Clinical clustering with prognostic implications in Japanese COVID-19 patients: report from Japan COVID-19 Task Force, a nation-wide consortium to investigate COVID-19 host genetics

Shiro Otake1, Shotaro Chubachi1*, Ho Namkoong1, Kensuke Nakagawa1, Hiromu Tanaka1, Ho Lee1, Atsuho Morita1, Takahiro Fukushima1, Mayuko Watase1, Tatsuya Kusumoto1, Katsunori Masaki1, Hirofumi Kamata1, Makoto Ishii1, Naoki Hasegawa1, Norihito Harada1, Tetsuya Ueda4, Solicho Ueda5, Takashi Ishiguro6, Ken Arimura7, Fukuki Saito8, Takashi Yoshiyama6, Yasushi Nakano10, Yoshikazu Mutoh11, Yusuke Suzuki12, Koji Murakami13, Yukinori Okada14, Ryuji Koike15, Yuko Kitagawa16, Akinori Kimura17, Seiya Imoto18, Satoru Miyano19, Seishi Ogawa20, Takanori Kanai21, Koichi Fukunaga1 and The Japan COVID-19 Task Force

Abstract

Background: The clinical course of coronavirus disease (COVID-19) is diverse, and the usefulness of phenotyping in predicting the severity or prognosis of the disease has been demonstrated overseas. This study aimed to investigate clinically meaningful phenotypes in Japanese COVID-19 patients using cluster analysis.

Methods: From April 2020 to May 2021, data from inpatients aged ≥ 18 years diagnosed with COVID-19 and who agreed to participate in the study were collected. A total of 1322 Japanese patients were included. Hierarchical cluster analysis was performed using variables reported to be associated with COVID-19 severity or prognosis, namely, age, sex, obesity, smoking history, hypertension, diabetes mellitus, malignancy, chronic obstructive pulmonary disease, hyperuricemia, cardiovascular disease, chronic liver disease, and chronic kidney disease.

Results: Participants were divided into four clusters: Cluster 1, young healthy (n = 266, 20.1%); Cluster 2, middle-aged (n = 245, 18.5%); Cluster 3, middle-aged obese (n = 435, 32.9%); and Cluster 4, elderly (n = 376, 28.4%). In Clusters 3 and 4, sore throat, dysosmia, and dysgeusia tended to be less frequent, while shortness of breath was more frequent. Serum lactate dehydrogenase, ferritin, KL-6, d-dimer, and C-reactive protein levels tended to be higher in Clusters 3 and 4. Although Cluster 3 had a similar age as Cluster 2, it tended to have poorer outcomes. Both Clusters 3 and 4 tended to exhibit higher rates of oxygen supplementation, intensive care unit admission, and mechanical ventilation, but the mortality rate tended to be lower in Cluster 3.
Background
In December 2019, a disease outbreak was noticed after a massive admission of patients with common clinical symptoms of pneumonia in the local hospitals of Wuhan City, China. Upon further investigations, the World Health Organization confirmed that the novel coronavirus, Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), was responsible for these clinical symptoms and further denominated this disease as coronavirus disease (COVID-19) [1]. Its clinical course is diverse, ranging from mild self-limited illness to life-threatening organ dysfunctions [2–4].

Identifying disease sub-phenotypes could improve the understanding of the pathophysiology of critical care syndromes and lead to the discovery of new treatment targets by allowing future therapeutic trials to focus on predicted responders [5]. COVID-19 cluster analysis was previously used to identify distinct sub-phenotypes based on clinical and biochemical characteristics [6–10], for other heterogeneous syndromes, such as acute respiratory distress syndrome, sepsis, and acute kidney injury [11]. However, the main factors in the cluster analysis and methodology differed among these studies, as did the characteristics of the sub-phenotypes. Moreover, most studies used not only baseline characteristics but also laboratory test results and radiographic patterns [7–10].

Previous reports, including ours, revealed that baseline characteristics, such as age, sex, and comorbidities, can predict meaningful outcomes of COVID-19 [12, 13]. The clinical characteristics of COVID-19 may differ depending on the population. For instance, COVID-19 is milder in Japan than in other countries [14, 15]. Population differences may be influenced by complex factors, including the number of patients, medical infrastructure, resources of medical personnel, and patient background [15]. To the best of our knowledge, no clinical studies to date have examined the phenotypes of COVID-19 patients in Japan.

Based on the above, we hypothesized that cluster analysis using baseline characteristics reportedly related to COVID-19 outcomes may allow for simple meaningful phenotyping of Japanese COVID-19 patients, and that sub-phenotypes may differ according to population differences and cluster analysis methods. The present study aimed to demonstrate the usefulness of phenotyping in predicting meaningful outcomes of Japanese COVID-19 patients and to capture the patients’ post-hospitalization course.

Methods
Study design and settings
All COVID-19 cases in this retrospective cohort study were recruited through the Japan COVID-19 Task Force [16, 17]. From April 2020 to May 2021, data from consecutive inpatients aged ≥ 18 years diagnosed with COVID-19, using SARS-CoV2 polymerase chain reaction (PCR) test results at one among the > 100 affiliated hospitals, and who agreed to cooperate in the study were registered in an electronic case record form by the study subspecialist at the affiliated research institute. Patients meeting any of the following exclusion criteria were excluded: (i) non-Japanese patients, (ii) patients with incomplete medical records, such as missing outcome information, and (iii) patients lacking any of the selected 12 variables for cluster analysis (Fig. 1). All patients provided written informed consent. This study was approved by the ethics committees of Keio University School of Medicine (20200061) and related research institutions. All aspects of the study conformed to the principles of the Declaration of Helsinki adopted by WMA General Assembly, Fortaleza, Brazil, October 2013.
Data collection
The following information was extracted from the electronic case record form: age, sex, height, weight, clinical symptoms and signs, laboratory findings on admission, comorbidities, disease severity (supplementary oxygen, intensive care unit (ICU) entry, need for invasive mechanical ventilation, and survival status), and treatment details. We defined disease severity as follows: most severe, need for support by high-flow oxygen devices, invasive mechanical ventilation, extracorporeal membrane oxygenation, or death; severe, need for support of low-flow oxygen devices; mild, symptomatic patients not requiring oxygen support; asymptomatic, asymptomatic patients without oxygen support [18]. All laboratory tests were performed according to the patients’ clinical care needs. Symptoms and signs were included not only at the time of referral and admission, but also during hospitalization. Blood tests such as biochemistry, peripheral blood analysis, and coagulation were performed within 48 h of the initial visit or admission. The collected data were reviewed by a team of respiratory clinicians. If core data were missing, the clinician who first diagnosed the disease was contacted to collect it. Missing or absent data in the patient background were noted as unknown.

