International Journal of Infectious Diseases 171 (2026) 108963

Contents lists available at ScienceDirect

International Journal of Infectious Diseases

journal homepage: www.elsevier.com/locate/ijid

Time to recovery from long COVID: a longitudinal analysis of symptom duration and risk factors using accelerated failure time models

Youngji Jo[1], Zeyu Hu[2], Hyejin Joo[3], Jaehun Jung[3], Jacob Lee[4,*]

[1] Departments of Public Health Sciences, School of Medicine, UConn Health, Farmington, CT, USA
[2] Department of Statistics, University of Connecticut, Storrs, CT, USA
[3] Department of Preventive Medicine, Korea University College of Medicine, Seoul, Republic of Korea
[4] Division of Infectious Diseases, Department of Internal Medicine, Kangnam Sacred Heart Hospital, Hallym University College of Medicine, Seoul, Republic of Korea

ARTICLE INFO

Article history:
Received 20 April 2026
Revised 21 June 2026
Accepted 2 July 2026

Keywords:
Post-acute COVID-19 syndrome
Accelerated Failure Time
Republic of Korea
Fatigue
Sleep disorder

ABSTRACT

Objectives: Long COVID comprises heterogeneous symptoms that may persist for months to years after acute SARS-CoV-2 infection. While prevalence is well described, less is known about symptom duration and determinants of recovery.
Methods: We conducted a multicenter longitudinal cohort study across 12 institutions in Korea (December 2022-March 2025), enrolling adults with confirmed infection and uninfected controls. Participants were followed for up to 22 months with repeated assessments of symptoms, vaccination, and comorbidities. Eight common symptoms were analyzed. Time to resolution was estimated using Kaplan–Meier methods and log-normal accelerated failure time models, adjusting for demographic and clinical factors. Results are reported as time ratios (TRs).
Results: Among 757 infected participants and 558 controls, respiratory and physical symptoms resolved similarly between groups, whereas systemic and neurological symptoms persisted longer in infected individuals. At one year, 20-25% reported fatigue and 15-20% sleep disturbance versus ~10% and ~5% in controls. Infection was associated with ~50% longer duration of fatigue and sleep disturbance. Individuals aged 40-59 recovered more slowly; other factors showed minimal associations.
Conclusion: Persistent long COVID burden is driven by systemic neuro-related symptoms, highlighting the need for targeted, symptom-specific management strategies.
© 2026 The Author(s). Published by Elsevier Ltd on behalf of International Society for Infectious Diseases. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

Introduction

The COVID-19 pandemic has evolved from an acute public health crisis into a complex challenge characterized by persistent post-acute sequelae, commonly known as long COVID or post-acute sequelae of SARS-CoV-2 infection (PASC) [1,2]. While initial research focused on acute disease management and mortality prevention, accumulating evidence reveals that a substantial proportion of infected individuals experience prolonged symptoms ex10ding weeks, months, or even years beyond the initial infection [3–5]. These persistent manifestations affect multiple organ systems and significantly impact quality of life, functional capacity, and healthcare utilization [6].

Long COVID encompasses a heterogeneous constellation of symptoms, including fatigue, cognitive impairment, respiratory difficulties, musculoskeletal complaints, and neuropsychiatric manifestations. Despite extensive descriptive literature documenting the prevalence and types of long-COVID symptoms, substantial gaps remain in our understanding of symptom duration and the factors that influence recovery trajectories [6,7]. Current estimates suggest that 10-30% of COVID-19 survivors eFxperience at least one symptom lasting beyond 12 weeks, though prevalence varies considerably across studies due to differences in case definitions, follow-up duration, and population characteristics [3,5,8,9]. The pathophysiology of long COVID remains incompletely understood, but risk factors for symptom persistence have been consistently identified [7,10]. Female sex, older age, obesity, pre-existing comorbidities (such as diabetes, hypertension, asthma, and cancer), and greater severity of acute illness are associated with increased risk and longer duration of symptoms. Notably, the number and type of

* Corresponding author.
E-mail address: litjacob@gmail.com (J. Lee).

https://doi.org/10.1016/j.ijid.2026.108963
1201-9712/© 2026 The Author(s). Published by Elsevier Ltd on behalf of International Society for Infectious Diseases. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

symptoms during the acute phase are predictive of prolonged recovery, with fatigue and cognitive symptoms among the most persistent [1]. The impact of vaccination and other behavioral factors remains an area of ongoing investigation. Understanding the natural history and determinants of symptom resolution is critical for clinical management, resource allocation, and public health planning as healthcare systems worldwide grapple with the growing burden of post-COVID care [11,12].

