Archives of Public Health

SYSTEMATIC REVIEW Open Access

Trends in symptom prevalence and sequential onset of SARS-CoV-2 infection from 2020 to 2022 in East and Southeast Asia: a trajectory pattern exploration based on summary data

Tian Song[1†], Jing Guo[1†], Bin Liu[1], Lu Yang[1], Xiangwei Dai[1], Fuqiang Zhang[1], Zhaoyuan Gong[1], Mingzhi Hu[1], Qianzi Che[1*†] and Nannan Shi[1*†]

Abstract

Background The COVID-19 pandemic's diverse symptomatology, driven by variants, underscores the critical need for a comprehensive understanding. Employing stochastic models, our study evaluates symptom sequences across SARS-CoV-2 variants on aggregated data, yielding essential insights for targeted interventions.

Methods We conducted a meta-analysis based on research literature published before December 9, 2022, from PubMed, LitCovid, Google Scholar, and CNKI databases, to investigate the prevalence of COVID-19 symptoms during the acute phase. Registered in PROSPERO (CRD42023402568), we performed random-effects meta-analyses using the R software to estimate pooled prevalence and 95% CI. Based on our findings, we introduced the Stochastic Progression Model and Sequential Pattern Discovery using Equivalence classes (SPADE) algorithm to analyze patterns of symptom progression across different variants.

Results Encompassing a total of 430,100 patients from east and southeast Asia, our results reveal the highest pooled estimate for cough/dry cough across wild-type, Delta, and Omicron variants, with fever (78.18%; 95% CI: 67–89%) being the most prominent symptom for the Alpha variant. Symptoms associated with the Omicron variant primarily manifested in upper respiratory tracts, cardiovascular, and neuropsychiatric systems. Stochastic models indicate early symptoms including dry cough and fever, followed by subsequent development of sleep disorders, fatigue, and more.

†Tian Song and Jing Guo contributed equally to this work and share first authorship.

†Qianzi Che and Nannan Shi contributed equally to this work and share last authorship.

*Correspondence:
Qianzi Che
cheqianzi123@126.com
Nannan Shi
13811839164@vip.126.com

Full list of author information is available at the end of the article

Conclusion Our study underscores the evolving symptomatology across SARS-CoV-2 variants, emphasizing similarities in fever, cough, and fatigue. The Omicron variant presents a distinct profile characterized by milder symptoms yet heightened neuropsychological challenges. Advanced analytical models validate the observed sequential progression of symptoms, reinforcing the consistency of disease trajectory.

Keywords SARS-CoV-2 variants, Disease progression, Symptom trajectories, Stochastic model

Text box 1. Contributions to the literature

- Through the examination of symptom trajectory patterns, our research has elucidated the substantial neuropsychological challenges linked to the Omicron variant of COVID-19, particularly during the acute phase and across a diverse array of symptoms.
- We employed the stochastic models and SPADE algorithm to analyze symptom progression from summary data, offering valuable insights for targeted interventions.
- We conducted a meta-analysis that synthesized the results from studies involving 430,100 COVID-19 patients, ensuring the robustness and reliability of our study findings.

Introduction

The COVID-19 pandemic, caused by the novel strain of coronavirus SARS-CoV-2 coronavirus, emerged in late 2019 in China and rapidly spread worldwide, leading to a global health crisis [1]. Common symptoms associated with the wild-type strain include fever, cough, headache, fatigue, breathing difficulties, and loss of smell and taste, ranging from mild to severe manifestations [2]. In more severe cases, patients may develop complications like pneumonia, acute respiratory distress syndrome, blood clot formation, organ failure, and even mortality [3]. However, the Wild-type virus has undergone significant evolution, giving rise to several Variants of Concerns (VOCs) such as Alpha, Beta, Delta, Gamma, and Omicron variants [4]. Emerging evidence suggests that these variants may be associated with varying prevalence of specific symptoms [5].

Understanding the progression of COVID-19 symptoms holds considerable importance as it offers insights into viral load dynamics, prognostic outcomes, organ tropism, and can guide the development of targeted prevention and treatment strategies [6]. However, comprehensive research in these domains faces obstacles, primarily due to restricted access to patient-level data. Such limitations often arise from protective policies implemented by nations and institutions, concerns about privacy, legal constraints, and the sensitive nature of medical information. Notably, regulations like the US Health Insurance Portability and Accountability Act (HIPAA) exemplify the efforts to safeguard patient data [7]. Even when data is accessible, it tends to be fragmented across multiple sources and may lack regular updates, posing additional challenges. To address this, a pragmatic and effective approach involves harnessing information derived from peer-reviewed studies and scientific literature. These well-vetted publications offer a wealth of reliable, accurate, and pertinent insights.

Efforts have been made to leverage aggregated data for the study of disease characteristics. For instance, Larsen et al. have pioneered a stochastic progression model to predict the sequential order of symptoms in Alzheimer's and other diseases, utilizing patient characteristics from prior studies [8–10]. In light of this, our study aims to adapt and implement a similar model using aggregated data from publications centered around the east and southeast Asian region. Our objective is to investigate whether the sequence of symptoms remains consistent across different SARS-CoV-2 variants. To achieve this, we initially extracted descriptive data from academic databases and conducted a meta-analysis to assess symptom distributions. Subsequently, we introduced a stochastic progression model to analyze symptom sequences across various variants. To bolster the model's robustness and address inherent logical challenges, we incorporated a frequent sequence mining technique known as Sequential PAttern Discovery using Equivalence classes (SPADE) [11].

Method

Registration and protocol

This meta-analysis was pre-registered in the PROSPERO database (CRD42023402568). The initial literature search was independently conducted by three investigators (CQZ, GJ, and YL).

Search strategy

We conducted a systematic search of both Chinese and English literature, focusing on PubMed, Lit Covid from NLM, Google Scholar, and CNKI databases. Our search included studies from east and southeast Asia area, specifically including China, Japan, South Korea, Singapore, India, up to December 9, 2022.