Identification of COVID-19 phenotypes using cluster analysis
We selected 12 clinically relevant patient baseline characteristics reportedly associated with the severity or prognosis of COVID-19 [12, 19–25], namely, age, sex, obesity, smoking history, hypertension, diabetes mellitus, malignancy, chronic obstructive pulmonary disease, hyperuricemia, cardiovascular disease, chronic liver disease, and chronic kidney disease. We defined obesity as body mass index (BMI) > 25 and treated it as a nominal variable.

Statistical analysis
Data are presented as means ± standard deviation (SD). Data were compared among groups using analysis of variance (ANOVA) and χ² tests. Hierarchical cluster analysis using the 12 variables mentioned above was performed using the Ward’s minimum-variance method [26, 27]. The results are graphically depicted by a dendrogram. Statistical significance was set at p < 0.05. All data were analyzed using the JMP 16 software (SAS Institute, Cary, NC, USA).

Results
Characteristics of the study population
Table 1 shows the baseline clinical characteristics of the participants. A total of 1322 inpatients (men, 65.1%; mean age, 58 ± 18.1 years) were enrolled in this study. The mean BMI was 24.4 ± 4.7 kg/m², and 597 (45.2%) had a history of smoking. Based on their clinical presentation, participants were classified into the most severe (n = 63, 4.8%), severe (n = 426, 32.2%), mild (n = 777, 58.8%), and asymptomatic (n = 56, 4.2%) disease groups. The most common comorbidities were hypertension (n = 449, 34%), diabetes mellitus (n = 263, 19.9%), and hyperuricemia (n = 134, 10.1%).

Comparison of baseline characteristics among clusters
We performed Ward's cluster analysis based on 12 factors reportedly associated with the severity or prognosis of COVID-19 [12, 19–25]. Based on visual assessment
of the resulting dendrogram (Fig. 2), data could be optimally grouped into four clusters, with each cluster corresponding to a potential phenotype. Table 2 presents the baseline characteristics of each cluster. Cluster 1 (young healthy cluster: n = 266) included the youngest population and tended to have fewer comorbidities than the other clusters. Cluster 3 (middle-aged obese cluster: n = 435) included mostly middle-aged patients, had the highest percentage of men with higher BMI and numerous comorbidities, such as hypertension, diabetes mellitus, and hyperuricemia. Although patients in Cluster 2 (middle-aged cluster: n = 245) were in the same age group as those in Cluster 3, they tended to have a lower BMI and fewer comorbidities compared to those in Cluster 3. Compared to other clusters, Cluster 4 (elderly: n = 376) included the oldest patients who tended to have numerous comorbidities, such as malignancy, cardiovascular diseases, and chronic kidney disease.

Comparison of clinical characteristics and laboratory findings among clusters
Table 3 shows a comparison of the subjective symptoms and physical findings among the four clusters. Sore throat, dysosmia, and dysgeusia, all reported as good prognostic factors [12, 28, 29], tended to be more frequent in Cluster 1 than in other clusters. In contrast, shortness of breath, reported as a poor prognostic factor [30], tended to be less frequent in Cluster 1 than in other clusters. Cluster 4 exhibited the lowest prevalence of sore throat, dysosmia, and dysgeusia among the four clusters, but more frequent consciousness disturbance, reportedly a poor prognostic factor [31], than other clusters. Table 4 shows a comparison of the laboratory findings among the clusters. Platelet count, reported as a poor prognostic factor [32], tended to be lower in Clusters 3 and 4, while lactate dehydrogenase (LDH), ferritin, Krebs von den Lungen-6 (KL-6), D-dimer, and C-reactive protein (CRP), also considered poor prognostic factors [33–35], tended to be lower in Cluster 1 and higher in Clusters 3 and 4. These results imply that Cluster 1 had COVID-19 related symptoms and laboratory findings associated with good prognosis, while Clusters 3 and 4 had poor prognosis.

Comparison of clinical outcomes between the four clusters
A comparison of the rate of supplemental oxygen needs, ICU admission, mechanical ventilation, and mortality is shown in Fig. 3. Cluster 3 exhibited a higher rate of patient receiving supplementary oxygen and/or mechanical ventilation, admitted to the ICU, and mortality compared to Clusters 1 and 2. Cluster 2 had intermediate rates of the above factors, between Clusters 1 and 3, and Cluster 1 exhibited the most favorable outcomes among all the clusters. Similar to Cluster 3, Cluster 4 also tended to have poor outcomes, coupled with a higher mortality rate. These results suggest that middle-aged obese men tend to have a similarly serious course as the elderly but with a lower risk of death. Consistent with the high rate of severe disease in Clusters 3 and 4, patients in these clusters received intensive drug treatment, including remdesivir and glucocorticoids, of current frequent use and considered to be effective in the treatment of COVID-19 [36] (Table 5).

