebleau, MSPU, Fontainebleau, France

<sup>7</sup>Maison de Sante Pluriprofessionnelle Universitaire (MSPU) Nemours, MSPU, Nemours, France

<sup>8</sup>APHP, Hôpital Henri-Mondor, Public Health Department, AP-HP, Creteil, France

**Acknowledgements** We thank Camille Hug and Faïka Bacar for help with data collection and interpretation. We also thank all the study participants and the participating health centres. Lastly, we thank Dr David Fraser (Biotech Communication SARL, Ploudalmézeau, France) for copy-editing assistance.

**Contributors** Conceptualisation: EF. Methodology: EF and SB-G. Data collection: EM. Data analysis: EB. Data interpretation: EF, SB-G and EB. Drafting and revision of the manuscript: WM, EF and SB-G. Final revision of the manuscript: EF, WM, EB, EM, SB, J-DH, EA, T-TP and SB-G. All authors have read and agreed to the published version of the manuscript. EF is the guarantor.

**Funding** The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interests** None declared.

**Patient and public involvement** Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

**Patient consent for publication** Not applicable.

**Ethics approval** This study involves human participants and all patients received written information about the study and provided their verbal consent to participation. The study database was registered with the French National Data Protection Commission (Commission nationale de l'informatique et des libertés (Paris, France); reference: 2211627 v0). The study protocol was approved by an independent ethics committee (Comité de Protection des Personnes Est IV (Strasbourg, France); reference: IDRCB 2020-A01693-36. Participants gave informed consent to participate in the study before taking part.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** All data relevant to the study are included in the article or uploaded as online supplemental information.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