Our search strategy included a wide range of terms associated with COVID-19 and relevant countries. Keywords used in our search included 'SARS-CoV-2 infection', 'SARS coronavirus 2 infection', '2019 novel coronavirus disease', '2019 novel coronavirus infection', '2019-nCoV disease', '2019-nCoV infection', 'COVID-19 Pandemic', 'COVID-19 Pandemics', 'COVID-19 Virus Disease', 'COVID-19 Virus Infection', 'COVID19', '2019 Coronavirus Disease', 'Coronavirus Disease-19', 'SARS-CoV-2

strains', 'alpha variants', 'Delta variants', 'Omicron variant', 'symptoms', 'alpha', 'beta', 'Delta', 'Omicron', 'case series', 'cross-sectional study', 'case-control study', 'cohort study', 'registry study', 'case-cohort study', 'case-crossover study', 'randomized control trial', and 'pragmatic clinical trial'.

Inclusion criteria for literature

We used the PICO model to assess the eligibility of studies for our meta-analysis. This included (P) participants: patients with COVID-19; (I) intervention: infection with different SARS-CoV-2 strains; (C) comparison: not applicable; and (O) outcome measures: clinical characteristics of COVID-19, including general, respiratory, cardiovascular, gastrointestinal, neurological, and psychiatric symptoms.

Our study population included laboratory-confirmed cases of COVID-19, patients diagnosed with COVID-19 by hospital diagnosis codes, and clinically diagnosed patients with COVID-19. We included original peer-reviewed clinical research studies published in either Chinese or English that provided details on the clinical characteristics and relevant statistics of COVID-19. Studies were excluded if they only reported patient imaging without corresponding symptom or laboratory data, contained data that overlapped with data from previously included studies, or had a sample size of less than 50 individuals. Search results were cross-checked, and discrepancies were resolved by discussion.

Data extraction

One of the main variables of interest in our study was the type of variant, which included the wild-type strain, Alpha, Delta, and Omicron variants. If a Chinese study did not specify a particular variant but was completed before September 2021, we categorized it as the wild-type strain using data from the 2019 Novel Coronavirus Resource (2019nCoVR) [12].

From the qualifying studies, we extracted several key variables, particularly the number of cases presenting with various COVID-19 symptoms. These symptoms included: fever, fatigue, cough/dry cough, dyspnea (including respiratory difficulty, respiratory distress, breathlessness, shortness of breath, gasp, wheeze), hypoxemia, chest tightness, chest pain, palpitations, arthralgia, myalgia, cognitive impairment, sleep disorders, headache, dizziness, nasal congestion, rhinorrhea, tinnitus, otalgia, sore throat, anosmia, ageusia, diarrhea, vomiting, abdominal pain, anorexia, decreased appetite, depression, anxiety, and rash. These symptoms were then further categorized as respiratory, cardiovascular, gastrointestinal, neurological, or psychiatric symptoms.

The selected articles underwent manual interpretation and were organized based on a predetermined template. Discrepancies or uncertainties in the extracted data were resolved by consensus among the three reviewers: CQZ, GJ, and YL.

Meta-analysis

We performed a meta-analysis by pooling the results of all the included trials. For clinical trials with more than one group, we combined them into one large single-arm trial. We used proportions and proportion ratios as effect size measures, which allowed us to compare specific symptom proportions across trials. A detailed flowchart of the trials is shown in Fig. 1. Specifically, "Not available for information extraction" refers to studies where the reported symptoms were influenced by where patients' information was subject to at least one treatment or intervention, thus not reflecting the normal prevalence of COVID-19 symptoms. In addition, "Not meet statistical data requirements" means that the studies did not report symptom prevalence or reported symptoms in a manner that could not be converted to prevalence. Details of included publications could be found in supplementary Table S5.

The analysis was conducted using a meta-analytic random effects model with maximum likelihood estimation. This probabilistic method estimates model parameters by optimizing the likelihood function. Model outputs included the overall estimated effect size, heterogeneity statistics, test results and confidence intervals. In addition, we performed a sub-group analysis by stratifying studies or groups based on the specific COVID-19 variants they investigated. We used Cochran's Q heterogeneity test and I^2 statistics to assess statistical heterogeneity. The Cochrane Risk of Bias Tool and the Newcastle-Ottawa Scale were used to assess the quality of the included studies and the potential risk of bias.

All data analyses were performed with R [13], version 4.2.0, using the 'metafor' package [14].

Stochastic progression model

We executed a Stochastic Progression Model, drawing inspiration from prior studies and modifying it to align with the specificity of our research [9, 15]. Given the rising number of symptoms, it becomes impractical to precisely replicate the methodology of the original study. Hence, we employed greedy algorithms to determine the most probable sequence of symptom progression instead of simulating the entire population using an enumeration method. The distribution rate of these symptoms is shaped by the prevalence rates sourced from our meta-analysis.

The outcome of this method provides a predicted sequence of the most probable symptom progression. This result is then evaluated using a confusion matrix. Additionally, we computed transition probabilities between every pair of symptoms. These