Table 2 Baseline characteristics for each cluster

|                      | Cluster 1: Young healthy | Cluster 2: Middle aged | Cluster 3: Middle aged obese | Cluster 4: Elderly | p-value  |
|----------------------|--------------------------|------------------------|-------------------------------|-------------------|----------|
| n                    | 266                      | 245                    | 435                           | 376               | <0.0001  |
| Age, years           | 31.7 ± 0.6               | 61.1 ± 0.6             | 56.9 ± 0.5                    | 76.1 ± 0.5        | <0.0001  |
| Male, n (%)          | 135 (50.8)               | 180 (73.5)             | 206 (47.4)                    | 181 (48.1)        | <0.0001  |
| BMI, kg/m²           | 22.9 ± 0.2               | 22.2 ± 0.2             | 28.3 ± 0.2                    | 22.3 ± 0.2        | <0.0001  |
| Smoking history, n (%)| 98 (37.2)                | 117 (47.8)             | 247 (56.8)                    | 134 (35.6)        | <0.0001  |
| Hypertension, n (%)  | 0 (0)                    | 13 (5.3)               | 18 (4.1)                      | 62 (16.5)         | <0.0001  |
| Diabetes mellitus, n (%)| 2 (0.8)                | 56 (22.9)              | 124 (28.5)                    | 81 (21.5)         | <0.0001  |
| Malignancy, n (%)    | 6 (2.3)                  | 13 (5.3)               | 18 (4.1)                      | 62 (16.5)         | <0.0001  |
| COPD, n (%)          | 0 (0)                    | 31 (12.7)              | 16 (3.7)                      | 17 (4.5)          | <0.0001  |
| Hyperuricemia, n (%) | 4 (1.5)                  | 4 (1.6)                | 98 (22.5)                     | 28 (7.5)          | <0.0001  |
| Cardiovascular disease, n (%)| 1 (0.4)          | 1 (0.4)                | 47 (10.8)                     | 65 (17.3)         | <0.0001  |
| Chronic liver disease, n (%)| 0 (0)                  | 4 (1.6)                | 29 (6.7)                      | 10 (2.7)          | <0.0001  |
| Chronic kidney disease, n (%)| 1 (0.4)              | 14 (5.7)               | 34 (7.8)                      | 42 (11.2)         | <0.0001  |

Data are shown as mean ± SD. Data were compared among groups using analysis of variance (ANOVA) and χ² tests.
BMI, body mass index; COPD, chronic obstructive pulmonary disease.
Discussion

This study was the first in Japan to perform a cluster analysis of COVID-19 patients. We identified four clinical sub-phenotypes, namely the “young healthy cluster” (Cluster 1), “middle-aged cluster” (Cluster 2), “middle-aged obese cluster” (Cluster 3), and “elderly cluster” (Cluster 4), which were associated with different outcomes in Japanese patients with COVID-19. Previous reports, including ours, have shown that comorbidities and mortality rates in Japan differed from inpatient studies in other countries [15, 17]. Thus, the identification of the meaningful sub-phenotypes of Japanese COVID-19 patients is important. Notably, our study used simple baseline characteristics as variables for cluster analysis. Several previous studies have shown that cluster analysis is useful for phenotyping and predicting COVID-19 outcomes [6–10]. However, most of these studies used complicated variables, combining a wide range of blood test results for clustering. Promptly indefinable is an important feature for defining COVID-19 sub-phenotypes [37]. We believe that the present simple clustering may be of great help to clinicians in predicting prognosis and performing individualized therapy.

Cluster 3 included mainly middle-aged patients with a high BMI, and a high rate of complications from lifestyle-related diseases, such as hypertension, diabetes, and hyperuricemia. Even though hyperuricemia has been previously reported to be associated with prognosis [38, 39], its rate was higher in Cluster 3 than in Cluster 4, which showed the highest mortality rate. This finding may be due to a possible association between obesity and hyperuricemia [40, 41]. Cluster 2 patients were similarly middle-aged but had lower BMI and lifestyle-related diseases. Cluster 3 revealed poorer outcomes, including need for oxygen, ICU admission, and intubation, than Cluster 2. This result is consistent with the fact that obesity has already been reported as a poor prognostic factor for COVID-19 [20], as have lifestyle-related diseases [12, 21, 22]. However, the mortality rate of Cluster 3 was lower than that of Cluster 4. Despite the high risk of severe disease, there is still life-saving potential, suggesting that this cluster is likely to benefit from aggressive intensive care.

Cluster 1 consisted mainly of younger patients with fewer comorbidities. They showed the highest frequency of sore throat, dysosmia, and dysgeusia of all the clusters, and the outcomes were generally the most favorable. In addition, several biomarkers (LDH, ferritin, KL-6, d-dimer, and CRP) [33–35] reported as poor prognosis predictors were lower in Cluster 1 than in other clusters. A majority of young people with COVID-19 are reported to be asymptomatic or have few symptoms [42], and this cluster also tended to have fewer symptoms than other clusters, except for upper respiratory tract symptoms. It is possible that this group may have contributed to the spread of the disease.

Table 3: Comparison of subjective symptoms and physical findings among the four clusters

|                   | Cluster 1: Young healthy | Cluster 2: Middle aged | Cluster 3: Middle aged obese | Cluster 4: Elderly | p-value |
|-------------------|--------------------------|------------------------|-------------------------------|-------------------|---------|
| Consciousness disturbance, n (%) | 1 (0.4)                  | 4 (1.6)                | 3 (0.7)                       | 16 (4.3)          | 0.0004  |
| Cough, n (%)      | 140 (52.8)               | 158 (64.5)             | 283 (66)                      | 198 (53.2)        | 0.0001  |
| Sputum, n (%)     | 47 (17.7)                | 59 (24.5)              | 112 (25.9)                    | 78 (21)           | 0.0634  |
| Sore throat, n (%)| 91 (35)                  | 63 (26)                | 115 (26.6)                    | 64 (17.3)         | <0.0001 |
| Nasal discharge, n (%)| 62 (23.7)            | 43 (17.7)              | 71 (16.4)                     | 46 (12.4)         | 0.0029  |
| Taste disorder, n (%)| 86 (33.3)             | 39 (16)                | 75 (17.4)                     | 39 (10.5)         | <0.0001 |
| Smell disorder, n (%)| 90 (34.9)              | 32 (13.1)              | 70 (16.3)                     | 25 (6.7)          | <0.0001 |
| Shortness of breath, n (%)| 52 (20.4)            | 65 (27.1)              | 140 (32.6)                    | 93 (25.4)         | 0.005   |
| Malaise, n (%)    | 105 (39.8)               | 113 (46.3)             | 225 (52.3)                    | 147 (39.7)        | 0.0009  |
| Body temperature $\geq 37.5 ~{^\circ}C$, n (%)  | 186 (70.5)             | 213 (86.9)             | 370 (86.1)                    | 260 (69.9)        | <0.0001 |
| Systolic pressure, mmHg  | 120 ± 1.2               | 129.4 ± 1.2            | 131.6 ± 0.9                   | 132 ± 1           | <0.0001 |
| Diastolic pressure, mmHg  | 78.5 ± 0.8              | 81.4 ± 0.8             | 85.1 ± 0.6                    | 77.7 ± 0.7        | <0.0001 |
| Heart rate, bpm    | 84.4 ± 1                | 88.6 ± 1               | 90 ± 0.8                      | 84.3 ± 0.8        | <0.0001 |
| Respiratory rate, bpm | 17.5 ± 0.3             | 19.3 ± 0.3             | 19.4 ± 0.2                    | 19 ± 0.2          | <0.0001 |
| SpO$_2$, %         | 97.6 ± 0.2              | 96.3 ± 0.2             | 95.9 ± 0.1                    | 95.5 ± 0.1        | <0.0001 |

Data are shown as mean ± SD. Data were compared among groups using analysis of variance (ANOVA) and $\chi^2$ tests.