## ORCID iDs

William Mirat <http://orcid.org/0000-0001-6143-7837>

Tan-Trung Phan <http://orcid.org/0000-0001-9990-7398>

## REFERENCES

- 1 Coronavirus: chiffres clés et évolution de la COVID-19 en France et dans le Monde, 2023. Available: <https://www.santepubliquefrance.fr/dossiers/coronavirus-covid-19/coronavirus-chiffres-cles-et-evolution-de-la-covid-19-en-france-et-dans-le-monde>
- 2 Sudre CH, Lee KA, Lochlainn MN, *et al.* Symptom clusters in COVID-19: A potential clinical prediction tool from the COVID symptom study App. *Sci Adv* 2021;7:eabd4177.
- 3 Ferrat E, Audureau E, Paillaud E, *et al.* Four distinct health profiles in older patients with cancer: latent class analysis of the prospective ELCAPA cohort. *GERONA* 2016;71:1653–60.
- 4 Phan T-T, Mirat W, Brossier S, *et al.* Initial characteristics and course of disease in patients with suspected COVID-19 managed in general practice: a prospective, Multicentre cohort study. *BMJ Open* 2023;13:e068424.
- 5 Haut Conseil de la Santé Publique, 2020. Available: [https://solidarites-sante.gouv.fr/IMG/pdf/hcsp\\_sars-cov2-priorisationtest-sdiagnostiques-2020-03-11-2.pdf](https://solidarites-sante.gouv.fr/IMG/pdf/hcsp_sars-cov2-priorisationtest-sdiagnostiques-2020-03-11-2.pdf)
- 6 Haute Autorité de Santé. Réponses rapides dans le cadre de la Covid-19. Prise en charge de premier recours des patients suspectés de Covid-19, 2020. Available: [https://www.has-sante.fr/upload/docs/application/pdf/2020-11/reco410\\_rr\\_covid\\_premier\\_recours\\_maj\\_mel\\_vd\\_2020-11-06\\_19-00-27\\_357.pdf](https://www.has-sante.fr/upload/docs/application/pdf/2020-11/reco410_rr_covid_premier_recours_maj_mel_vd_2020-11-06_19-00-27_357.pdf)
- 7 Société Française de Radiologie (2020 March) Recommandations dans un contexte de Covid-19, 2020. Available: <http://bulletin.radiologie.fr/actualites-covid-19/epidemie-covid-19-point-limagerie>
- 8 McCutcheon AL. Latent class analysis. In: *Latent Class Analysis*. 2455 Teller Road, Thousand Oaks California 91320 United States of America : Sage Publications, 1987. Available: <https://methods.sagepub.com/book/latent-class-analysis>
- 9 Vermunt JK, Magidson J. Technical Guide for Latent GOLD 5.0. Basic, Advanced, and Syntax. Belmont, MA: Statistical Innovations Inc, 2013.
- 10 Benjamini Y, Hochberg Y. Hochberg Y(1995) controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Royal Stat Soc Series B* 1995;57:289–300.
- 11 Menni C, Valdes AM, Polidori L, *et al.* Symptom prevalence, duration, and risk of hospital admission in individuals infected with SARS-Cov-2 during periods of Omicron and Delta variant dominance: a prospective observational study from the ZOE COVID study. *Lancet* 2022;399:1618–24.
- 12 Watanabe A, Iwagami M, Yasuhara J, *et al.* Protective effect of COVID-19 vaccination against long COVID syndrome: A systematic review and meta-analysis. *Vaccine* 2023;41:1783–90.
- 13 Gottlieb M, Wang RC, Yu H, *et al.* Innovative support for patients with SARS-Cov-2 infections Registry (INSPIRE) group. severe fatigue and persistent symptoms at 3 months following severe acute respiratory syndrome Coronavirus 2 infections during the pre-Delta, Delta, and Omicron time periods: A multicenter prospective cohort study. *Clin Infect Dis* 2023;76:1930–41.
- 14 Fernández-de-Las-Peñas C, Notarte KI, Peligro PJ, *et al.* Long-COVID symptoms in individuals infected with different SARS-Cov-2 variants of concern: A systematic review of the literature. *Viruses* 2022;14:2629.
- 15 Du M, Ma Y, Deng J, *et al.* Comparison of long COVID-19 caused by different SARS-Cov-2 strains: A systematic review and meta-analysis. *Int J Environ Res Public Health* 2022;19:16010.
- 16 Pancaldi F, Pezzuto GS, Cassone G, *et al.* VECTOR: an algorithm for the detection of COVID-19 pneumonia from Velcro-like lung sounds. *Comput Biol Med* 2022;142.
- 17 Giacomelli A, Pezzati L, Conti F, *et al.* Self-reported olfactory and taste disorders in patients with severe acute respiratory Coronavirus 2 infection: A cross-sectional study. *Clin Infect Dis* 2020;71:889–90.
- 18 Moreira R da S. Latent class analysis of COVID-19 symptoms in Brazil: results of the PNAD-Covid19 survey. *Cad Saude Publica* 2021;37.
- 19 Yordanov Y, Dinh A, Bleibtreu A, *et al.* Clinical characteristics and factors associated with hospital admission or death in 43 103 adult outpatients with Coronavirus disease 2019 managed with the Covidom Telesurveillance solution: a prospective cohort study. *Clin Microbiol Infect* 2021;27:1158–66.
- 20 Huang C, Wang Y, Li X, *et al.* Clinical features of patients infected with 2019 novel Coronavirus in Wuhan, China. *Lancet* 2020;395:497–506.
- 21 Docherty AB, Harrison EM, Green CA, *et al.* Features of 20 133 UK patients in hospital with COVID-19 using the ISARIC WHO clinical Characterisation protocol: prospective observational cohort study. *BMJ* 2020;369:m1985.
- 22 Khan M, Khan H, Khan S, *et al.* Epidemiological and clinical characteristics of Coronavirus disease (COVID-19) cases at a screening clinic during the early outbreak period: a single-centre study. *J Med Microbiol* 2020;69:1114–23.
- 23 Bernard Stoecklin S, Rolland P, Silue Y, *et al.* Investigation team. First cases of Coronavirus disease 2019 (COVID-19) in France: surveillance, investigations and control measures. *Euro Surveill* 2020;25:2000094.
- 24 Dres M, Hajage D, Lebbah S, *et al.* Characteristics, management, and prognosis of elderly patients with COVID-19 admitted in the ICU during the first wave: insights from the COVID-ICU study: prognosis of COVID-19 elderly critically ill patients in the ICU. *Ann Intensive Care* 2021;11:77.
- 25 Peckham H, de Grujter NM, Raine C, *et al.* Male sex identified by global COVID-19 meta-analysis as a risk factor for death and ICU admission. *Nat Commun* 2020;11:6317.
- 26 Semenzato L, Botton J, Drouin J, *et al.* Maladies chroniques, états de santé et risque d'hospitalisation et de décès hospitalier pour COVID-19: analyse comparative de données des deux vagues épidémiques de 2020 en France à partir d'une cohorte de 67 millions de personnes, 2021. Available: [https://www.epi-phare.fr/app/uploads/2021/07/epi-phare\\_risques\\_hospitalisations\\_deces\\_covid19\\_20210721.pdf](https://www.epi-phare.fr/app/uploads/2021/07/epi-phare_risques_hospitalisations_deces_covid19_20210721.pdf)
- 27 Humberston B. Prevalence of ongoing symptoms following Coronavirus (COVID-19) infection in the UK. 2021. Available: <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/prevalenceofongoingsymptomsfollowingcoronaviruscovid19infectionintheuk/1april2021>
- 28 Silverberg JI, Zyskind I, Naiditch H, *et al.* Predictors of chronic COVID-19 symptoms in a community-based cohort of adults. *PLoS One* 2022;17:e0271310.
- 29 Santé publique France. Coronavirus: chiffres clés et évolution de la COVID-19 en France et dans le Monde, Available: <https://www.santepubliquefrance.fr/dossiers/coronavirus-covid-19/coronavirus-chiffres-cles-et-evolution-de-la-covid-19-en-france-et-dans-le-monde#block-266151>
- 30 Carfi A, Bernabei R, Landi F. Against COVID-19 post-acute care study group. persistent symptoms in patients after acute COVID-19. *JAMA* 2020;324:603–5.
- 31 Nasserie T, Hittle M, Goodman SN. Assessment of the frequency and variety of persistent symptoms among patients with COVID-19: A systematic review. *JAMA Netw Open* 2021;4:e2111417.
- 32 Ghosn J, Piroth L, Epaulard O, *et al.* Persistent COVID-19 symptoms are highly prevalent 6 months after hospitalization: results from a large prospective cohort. *Clin Microbiol Infect* 2021;27.
- 33 Lechien JR, Chiesa-Estomba CM, Place S, *et al.* COVID-19 task force of YO-IFOS. clinical and Epidemiological characteristics of 1420 European patients with mild-to-moderate Coronavirus disease 2019. *J Intern Med* 2020;288:335–44.
- 34 Sarti TD, Lazarini WS, Fontenelle LF, *et al.* What is the role of primary health care in the COVID-19 pandemic. *Epidemiol Serv Saude* 2020;29.
- 35 Greenhalgh T, Koh GCH, Car J. Covid-19: a remote assessment in primary care. *BMJ* 2020;368:m1182.
- 36 Stewart M, Brown JB, Weston WW, *et al.* Patient-centered medicine. In: *Patient-Centered Medicine. Transforming the Clinical Method*. 3rd edn. London: CRC Press, 2019. Available: <https://www.taylorfrancis.com/books/9781909368040>