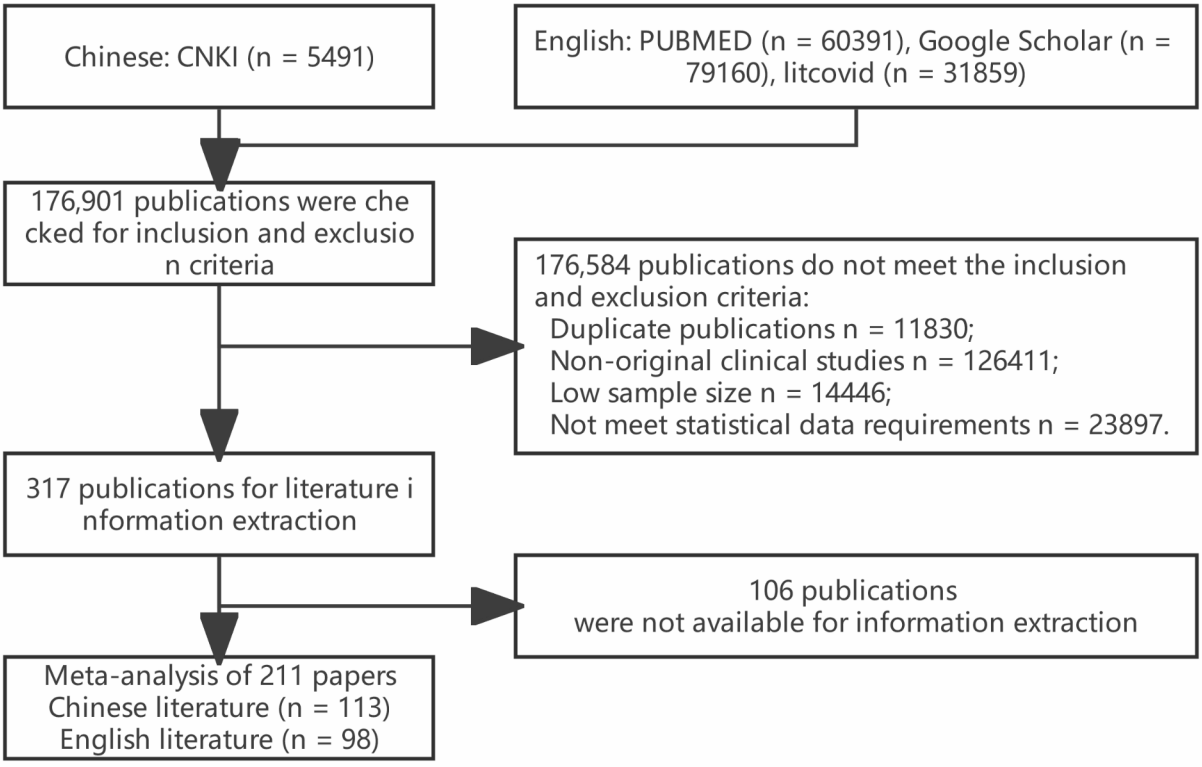

Fig. 1 Flowchart illustrating the literature search process for the main variants of COVID-19

probabilities indicate the likelihood of one symptom leading to another. Furthermore, we measured the deviation between our simulated results and mathematical expectations as an error. To ensure robustness, we simulated a population of 10 million individuals, providing an adequate sample size even for infrequently occurring symptoms. For the visualization of the sub-group analysis concerning COVID-19 variants, we displayed only the predominant branch of symptom evolution for each variant, due to the complexities of showcasing a directed acyclic graph comprising 29 symptoms.

SPADE algorithm

Our model assumes that symptoms with a higher prevalence rate will appear earlier in the progression of symptom development. This is predicated on the notion the sequence of symptoms maybe halts further symptom development at any point as the success of treatment. Hence, early symptoms should have a higher prevalence in cross-sectional panel data.

Therefore, we introduced Sequential PAttern Discovery using Equivalence classes (SPADE) algorithm to identify frequent sequences of two symptoms in the sequence database, a parent symptom and a child symptom [16]. It then moves on to form sequences composed of two or more parent symptoms. All sequences identified must exceed our minimum support threshold (minsup) to be classified as valid. The process iterates, generating and scoring sequences until no more common sequences can be detected. As all higher-level sequences are built on lower level sequences, the sequence with more than one parent symptom could be broken down into a chain of symptoms, making it easier to interpret and visualize. In our study, SPADE uses depth-first search methods to discover new sequences. The confidence value was calculated by dividing the frequency of the whole sequence by the frequency of the sequence without symptoms. Lift was calculated by dividing the frequency of the whole sequence by the product of the frequency of the parent symptoms and the offspring symptoms. Rules with a lift value greater than one were selected.

To visualize the rules identified by the algorithm, we used a Hasse diagram to represent the partial order relationship between symptoms provides an intuitive illustration of the sequence of symptom development. Each node in the diagram represented a symptom, while the edges indicated the number of rules based on those edges. The arrangement of nodes and edges represented the hierarchical structure of the sequences, effectively highlighting the more frequent sequences or clusters of symptoms. In addition, the Hasse diagram visually delineated the dependencies and relationships between

different rules or sequences, making it easier to interpret the complex results generated by the SPADE algorithm. And the level of symptoms is assumed by found rules to better understand the sequences.

Results

Our initial search had returned a total of 176,901 articles, sourced from CNKI (5,491 articles), PubMed (60,391 articles), Google Scholar (79,160 articles), and Lit Covid (31,859 articles). After the removal of duplicates, which amounted to 11,830 articles, we had screened the remaining 165,071 articles. Of these, 126,411 articles were excluded as they were not clinical research articles. Further, 1,446 articles were excluded due to insufficient sample size, and 23,897 articles were disregarded as they did not provide descriptive data related to symptoms.

After this screening process, 317 articles were left for further analysis. From this set, 89 articles were excluded as they did not form an appropriate analysis dataset. After a thorough review of full-text forms, a total of 210 articles were included in our meta-analysis (refer to Table 1 for details). Based on the Newcastle-Ottawa Scale assessment and Risk of Bias tool from Cochrane, we found the majority of the studies are assessed as having good quality (green) across most domains. A few studies are rated as fair quality in certain domains, particularly in comparability (yellow). Very few studies are rated as poor quality (red).

The RoB figure indicates that while most RCTs included in the meta-analysis exhibit a low risk of bias, there are concerns in areas such as allocation concealment, blinding, and selective reporting for some studies. The NOS figure shows that the majority of non-RCTs included are of good quality, with some studies having fair quality in terms of comparability. Detailed results are included in supplementary materials Table S2, S3, S4, and S5.

Results of the meta-analysis

The meta-analysis of data from 430,100 patients showed the prevalence rates for different symptoms. Cough/dry cough was the most common symptom, reported by 55.48% of patients (40,259/83,298), followed by fever (48.08%, 78,936/250,986), hypoxemia (41.16%, 7,702/22,023), sore throat (25.16%, 8,500/47,557), and fatigue (24.26%, 78,729/170,787). Less common symptoms ranged from myalgia at 15.95% to rash at 1.53%. A comprehensive summary of all symptoms and their respective prevalence rates from the meta-analysis present in supplementary Table S1.

Table 1 Distribution of viral variants included in the literature

Virus Variant	Number of Articles	Sample Size (%)
Wild-type strain	71	25,768 (5.99%)
Alpha	1	30 (0.01%)
Delta	45	17,652 (4.10%)
Omicron	35	315,146 (73.27%)
Not described/other	58	71,504 (16.62%)
Total	210	430,100 (100%)

COVID-19 symptom prevalence varied distinctly across different strains. Core symptoms like cough/dry cough (Wild-type: 61.46%, Alpha: 61.82%, Delta: 51.09%, Omicron: 49.96%), fever (Wild-type: 59.68%, Alpha: 78.18%, Delta: 40.58%, Omicron: 35.39%), and fatigue (Wild-type: 31.03%, Alpha: 67.43%, Delta: 15.32%, Omicron: 24.5%) consistently manifested across strains, serving as primary indicators.