SpO$_2$, saturation of percutaneous oxygen.
Cluster 4 included predominantly older patients with comorbidities such as hypertension, diabetes, malignant disease, cardiovascular disease, and chronic kidney disease. They had the poorest outcomes in terms of oxygen demand, ICU admission, ventilator use, and death. These results were consistent with previous reports showing that old age and comorbidities are related with poor prognosis [12, 19, 21–24]. In addition, several poor prognostic biomarkers (LDH, ferritin, KL-6, D-dimer, and CRP) [33–35] were higher than those in Clusters 1 and 2. Lymphocyte count, which has been linked to severe disease and mortality,
was also lowest in Cluster 4 [43]. The mechanism of this lymphocytopenia has been previously reported to be hypercytokinemia, leading to inhibition of hematopoiesis by TNF-α [44]. In fact, Cluster 4 patients with low lymphocyte count also showed a trend toward low hemoglobin level and platelet count, consistent with previous reports. Among patients in Cluster 4, 4% were admitted to the ICU and 17.6% of intubated patients died, indicating their potential as a target for future development of COVID-19 therapy.

One of the characteristics of the present study is the inclusion of a single racial group only. Many of the previous studies on cluster analysis of COVID-19 patients included multiple racial groups in their analyses [6, 7], and each cluster had different proportions of racial groups, suggesting that the clinical characteristics also reflect the racial differences. In contrast, since only Japanese patients were analyzed in this study, we focused more on basic clinical information, such as age, weight, and comorbidities, and the characteristics of the clusters can be easily grasped.

Some potential limitations of our study need to be discussed. First, the phenotyping of infectious diseases requires consideration of both the host and pathogen. SARS-CoV-2 is prone to genetic evolution, resulting in multiple variants with different characteristics compared to ancestral strains. Specifically, the transmissibility and virulence of these variants can greatly differ [45]. However, our study had no detailed data on viral load and/or strain. Second, we had no validation cohort data, necessitating additional studies. Third, we could not compare the differences in treatment response among the clusters. Five essential criteria could help define COVID-19 subtypes: (1) biologically plausible, (2) promptly identifiable, (3) nonsynonymous, (4) reproducible, and most importantly, (5) treatment responsive. To establish precision medicine against COVID-19 disease, further studies with more detailed and representative data are warranted.

**Conclusions**

We developed a simplified tool for clustering COVID-19 patients with diverse characteristics into sub-phenotypes. We identified four clusters that predicted

---

**Table 5** Comparison of drug treatment among the four clusters

|                      | Cluster 1 Young healthy | Cluster 2 Middle aged | Cluster 3 Middle aged obese | Cluster 4 Elderly | p-value   |
|----------------------|-------------------------|-----------------------|------------------------------|-------------------|-----------|
| Antibiotics, n (%)   | 17 (6.5)                | 49 (20.3)             | 83 (19.2)                    | 90 (24.1)         | <0.0001   |
| Azithromycin, n (%)  | 21 (7.9)                | 29 (12)               | 50 (11.6)                    | 63 (16.9)         | 0.0066    |
| Ciclesonide, n (%)   | 35 (13.3)               | 48 (19.9)             | 77 (17.9)                    | 52 (14)           | 0.0957    |
| Favipiravir, n (%)   | 29 (10.9)               | 92 (38)               | 168 (38.8)                   | 130 (34.8)        | <0.0001   |
| Hydroxychloroquine, n (%) | 0 (0)                  | 2 (0.8)               | 2 (0.5)                      | 2 (0.5)           | 0.5759    |
| Lopinavir and Ritonavir, n (%) | 1 (0.4)          | 2 (0.8)               | 0 (0)                        | 2 (0.5)           | 0.3684    |
| Remdesivir, n (%)    | 10 (3.8)                | 53 (22)               | 85 (19.8)                    | 68 (18.5)         | <0.0001   |
| Nafamostat, n (%)    | 3 (1.1)                 | 15 (6.2)              | 39 (9.1)                     | 26 (7.1)          | <0.0001   |
| Anticoagulant, n (%) | 15 (5.6)                | 48 (19.8)             | 86 (19.9)                    | 98 (26.1)         | <0.0001   |
| Glucocorticoids, n (%) | 25 (9.4)               | 100 (40.8)            | 219 (50.6)                   | 179 (48)          | <0.0001   |

Data were compared among groups using χ² tests
in-hospital outcomes in a large nationwide series of Japanese COVID-19 patients. This simple clustering will be needed to develop precision medicine for COVID-19.

Acknowledgements

We would like to thank all the participants involved in this study, and all members of the Japan COVID-19 Task Force engaged in daily clinical and research work on COVID-19. All members contributed cases to this study.

Japan COVID-19 Task Force is composed of more than 70 institutions nationwide in Japan. The members who contributed to the collection and analysis of cases at each institution are shown as coauthors in the following list.

The members who contributed to the collection and analysis of cases in each institution are shown as coauthors in the following list.

The COVID-19 pandemic has highlighted the importance of precision medicine, particularly for COVID-19. The results of this study will be helpful to clinicians in predicting prognosis and outcomes in COVID-19 patients. This simple clustering will be needed to develop precision medicine for COVID-19.

We would like to thank all the participants involved in this study, and all members of the Japan COVID-19 Task Force engaged in daily clinical and research work on COVID-19. All members contributed cases to this study.