Simultaneously, each strain exhibited unique symptom profiles. The Wild-type strain prominently displayed hypoxemia at 45.48%, a symptom that was not observed in other strains. In contrast, the Omicron variant uniquely presented with anxiety (12.53%) and cognitive impairment (18.27%). Notably, the prevalence of fever, which peaked in the Alpha variant at 78.18%, decreased markedly by the time of the Omicron variant, reaching 35.39%. Meanwhile, the prevalence of nasal congestion was notably higher in the Omicron at 24.24%, distinguishing it from its lower rate of 5.51% in the wild-type strain. The high prevalence of fatigue in the Alpha variant (67.43%) saw a sharp decline in the Delta variant (15.32%), but it was still relatively higher in the Omicron variant (24.5%). Dyspnea, which was notably prevalent in the Alpha at 70.91%, also saw a significant reduction in both the Delta (8.57%) and Omicron strains (21.44%).

Additionally, the wide confidence intervals for certain symptoms, such as hypoxemia in the wild-type strain and palpitations in the Omicron, hinted at potential variability in the reported prevalence across different studies. A detailed table presenting the prevalence of symptoms, stratified by variants, can be found in Table 2. A visualization comparing the prevalence of each symptom is depicted in Fig. 2 using a radar chart.

Symptom sequence analysis

Stochastic progression model

The most likely orders of symptoms in the wild-type strain appeared to begin with a dry cough, followed by a range of other symptoms, including fever, fatigue, and loss of appetite. The Alpha strain had a distinct progression, beginning with fever, which contrasted with the wild-type strain. Additionally, the appearance of loss of smell in its sequence was unique to this variant. The Delta variant shared the wild-type strain's initial symptom of a dry cough. However, symptoms such as sleep disturbance and sore throat appeared relatively early compared to the other variants. The Gamma variant's first noticeable symptom was fever. This variant had a unique sequence in which cognitive dysfunction and anxiety appeared later in the progression, a pattern not

Table 2 Symptoms and prevalence in all patients included in the study

Symptoms	Symptom Prevalence Rate (95% CI)			
	Wild-type strain	Alpha Variant	Delta Variant	Omicron Variant
Cough/dry cough	61.46% (57–66%)	61.82% (49–75%)	51.09% (46–56%)	49.96% (44–56%)
Fever	59.68% (53–66%)	78.18% (67–89%)	40.58% (34–47%)	35.39% (29–42%)
Hypoxemia	45.48% (12–79%)	NA	NA	NA
Fatigue	31.03% (26–36%)	67.43% (58–77%)	15.32% (12–19%)	24.5% (18–31%)
Anorexia	28.01% (16–40%)	NA	2.79% (1–5%)	NA
Decreased appetite	27.98% (20–36%)	53.33% (35–71%)	7.05% (5–9%)	20.52% (12–29%)
Dyspnea	25.81% (21–30%)	70.91% (59–83%)	8.57% (3–14%)	21.44% (9–34%)
Sleep disorders	24.86% (19–31%)	40% (22–58%)	26.98% (22–32%)	12.71% (7–19%)
Chest tightness	24.05% (18–30%)	36.67% (19–54%)	10.1% (4–16%)	8.02% (4–12%)
Sore throat	17.11% (13–21%)	39.64% (8–71%)	26.72% (23–31%)	30.38% (26–35%)
Myalgia	15.15% (12–19%)	30.08% (16–44%)	16.62% (12–21%)	14.3% (7–21%)
Diarrhea	11.62% (9–14%)	28.2% (19–38%)	6.7% (5–8%)	4.13% (2–6%)
Rhinorrhea	8.34% (5–12%)	NA	17.74% (15–21%)	14.5% (10–19%)
Anosmia	8.23% (2–15%)	50% (32–68%)	12.75% (11–15%)	13.34% (4–23%)
Headache	7.32% (5–9%)	20% (9–31%)	19.05% (15–23%)	13.77% (10–18%)
Ageusia	6.28% (1–11%)	NA	11.07% (9–13%)	16.25% (5–27%)
Nasal congestion	5.51% (3–8%)	3.64% (0–9%)	16.72% (13–20%)	24.24% (13–35%)
Chest pain	5.41% (2–9%)	10% (0–21%)	3.26% (2–5%)	13.38% (1–26%)
Dizziness	5.33% (3–7%)	NA	2.43% (1–3%)	15.34% (9–22%)
Vomiting	5.1% (4–7%)	20% (9–31%)	4.45% (3–6%)	4.85% (3–7%)
Palpitations	3.69% (0–7%)	NA	NA	16.78% (0–37%)
Abdominal pain	2.8% (1–4%)	NA	1.75% (1–2%)	6.15% (3–9%)
Arthralgia	1.81% (0–4%)	NA	6.92% (4–10%)	15.82% (0–34%)
Rash	0.09% (0–0%)	NA	1.23% (0–2%)	2.09% (0–4%)
Anxiety	NA	NA	NA	12.53% (4–21%)
Cognitive impairment	NA	NA	NA	18.27% (16–21%)

seen in the other variants. The progression of the Omicron variant was similar to the Wild-type and Delta variants, beginning with a dry cough. However, compared to the Gamma variant, the onset of cognitive dysfunction occurred relatively early, while anxiety surfaced later in the progression.

Each variant had a unique sequence of symptom transitions. In particular, cognitive dysfunction and anxiety were observed only in the Gamma and Omicron variants, suggesting potential differences in how these variants interacted with the nervous system. Figure 3 shows the top five symptoms of each simulation, and a comprehensive Figure S1 can be found in the supplementary material.

Our modifications compromised the accuracy of transition probability and error. Therefore, we utilized a simplified confusion matrix to present and analysed the performance of the sequences we predicted using the Stochastic Progression Model. Table 3 displayed the number of correct ratios.