Japan COVID-19 Task Force is composed of more than 70 institutions nationwide in Japan. The members who contributed to the collection and analysis of cases at each institution are shown as coauthors in the following list.

Japanese COVID-19 patients. This simple clustering will be needed to develop precision medicine for COVID-19.

Acknowledgements

We would like to thank all the participants involved in this study, and all members of the Japan COVID-19 Task Force engaged in daily clinical and research work on COVID-19. All members contributed cases to this study.

Japan COVID-19 Task Force is composed of more than 70 institutions nationwide in Japan. The members who contributed to the collection and analysis of cases at each institution are shown as coauthors in the following list.

The members who contributed to the collection and analysis of cases in each institution are shown as coauthors in the following list.

The COVID-19 pandemic has highlighted the importance of precision medicine, particularly for COVID-19. The results of this study will be helpful to clinicians in predicting prognosis and outcomes in COVID-19 patients. This simple clustering will be needed to develop precision medicine for COVID-19.

We would like to thank all the participants involved in this study, and all members of the Japan COVID-19 Task Force engaged in daily clinical and research work on COVID-19. All members contributed cases to this study.

Japan COVID-19 Task Force is composed of more than 70 institutions nationwide in Japan. The members who contributed to the collection and analysis of cases at each institution are shown as coauthors in the following list.

The members who contributed to the collection and analysis of cases in each institution are shown as coauthors in the following list.

The COVID-19 pandemic has highlighted the importance of precision medicine, particularly for COVID-19. The results of this study will be helpful to clinicians in predicting prognosis and outcomes in COVID-19 patients. This simple clustering will be needed to develop precision medicine for COVID-19.

We would like to thank all the participants involved in this study, and all members of the Japan COVID-19 Task Force engaged in daily clinical and research work on COVID-19. All members contributed cases to this study.

Japan COVID-19 Task Force is composed of more than 70 institutions nationwide in Japan. The members who contributed to the collection and analysis of cases at each institution are shown as coauthors in the following list.

The members who contributed to the collection and analysis of cases in each institution are shown as coauthors in the following list.

The COVID-19 pandemic has highlighted the importance of precision medicine, particularly for COVID-19. The results of this study will be helpful to clinicians in predicting prognosis and outcomes in COVID-19 patients. This simple clustering will be needed to develop precision medicine for COVID-19.
4. Department of General Medicine, Juntendo University Faculty of Medicine and Graduate School of Medicine, Tokyo, Japan.
5. Department of Emergency and Disaster Medicine, Juntendo University Faculty of Medicine and Graduate School of Medicine, Tokyo, Japan.
6. Department of Cardiovascular Biology and Medicine, Juntendo University Faculty of Medicine and Graduate School of Medicine, Tokyo, Japan.
7. Department of Internal Medicine and Rheumatology, Juntendo University Faculty of Medicine and Graduate School of Medicine, Tokyo, Japan.
8. Department of Nephrology, Juntendo University Faculty of Medicine and Graduate School of Medicine, Tokyo, Japan.
9. Atopy Allergy Research Center, Juntendo University Graduate School of Medicine, Tokyo, Japan.
10. Department of Emergency and Critical Care Medicine, Keio University School of Medicine, Tokyo, Japan.
11. Department of Anesthesiology, Keio University School of Medicine, Tokyo, Japan.
12. Department of Laboratory Medicine, Keio University School of Medicine, Tokyo, Japan.
13. Division of Gastroenterology and Hepatology, Department of Medicine, Keio University School of Medicine, Tokyo, Japan.
14. Keio University Health Center, Keio University School of Medicine, Tokyo, Japan.
15. Department of Organoid Medicine, Keio University School of Medicine, Tokyo, Japan.
16. Department of Respiratory Medicine, Osaka Saiseikai Nakatsu Hospital, Osaka, Japan.
17. Department of Infection Control, Osaka Saiseikai Nakatsu Hospital, Osaka, Japan.
18. JCHO (Japan Community Health Care Organization) Saitama Medical Center, Internal Medicine, Saitama, Japan.
19. Department of Respiratory Medicine, Saitama Cardiovascular and Respiratory Center, Kamagaya, Japan.
20. Department of Respiratory Medicine, Tokyo Women's Medical University, Tokyo, Japan.
21. Department of General Medicine, Tokyo Women's Medical University, Tokyo, Japan.
22. Department of Emergency and Critical Care Medicine, Kansai Medical University General Medical Center, Morinohi, Japan.
23. M&D Data Science Center, Tokyo Medical and Dental University, Tokyo, Japan.
24. Clinical Research Center, Tokyo Medical and Dental University Hospital of Medicine, Tokyo, Japan.
25. Department of Medical Informatics, Tokyo Medical and Dental University Hospital of Medicine, Tokyo, Japan.
26. Respiratory Medicine, Tokyo Medical and Dental University, Tokyo, Japan.
27. Clinical Laboratory, Tokyo Medical and Dental University Hospital of Medicine, Tokyo, Japan.
28. Department of Insured Medical Care Management, Tokyo Medical and Dental University Hospital of Medicine, Tokyo, Japan.
29. Fukuijuji Hospital, Kiyose, Japan.
30. Kawasaki Municipal Ida Hospital, Department of Internal Medicine, Kawasaki, Japan.
31. Department of Infectious Diseases, Tosei General Hospital, Seto, Japan.
32. Department of Respiratory Medicine and Allergy, Tosei General Hospital, Seto, Japan.
33. Department of Respiratory Medicine, Kitasato University Kitasato Institute Hospital, Tokyo, Japan.
34. School of Veterinary Medicine, Kitasato University, Toyoda, Japan.
35. Laboratory of Viral Infection I, Department of Infection Control and Immunology, Omura Satoshi Memorial Institute & Graduate School of Infection Control Sciences, Kitasato University, Tokyo, Japan.
36. Department of Respiratory Medicine, Tohoku University Graduate School of Medicine, Sendai, Japan.
37. Department of Infectious Diseases, Tohoku University Graduate School of Medicine, Sendai, Japan.
38. Saiseikai Utsumoiyama Hospital, Utsumoiyama, Japan.
39. Department of Pulmonary Medicine, Saitama City Hospital, Saitama, Japan.
40. Department of Infectious Diseases, Saitama City Hospital, Saitama, Japan.
41. Department of General Thoracic Surgery, Saitama City Hospital, Saitama, Japan.
42. Department of Pulmonary Medicine, Eiju General Hospital, Tokyo, Japan.
43. Division of Infection Control, Eiju General Hospital, Tokyo, Japan.
44. Department of Hematology, Eiju General Hospital, Tokyo, Japan.
45. Department of Respiratory Medicine and Clinical Immunology, Osaka University Graduate School of Medicine, Suita, Japan.
46. Department of Statistical Genetics, Osaka University Graduate School of Medicine, Suita, Japan.
47. Integrated Frontier Research for Medical Science Division, Institute for Open and Transdisciplinary Research Initiatives, Osaka University, Suita, Japan.
48. Department of Otorhinolaryngology-Head and Neck Surgery, Osaka University Graduate School of Medicine, Suita, Japan.
49. Department of Neurosurgery, Osaka University Graduate School of Medicine, Suita, Japan.
50. Department of Head and Neck Surgery, Aichi Cancer Center Hospital, Nagoya, Japan.
51. Division of Infection Control and Prevention, Osaka University Hospital, Suita, Japan.
52. Department of Biomedical Ethics and Public Policy, Osaka University Graduate School of Medicine, Suita, Japan.
53. Department of Otolaryngology and Head and Neck Surgery, Kansai Rosai Hospital, Hyogo, Japan.
54. Department of Immunopathology, Immunology Frontier Research Center (WPI-IFReC), Osaka University, Suita, Japan.
55. The Center for Infectious Disease Education and Research (CIDER), Osaka University, Suita, Japan.
56. Department of Respiratory Medicine, Saiseikai Yokohamashi Nanbu Hospital, Yokohama, Japan.
57. Department of Clinical Laboratory, Saiseikai Yokohamashi Nanbu Hospital, Yokohama, Japan.
58. Internal Medicine, Internal Medicine Center, Showa University Koto Toyosu Hospital, Tokyo, Japan.
59. Internal Medicine, Sano Kosei General Hospital, Sano, Japan.
60. Ishikawa Prefectural Central Hospital, Kanazawa, Japan.
61. Tachikawa Hospital, Tachikawa, Japan.
62. Department of Emergency and Critical Care Medicine, Tokyo Women's Medical University Medical Center East, Tokyo, Japan.
63. Department of Medicine, Tokyo Women's Medical University Medical Center East, Tokyo, Japan.
64. Department of Pediatrics, Tokyo Women's Medical University Medical Center East, Tokyo, Japan.
65. Japan Community Health Care Organization Kanazawa Hospital, Kanazawa, Japan.
66. Department of Respiratory Medicine, Japan Organization of Occupational Health and Safety, Kanto Rosai Hospital, Kawasaki, Japan.
67. Department of General Internal Medicine, Japan Organization of Occupational Health and Safety, Kanto Rosai Hospital, Kawasaki, Japan.
68. Sapporo City General Hospital, Sapporo, Japan.
69. Department of Emergency and Critical Care Medicine, Faculty of Medicine, Fukuoka University, Fukuoka, Japan.
70. Department of Infection Control, Fukuoka University Hospital, Fukuoka, Japan.
71. Kanagawa Cardiovascular and Respiratory Center, Yokohama, Japan.
72. Department of Respiratory Medicine, National Hospital Organization Tokyo Medical Center, Tokyo, Japan.
73. Department of Allergy, National Hospital Organization Tokyo Medical Center, Tokyo, Japan.
74. Department of General Internal Medicine and Infectious Diseases, National Hospital Organization Tokyo Medical Center, Tokyo, Japan.
75. Department of Respiratory Medicine, Toyohashi Municipal Hospital, Toyohashi, Japan.
76. Kiyu Hospital, Yokohama, Japan.
77. Division of Respiratory Medicine, Social Welfare Organization Saiseikai Imperial Gift Foundation, Inc., Saiseikai Kumamoto Hospital, Kumamoto, Japan.
78. KKR Sapporo Medical Center, Department of respiratory medicine, Sapporo, Japan.
79. Division of General Internal Medicine, Department of Internal Medicine, St. Marianna University School of Medicine, Kawasaki, Japan.
80. Department of Emergency and Critical Care Medicine, St. Marianna University School of Medicine, Kawasaki, Japan.
81. Japanese Red Cross Medical Center, Tokyo, Japan.
82. Matsumoto City Hospital, Matsumoto, Japan.
Author contributions

Conceptualization: SO, SC, HN, KM, HK, MI, NH, KF. Data curation: SO, KN, HT, HL, AM, TF, MW, TK. Formal analysis: SO, SC. Methodology: SO, SC, HN. Supervision: SC, HN, KM, HK, MI, NoH, Nai, TU, SU, TI, KA, FS, TY, YM, YN, YM, YS, KM, YO, RK, YK, AK, SI, SM, SO, TK. Visualization: SC, HN. Writing—original draft: SO, SC. Writing—review and editing: SO, SC, HN, KM, HK, MI, NH, Nai, NoH, NaH, TU, SU, TI, KA, FS, TY, YM, YN, YM, YS, KM, YO, RK, YK, AK, SI, SM, SO, TK, KF. All authors read and approved the final manuscript.
Funding
This study was supported by AMED (JP20km0101612, JP20km0108415, JP21km02143, JP21km0405217, JP21km0405217), JST CREST (JPMJCR20H2), MHLW (20CA2054), Takeda Science Foundation, Mitsubishi Foundation, and Bioinformatics Initiative of Osaka University Graduate School of Medicine, Osaka University. Precursory Research for Embryonic Science and Technology (JPMJPR21R7).

Availability of data and materials
The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate
This study was performed in accordance with the Declaration of Helsinki and was approved by the ethics committees of Keio University School of Medicine (20200061) and related research institutions. All adult participants provided written informed consent to participate in this study.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no conflicts of interest.