SPADE analysis

Only the wild-type strain, Delta, and Omicron variants subsets were examined and analyzed with the SPADE algorithm [11]. All three analyses were set to support at 0.1 and reported by lift value greater than one. Rules with lift values less than one were excluded from the visualization. Patients infected with the wild-type strain primarily began with a fever, followed by a cough. Post these initial symptoms, individuals might have experienced muscle pain, headache, fatigue, and chest tightness. Interestingly, both cough and fatigue could have potentially increased the likelihood of chest tightness. In regard to the Delta strain, fever led to a range of symptoms including fatigue, cough, runny nose, sore throat, difficulty breathing, and headache. Uniquely, a cough in this strain could contribute to the onset of fatigue, loss of smell and taste, runny nose, and sore throat. For Omicron, fever led to cough and headache. A cough in this case could also cause headache, sore throat, and fatigue. A special situation of the Omicron strain was that a sore throat could increase the likelihood of nasal congestion. These results are summarized in Fig. 4 and Tables S2, S3, and S4.

When compared with the results of the stochastic progression model, all SPADE symptom sequences were confirmed for all variants except fever and cough/dry cough. About 20–30% of the symptoms were found to be dependent on their parent symptoms. Fever and cough/dry cough could have been considered interchangeable due to their common, similar offspring symptoms, and

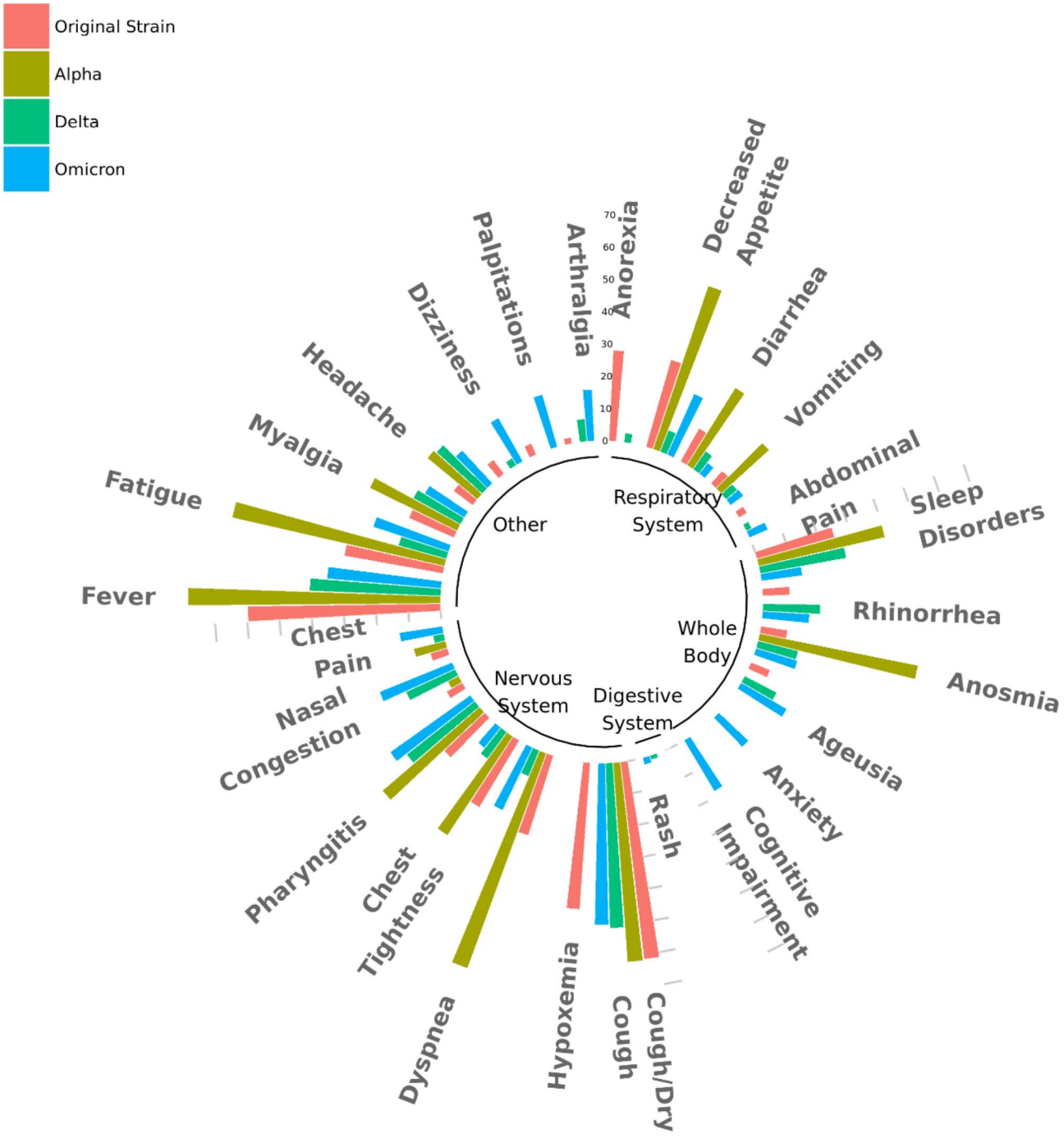

Fig. 2 Comparison of the incidence of some symptoms in different COVID-19 variants

the relationship between them was evident in both high and low lift values. The wild-type strain tended to have more classic flu-like symptoms, with fever and cough followed by myalgias and headaches. Additionally, coughing and fatigue might have increased the likelihood of chest tightness. The Delta strain was characterized by a unique association between cough and loss of smell and taste. It also had a more extensive set of symptoms associated with fever, including difficulty breathing and a runny nose. The Omicron strain was unique in that a sore throat could increase the likelihood of nasal congestion.