Author details
1 Division of Pulmonary Medicine, Department of Medicine, Keio University School of Medicine, 35 Shikanomachi, Tokyo 160-8582, Japan. 2 Department of Infectious Diseases, Keio University School of Medicine, Tokyo, Japan. 3 Department of Respiratory Medicine, Juntendo University Faculty of Medicine and Graduate School of Medicine, Tokyo, Japan. 4 Department of Respiratory Medicine, Osaka University Graduate School of Medicine, Osaka, Japan. 5 JCHO (Japan Community Health Care Organization) Saitama Medical Center, Internal Medicine, Saitama, Japan. 6 Department of Respiratory Medicine, Saitama Cardiovascular and Respiratory Center, Kamagaya, Japan. 7 Department of Respiratory Medicine, Tokyo Women's Medical University, Tokyo, Japan. 8 Department of Emergency and Critical Care Medicine, Kansai Medical University General Medical Center, Moriguchi, Japan. 9 Department of Respiratory Medicine, Fukujuji Hospital, Kiyose, Japan. 10 Department of Internal Medicine, Kawasaki Municipal Iida Hospital, Kawasaki, Japan. 11 Department of Infectious Diseases, Tosei General Hospital, Seto, Japan. 12 Department of Respiratory Medicine, Kitasato University Kitasato Institute Hospital, Tokyo, Japan. 13 Department of Respiratory Medicine, Toho University Graduate School of Medicine, Sendai, Japan. 14 Department of Statistical Genetics, Osaka University Graduate School of Medicine, Suita, Japan. 15 Medical Innovation Promotion Center, Tokyo Medical and Dental University, Tokyo, Japan. 16 Department of Surgery, Keio University School of Medicine, Tokyo, Japan. 17 Institute of Research, Tokyo Medical and Dental University, Tokyo, Japan. 18 Division of Health Medical Intelligence, Human Genome Center, The Institute of Medical Science, The University of Tokyo, Tokyo, Japan. 19 M&D Data Science Center, Tokyo Medical and Dental University, Tokyo, Japan. 20 Department of Pathology and Tumor Biology, Kyoto University, Kyoto, Japan. 21 Division of Gastroenterology and Hepatology, Department of Medicine, Keio University School of Medicine, Tokyo, Japan.

Received: 7 March 2022 Accepted: 23 August 2022
Published online: 14 September 2022

References
1. Samudrala PK, Kumar P, Choudhary K, Thakur N, Wadekar GS, Dayaramani R, et al. Virology, pathogenesis, diagnosis and in-line treatment of COVID-19. Eur J Pharmacol. 2020;883:173379.
2. Wiersinga WJ, Rhodes A, Cheng AC, Peacock SJ, Prescott HC. Pathophysiology, transmission, diagnosis, and treatment of coronavirus disease 2019 (COVID-19): a review. JAMA. 2020;324:782–93.
3. Tsai PH, Lai WY, Lin YY, Luo YH, Lin YT, Chen HK, et al. Clinical manifestation and disease progression in COVID-19 infection. J Chin Med Assoc. 2021;84(1):3–8.
4. Pollard CA, Moren MP, Nestor-Kalinoski AL. The COVID-19 pandemic: a global health crisis. Physiol Genomics. 2020;52(1):549–57.
5. Shankar-Hari M, Rubenfeld GD. Population enrichment for critical care trials: phenotypes and differential outcomes. Curr Opin Crit Care. 2019;25:489–97.
6. Rubio-Rivas M, Corbella X, Mora-Luján JM, Loureiro-Amigo J, López Sampaolo A, Yera Bergua C, et al. Predicting clinical outcome with phenotypic clusters in COVID-19 pneumonia: an analysis of 12,066 hospitalized patients from the Spanish registry SEMI-COVID-19. J Clin Med. 2020;9:3488.
7. Ye W, Lu W, Tang Y, Chen G, Li X, Ji C, et al. Identification of COVID-19 clinical phenotypes by principal component analysis-based cluster analysis. Front Med (Lausanne). 2020;7:507614.
8. Luszczek ER, Ingham NE, Karam BS, Proper J, Siegel L, Helgeson ES, et al. Characterizing COVID-19 clinical phenotypes and associated comorbidities and complication profiles. PLoS ONE. 2021;16:e0248956.
9. Dupont T, Caillaux-Zucman S, Fremeaux-Bacchi V, Morin F, Lengliné E, Darmon M, et al. Identification of distinct immunophenotypes in critically ill coronavirus disease 2019 patients. Chest. 2021;159:1884–93.
10. Vasquez CR, Gupta S, Miano TA, Roche M, Hsu J, Yang W, et al. Identification of distinct clinical subphenotypes in critically ill patients with COVID-19. Chest. 2021;160:929–43.
11. Reddy K, Sinha P, O’Kane CM, Gordon AC, Calfée CS, McAuley DF. Subphenotypes in critical care: translation into clinical practice. Lancet Respir Med. 2020;8:631–43.
12. Ishii M, Terai H, Kabata H, Masaki K, Chubachi S, Tateno H, et al. Clinical characteristics of 345 patients with coronavirus disease 2019 in Japan: a multicenter retrospective study. J Infect. 2020;8:1-3-5.
13. Williamson EJ, Walker AJ, Khaskaran K, Bacon S, Bates C, Morton CE, et al. Factors associated with COVID-19-related death using OpenSAFELY Nature. 2020;584:430–6.
14. Ministry of Health Laïk, Situation report. https://www.mhlw.go.jp/stf/covid-19/kokinunoshaseijyoukyou.html.
15. Matsunaga N, Hayakawa K, Terada M, Ohtsu H, Asai Y, Tszuki S, et al. Clinical epidemiology of hospitalized patients with COVID-19 in Japan: report of the COVID-19 REGISTRY JAPAN. Clin Infect Dis. 2020;73:e3677–89.
16. Namkoong H. Japan COVID-19 Task Force: a nation-wide consortium to elucidate host genetics of COVID-19 pandemic in Japan. medRxiv. 2021.
17. Tanaka H, Namkoong H, Lee H, Morita A, Chubachi S, Kabata H, et al. Clinical characteristics ofpatients with coronavirus disease (COVID-19): preliminary baseline report of Japan COVID-19 task force, a nation-wide consortium to investigate host genetics of COVID-19. Int J Infect Dis. 2021;113:74–81.
18. World Health Organization. Novel coronavirus: COVID-19 therapeutic trial synopsis. February 18, novel coronavirus: COVID-19 therapeutic trial synopsis. 2020. https://www.who.int/blueprint/priority-diseases-key-actions/COVID-19_Treatment_Trial_Design_Master_PROTOCOL_Synopsis_Final_18022020.pdf.
19. Zheng Z, Peng F, Xu B, Zhao J, Liu H, Peng J, et al. Risk factors of critical & mortal COVID-19 cases: a systematic literature review and meta-analysis. J Infect. 2020;81:e16–21.
20. Zhou Y, Chi J, Lu W, Wang Y. Obesity and diabetes as high-risk factors for severe coronavirus disease 2019 (COVID-19). Diabetes Metab Res Rev. 2021;37:e3377.
21. Li X, Xu S, Yu M, Wang K, Tao Y, Zhou Y, et al. Risk factors for severity and mortality in adult COVID-19 inpatients in Wuhan. J Allergy Clin Immunol. 2020;146:110–8.
22. Singh AK, Gupta R, Gosh A, Misra A. Diabetes in COVID-19: prevalence, pathophysiology, prognosis and practical considerations. Diabetes Metab Syndr. 2020;14:303–10.
23. Liang W, Guan W, Chen R, Wang W, Li J, Xu K, et al. Cancer patients in SARS-CoV-2 infection: a nationwide analysis in China. Lancet Oncol. 2020;21:335–7.
24. Zhao Q, Meng M, Kumar R, Wu Y, Huang J, Lian N, et al. The impact of COPD and smoking history on the severity of COVID-19: a systematic review and meta-analysis. J Med Virol. 2020;92:1915–21.
25. Kovalic AJ, Satapathy SK, Thuluvath PJ. Prevalence of chronic liver disease in patients with COVID-19 and their clinical outcomes: a systematic review and meta-analysis. Hepatol Int. 2020;14:612–20.

26. McLachlan GJ. Cluster analysis and related techniques in medical research. Stat Methods Med Res. 1992;1:27–48.

27. Ye W, Robbins RT. Cluster analysis of longidorus species (nematoda: longidoridae), a new approach in species identification. J Nematol. 2004;36:207–19.

28. Foster KJ, Jauregui E, Tajudeen B, Bishehsari F, Mahdavinia M. Smell loss is a prognostic factor for lower severity of coronavirus disease 2019. Ann Allergy Asthma Immunol. 2020;125:481–3.

29.Porta-Etessam J, Núñez-Gil IJ, González García N, Fernandez-Perez C, Viana-Llamas MC, Eid CM, et al. COVID-19 anosmia and gustatory symptoms as a prognosis factor: a subanalysis of the HOPE COVID-19 (Health Outcome Predictive Evaluation for COVID-19) registry. Infection. 2021;49:677–84.

30. Liang W, Liang H, Ou L, Chen B, Chen A, Li C, et al. Development and validation of a clinical risk score to predict the occurrence of critical illness in hospitalized patients with COVID-19. JAMA Intern Med. 2020;180:1081–9.

31. Sobhani S, Aryan R, Kalantari E, Soltani S, Malek N, Pirzadeh P, et al. Association between clinical characteristics and laboratory findings with outcome of hospitalized COVID-19 patients: a report from Northeast Iran. Interdiscip Perspect Infect Dis. 2021;2021:5552138.

32. Liao D, Zhou F, Luo L, Xu M, Wang H, Xia J, et al. Haematological characteristics and risk factors in the classification and prognosis evaluation of COVID-19: a retrospective cohort study. Lancet Haematol. 2020;7:e671–8.

33. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet. 2020;395:1054–62.

34. Wu C, Chen X, Cai Y, Xia J, Zhou X, Xu S, et al. Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 pneumonia in Wuhan, China. JAMA Intern Med. 2020;180:934–43.

35. Shi S, Qin M, Shen B, Cai Y, Liu T, Yang F, et al. Association of cardiac injury with mortality in hospitalized patients with COVID-19 in Wuhan, China. JAMA Cardiol. 2020;5:802–10.

36. Gavriatopoulou M, Ntanasis-Stathopoulos I, Korompoki E, Fotiou D, Migkou M, Tzanninis IG, et al. Emerging treatment strategies for COVID-19 infection. Clin Exp Med. 2021;21:167–79.

37. DeMerle K, Angus DC, Seymour CW. Precision medicine for COVID-19: phenotype anarchy or promise realized? JAMA. 2021;325:2041–2.

38. Chauhan K, Partharanitam P, Piani F, Johnson RJ, Urbairi J, Chan L, et al. Prevalence and outcomes associated with hyperuricemia in hospitalized patients with COVID-19. Am J Nephrol. 2022;53:78–86.

39. Deniz C, Selçuk Y, Kubilay I, Ahmed CG, Didar S, Ahmed BG, et al. Uric acid and mortality relationship in covid-19. Acta Medica Mediterranea. 2022;38(1):725–31.

40. Li C, Hsieh MC, Chang SJ. Metabolic syndrome, diabetes, and hyperuricemia. Curr Opin Rheumatol. 2013;25:210–6.

41. Song P, Wang H, Xiao W, Chang X, Wang M, An L. Prevalence and correlates of hyperuricemia in the middle-aged and older adults in China. Sci Rep. 2018;8:4314.

42. Ripabelli G, Sammarco ML, Cannizzaro F, Montano C, Ponziio GV, Tamburro M. A coronavirus outbreak linked to a funeral among a Romani community in Central Italy. Front Med (Lausanne). 2021;8:617264.

43. Figliozzis S, Masci PG, Ahmadi N, Tondi L, Koutli E, Aimo A, et al. Predictors of adverse prognosis in COVID-19: A systematic review and meta-analysis. Eur J Clin Invest. 2020;50:e13362.

44. Fathi N, Rezaei N. Lymphopenia in COVID-19: Therapeutic opportunities. Cell Biol Int. 2020;44:1792–7.

45. Raman R, Patel KJ, Ranjan K. COVID-19: unmasking emerging SARS-CoV-2 variants, vaccines and therapeutic strategies. Biomolecules. 2021;11:993.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.