Discussion

The emergence of unique variants of SARS-CoV-2 suggests that symptoms may differ due to changes in virulence among the variants. In this study, we

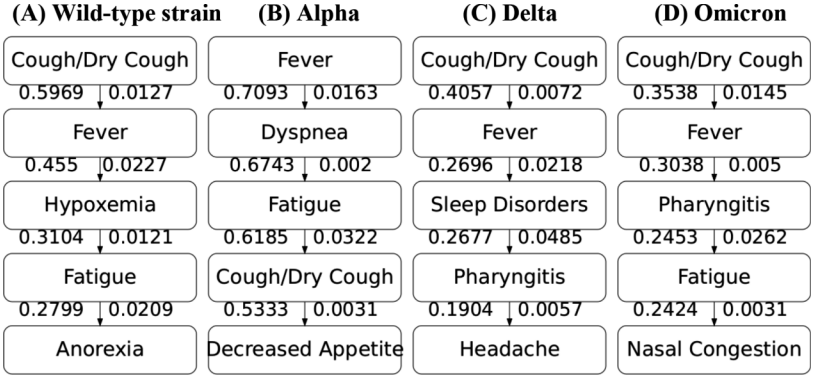

Fig. 3 Top 5 possible order of symptoms in different variants based on Stochastic Progression Model

Table 3 Confusion matrix showing the correct ratio of symptom prediction for each variant

Virus Variant	Recall/precision rate
Wild-type strain	0.8053830
Alpha	0.8641975
Delta	0.7926456
Omicron	0.7652068

comprehensively assessed the symptoms in the acute phase caused by different SARS-CoV-2 strains separately and in conjunction with other symptoms to supplement the limited evidence. Overall, fever, cough, and fatigue were common symptoms across wild-type strain, Alpha, Delta, and Omicron variants. As the virus iterated, the patients' symptoms involved more organs and systems. Symptoms in Omicron variant patients were concentrated in the upper respiratory tracts, followed by cardiovascular and neuropsychiatric symptoms (such as brain fog, lack of concentration, or memory problems). Utilizing Stochastic Progression Model and SPADE algorithm to analyze the disease transmission patterns of each variant, we found that the initial symptoms

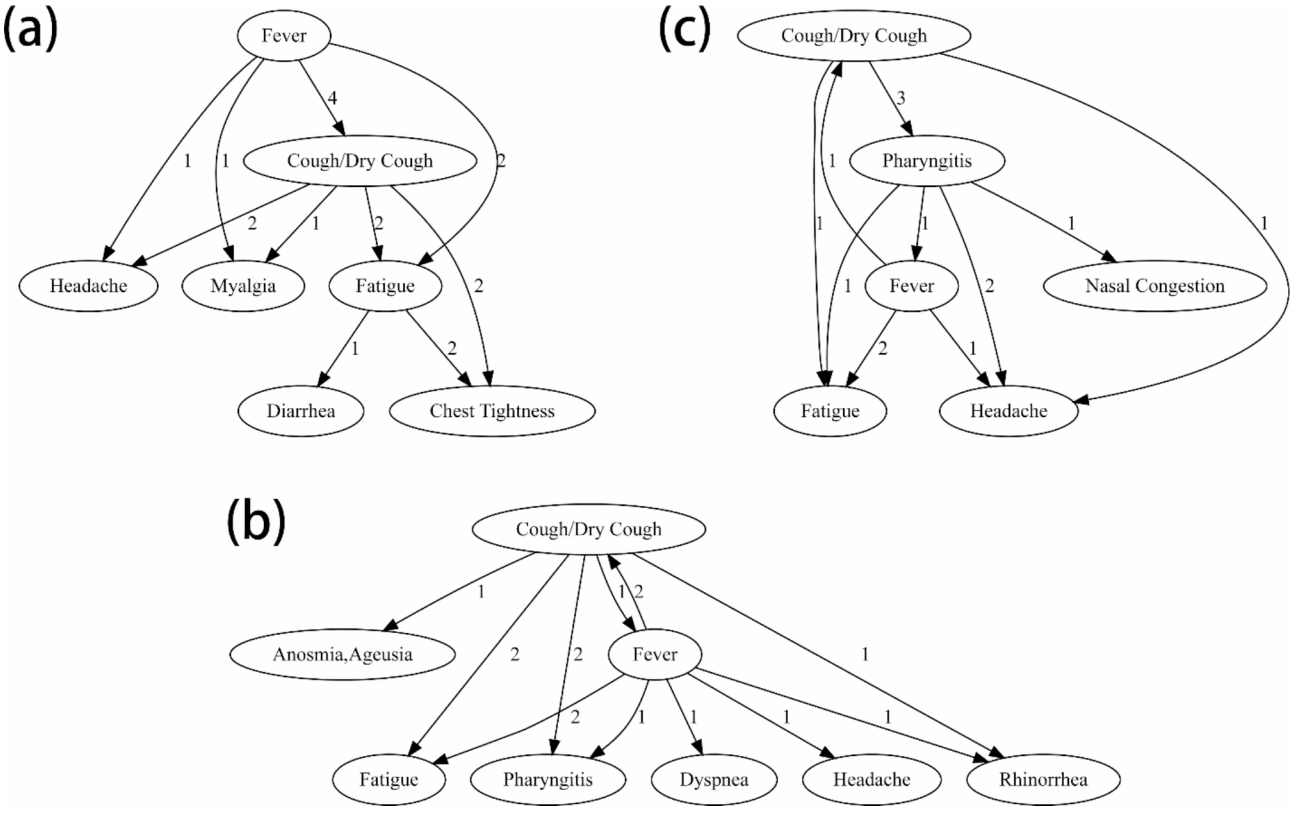

Fig. 4 Results from the SPADE Algorithm on Different COVID-19 Strains with 0.1 and Above Support and Confidence. (**a**) Rules for Wild-type strain; (**b**) Rules for Delta variant; (**c**) Rules for Omicron variant

were dry cough and fever. Subsequently, sleep disorder, decreased appetite, fatigue, and dyspnea occurred. The symptoms then gradually diversified, including anorexia, chest tightness, sore throat, diarrhea, nausea, cognitive impairment, and loss of smell and taste. Understanding the specific symptom progression patterns for different COVID-19 variants can enhance early detection, diagnosis, and treatment protocols. This knowledge allows for the development of tailored treatment plans and proactive symptom management, ultimately improving patient outcomes. Additionally, public health campaigns can educate the public about common symptoms and progression patterns, promoting early self-reporting and timely medical consultation.

In general, the wild-type strain and Alpha strain demonstrated higher severity. For patients infected with the wild-type strain, fever was the most common initial symptom, followed by cough, fatigue, and dyspnea. Our findings align with the results of other researchers who studied disease progression for COVID-19 using Internet search patterns in 32 countries across six continents from January 1 through April 20, 2020 [17, 18]. They reported that initial symptoms of fever, dry cough, sore throat, and chills were followed by shortness of breath, on average, 5.22 days (range 3.30–7.14) after initial symptom onset, which corroborates with the clinical course documented in medical literature [17, 19]. Besides the common symptoms, patients infected with the wild-type strain are also inclined to experience fatigue, loss of appetite, sleep difficulty, chest stuffiness, and diarrhea.

The Delta variant exhibited a relatively lower severity, and its symptom distribution appeared to be more evenly spread among patients. Interestingly, the pooled estimate of gastrointestinal symptoms in patients infected with the Delta variant was lower compared to those infected with other strains. A nationwide study encompassing > 99% of all claims in Japan reported that despite the Delta variant being associated with higher virulence than the Alpha variant, the overall severity appeared to have declined [5]. Several factors might have contributed to this change in the clinical profile of COVID-19 patients. Notably, the increased vaccination rate in east and southeast Asia emerged as a crucial factor [5]. The targeted vaccination efforts, with priority given to seniors, were implemented and expedited during the peak of the Delta-predominant wave in countries such as China, Japan, Korea, and Singapore [5, 20–22]. Furthermore, treatment strategies involving dexamethasone, baricitinib, and remdesivir were established before the Delta-predominant wave [23–26].

Our results reveal that although the symptoms of the Omicron infection are less severe than those of other strains, a higher number of patients may experience mental and neuropsychological health problems during the acute infection phase. Consistent with real-time prospective data, dyspnea and loss of the sense of smell were no longer considered core symptoms [27, 28]. However, non-specific neurological symptoms such as fatigue/weakness, myalgia, headaches, and dizziness were more prevalent in Omicron cases. The etiology of neuropsychiatric symptoms in COVID-19 patients is complex and multifactorial. Neuropsychiatric deficits during the acute infection may be associated with SARS-CoV-2-induced dysregulation of multiple organ systems, direct invasion of the central nervous system, cerebrovascular disease (including hypercoagulation), physiological compromise (hypoxia), side effects of medications, and social aspects, including anxiety about life risks and economic concerns [29–32].

Larsen et al. have utilized the Stochastic Progression Model to predict the order of symptoms using data from both the initial COVID outbreaks in China and the USA [15]. The model assumes that the occurrence of a symptom depends solely on the previous symptom, derived from the assumption of symptom independence. However, while some results in our model align with current publications, they still require further investigation and justification. Therefore, we employed the SPADE algorithm, which shares the same assumption but considers symptoms as dependent or interacting with each other, to validate and enhance our model. Both algorithms assume that symptoms with higher prevalence rates tend to occur earlier in the disease progression and that their sequence is linearly related to their prevalence. However, it is important to recognize that this may not hold universally true, as individual differences in factors such as age, health status, and immune response could lead to variations in symptom onset and progression [33].

Our study is subject to certain limitations. Firstly, our focus on research literature from China, Japan, South Korea, Singapore, and India may limit the generalizability of our findings to other regions. While these countries provide valuable insights into the early stages of the pandemic and diverse strategies for containment, variations in healthcare resources, cultural factors, and population demographics among different regions worldwide may influence symptom patterns differently. Therefore, caution should be exercised when extrapolating our results to other geographical areas. Secondly, the availability and quality of data in the selected countries could impact the comprehensiveness of our analysis. Variability in testing rates, reporting practices, and data accessibility could introduce biases and affect the accuracy of symptom prevalence estimates. Additionally, our study relies on the accuracy and consistency of the published literature, which could be influenced by publication biases and variations in reporting standards. Despite these limitations, our research provides valuable insights into the symptom

progression of COVID-19 across different SARS-CoV-2 variants within the studied east and southeast Asian countries, contributing to our understanding of the evolving nature of the disease.

In conclusion, our comprehensive investigation into the symptomatology of various SARS-CoV-2 variants underscores the dynamic nature of COVID-19 and its evolving clinical presentation. Consistent symptoms such as fever, cough, and fatigue transcend different strains, while the progression of the virus is marked by an expanding range of manifestations. The Omicron variant notably introduces a unique pattern, characterized by milder symptoms yet an increased prevalence of mental and neuropsychological challenges. Our employment of advanced analytical models like the Stochastic Progression Model and SPADE algorithm sheds light on the sequential emergence of symptoms and validates previously observed patterns, reinforcing the consistency of disease progression.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13690-024-01357-7.

Supplementary Material 1

Author contributions

Tian Song and Qianzi Che conceptualized and initiated the study. Tian Song developed the methodology and performed the statistical analysis. Bin Liu, Lu Yang, Xiangwei Dai and Fuqiang Zhang were responsible for data collection and validation. Tian Song, Qianzi Che and Bin Liu screened the literature and selected relevant studies for inclusion. Lu Yang, Xiangwei Dai, Zhaoyuan Gong, Mingzhi Hu, and Bin Liu conducted the risk of bias assessments and quality control. Jing Guo and Nannan Shi provided critical feedback and revised the manuscript for important intellectual content. Tian Song, Jing Guo and Qianzi Che wrote the initial draft of the manuscript, with significant contributions and revisions from all authors. Qianzi Che and Nannan Shi are the guarantors of this manuscript, accepting full responsibility for the work and the conduct of the study. They had access to all the data and controlled the decision to publish. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Funding

This research was financially supported by the China Academy of Chinese Medical Sciences Innovation Fund (No.CI2022C004-01), the Natural Science Foundation of Beijing, China (No. 7244507), and the 2022 Qi Huang Young Scholar programme of the National Administration of Traditional Chinese Medicine (Z0841). The funders had no role in the design, execution, interpretation, or decision to publish the results of this study.

Data availability

All of the data extracted for this review are included in the manuscript and associated supplementary files. Extracted data are available upon request to the corresponding author.

Declarations

Ethics approval and consent to participate

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]Institute of Basic Research in Clinical Medicine, China Academy of Chinese Medical Sciences, No. 16 Nanxiaojie, Dongzhimennei Ave, Dongcheng District, Beijing 100700, P.R. China

Received: 3 January 2024 / Accepted: 10 August 2024
Published online: 15 August 2024

References

1. Coronavirus disease (COVID-19). 2023.https://www.who.int/news-room/questions-and-answers/item/coronavirus-disease-covid-19 (accessed 8 Jun2023).
2. Whitaker M, Elliott J, Bodinier B et al. Variant-specific symptoms of COVID-19 in a study of 1,542,510 adults in England. Nature Communications 2022 13:1 2022 13: 1–10.
3. Tsai PH, Lai WY, Lin YY, et al. Clinical manifestation and disease progression in COVID-19 infection. J Chin Med Association. 2021;84:3–8.
4. Carabelli AM, Peacock TP, Thorne LG, et al. SARS-CoV-2 variant biology: immune escape, transmission and fitness. Nat Rev Microbiol. 2023;21:162.
5. Miyashita K, Hozumi H, Furuhashi K, et al. Changes in the characteristics and outcomes of COVID-19 patients from the early pandemic to the delta variant epidemic: a nationwide population-based study. Emerg Microbes Infect. 2023;12. https://doi.org/10.1080/22221751.2022.2155250.
6. Bature F, Guinn BA, Pang D, Pappas Y. Signs and symptoms preceding the diagnosis of Alzheimer's disease: a systematic scoping review of literature from 1937 to 2016. BMJ Open. 2017;7. https://doi.org/10.1136/BMJOPEN-2016-015746.
7. HIPAA Home | HHS.gov. https://www.hhs.gov/hipaa/index.html (accessed 10 Jun 2023).
8. Firth NC, Primativo S, Brotherhood E, et al. Sequences of cognitive decline in typical Alzheimer's disease and posterior cortical atrophy estimated using a novel event-based model of disease progression. Alzheimer's Dement. 2020;16:965.
9. Larsen JR, Martin MR, Martin JD, Kuhn P, Hicks JB. Modeling the onset of symptoms of COVID-19. Front Public Health. 2020;8:473.
10. Oxtoby NP, Leyland LA, Aksman LM, et al. Sequence of clinical and neurodegeneration events in Parkinson's disease progression. Brain. 2021;144:975.
11. Zaki MJ. SPADE: An Efficient Algorithm for Mining Frequent Sequences. 2001; 42: 31–60.
12. Zhao WM, Song SH, Chen ML, Zou D, Bao YM. The 2019 novel coronavirus resource. Hereditas (Beijing). 2020;42:212–21.
13. R Core Team. R: A Language and Environment for Statistical Computing. 2023.https://www.R-project.org/ (accessed 11 Jun2023).
14. Viechtbauer W. Conducting Meta-analyses in R with the metafor Package. J Stat Softw. 2010;36. https://doi.org/10.18637/jss.v036.i03.
15. Larsen JR, Martin MR, Martin JD, Hicks JB, Kuhn P. Modeling the onset of symptoms of COVID-19: effects of SARS-CoV-2 variant. PLoS Comput Biol. 2021;17. https://doi.org/10.1371/JOURNAL.PCBI.1009629.
16. Estiri H, Vasey S, Murphy SN. Transitive Sequential Pattern Mining for Discrete Clinical Data. 2020, pp 414–424.
17. Wang D, Hu B, Hu C, et al. Clinical characteristics of 138 hospitalized patients with 2019 Novel Coronavirus-infected pneumonia in Wuhan, China. JAMA. 2020;323:1061–9.
18. Lu T, Reis BY. Internet search patterns reveal clinical course of COVID-19 disease progression and pandemic spread across 32 countries. NPJ Digit Med. 2021;4:22.
19. Cohen PA, Hall LE, John JN, Rapoport AB. The early natural history of SARS-CoV-2 infection. Mayo Clin Proc. 2020;95:1124–6.
20. Chun JY, Jeong H, Kim Y. Age-varying susceptibility to the Delta variant (B.1.617.2) of SARS-CoV-2. JAMA Netw Open. 2022;5:e223064.
21. Wang G, Yao Y, Wang Y, et al. Determinants of COVID-19 vaccination status and hesitancy among older adults in China. Nat Med. 2023;29:623–31.
22. Ng OT, Koh V, Chiew CJ, et al. Impact of Delta variant and vaccination on SARS-CoV-2 secondary attack rate among Household Close contacts. Lancet Reg Health West Pac. 2021;17:100299.
23. Dexamethasone in Hospitalized Patients with Covid-19. N Engl J Med. 2021;384:693–704.
24. Kalil AC, Patterson TF, Mehta AK, et al. Baricitinib plus Remdesivir for hospitalized adults with Covid-19. N Engl J Med. 2021;384:795–807.

25. Beigel JH, Tomashek KM, Dodd LE, et al. Remdesivir for the treatment of Covid-19 — final report. N Engl J Med. 2020;383:1813–26.
26. Spinner CD, Gottlieb RL, Criner GJ, et al. Effect of Remdesivir vs Standard Care on Clinical Status at 11 days in patients with moderate COVID-19. JAMA. 2020;324:1048.
27. Bouzid D, Visseaux B, Kassasseya C, et al. Comparison of patients infected with Delta Versus Omicron COVID-19 variants presenting to Paris Emergency Departments. Ann Intern Med. 2022;175:831–7.
28. 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.
29. Rabaan AA, Smajlović S, Tombuloglu H, et al. SARS-CoV-2 infection and multi-organ system damage: a review. Bosn J Basic Med Sci. 2022. https://doi.org/10.17305/bjbms.2022.7762.
30. Spudich S, Nath A. Nervous system consequences of COVID-19. Sci (1979). 2022;375:267–9.
31. Baldini T, Asioli GM, Romoli M, et al. Cerebral venous thrombosis and severe acute respiratory syndrome coronavirus-2 infection: a systematic review and meta-analysis. Eur J Neurol. 2021;28:3478–90.
32. Rogers JP, Chesney E, Oliver D, et al. Psychiatric and neuropsychiatric presentations associated with severe coronavirus infections: a systematic review and meta-analysis with comparison to the COVID-19 pandemic. Lancet Psychiatry. 2020;7:611–27.
33. Bajaj V, Gadi N, Spihlman AP, Wu SC, Choi CH, Moulton VR. Aging, immunity, and COVID-19: how Age influences the host Immune Response to Coronavirus infections? Front Physiol. 2020;11:571416.

Publisher's Note