This study addresses these gaps by applying log-normal Accelerated Failure Time (AFT) models to a large prospective cohort with detailed longitudinal follow-up data. We analyze eight prevalent long-COVID symptoms—fatigue, cough, sputum production, rhinorrhea/throat symptoms, sleep disturbance, weakness, arthritis, and low concentration—to characterize their natural resolution patterns and identify demographic, clinical, and infection-related determinants of symptom duration. Our objectives are (1) to quantify the duration of major long-COVID symptoms and compare resolution patterns across the infection status; and (2) to identify demographic, behavioral, and clinical factors associated with prolonged symptom persistence using AFT modeling. The findings will inform clinical prognosis, identify high-risk populations requiring targeted interventions, and contribute to evidence-based guidelines for long-COVID management and surveillance.

Methods

Study Design and Participant Selection This prospective and retrospective cohort study was conducted across 12 hospitals in Korea between December 2022 and March 2025 to investigate long-term COVID-19 sequelae. The study enrolled 8761 adults who provided written informed consent, initially comprising 7208 infected individuals and 1553 uninfected controls. Controls were confirmed negative for SARS-CoV-2 N-protein antigen at enrollment and were followed up at 1, 3, and 6 months. If they were infected during the follow-up period, they were excluded from the control cohort. Infected participants were confirmed via PCR or rapid antigen testing at healthcare facilities (excluding self-test kits) and assessed at 1, 3, and 6 months. Because our objective was to estimate symptom duration following infection, we excluded participants whose most recent infection occurred more than 1 month before any visit to minimize left-censoring and selection bias and balance follow-up time with controls. The primary exposure was SARS-CoV-2 infection; Accordingly, the index date for the infection group was defined as the date of the most recent infection prior to any visits, while for controls, it was defined as the first study visit, corresponding to the N-protein test date at enrollment. This selection process resulted in a final analytic cohort of 757 infected and 558 uninfected participants with follow-up extending up to 22 months (Figure 1). Comprehensive data collection included demographics, Charlson Comorbidity Index (CCI) scores, 56 laboratory results, 29 self-reported symptoms, and mental health assessment results. More detailed information on study procedures, the data collection process, and the full list of survey and laboratory variables is described elsewhere [13].

Outcome measures. The primary outcome was symptom duration. Because some symptoms were infrequent and unsuitable for stable modeling, analyses were restricted to the eight most prevalent symptoms with a prevalence >20% among infected participants: fatigue, sleep disturbance, cough, sputum, weakness, rhinorrhea/throat symptoms, low concentration, and arthritis. This restriction ensured adequate event counts and clinically meaningful estimates while minimizing sparse-data bias.

Construction of time-to-event variables. For each long-COVID symptom under consideration, a person-level survival record was created. Within each participant, symptom onset was defined as the first visit at which the symptom was reported as present (value = 1). Resolution was defined as the first subsequent visit at which the symptom was reported as absent (value = 0). If a symptom remained present through the last available follow-up, the observation was treated as right-censored at the date of the final visit. Symptom duration was therefore defined as the number of days between the first recorded presence and the first subsequent absence of the symptom, with unresolved cases right-censored at the last follow-up. For each symptom, only participants who had ever reported that symptom were included in the analysis. This approach ensured that each eligible participant contributed a single time-to-event record for each symptom they experienced, representing the interval from the first report of that symptom to its resolution or the end of follow-up. Records with non-positive or missing duration values (e.g., ≤0 days or invalid dates) were excluded.

Statistical analysis. First, to visualize the unadjusted time-to-resolution patterns of long-COVID symptoms, we generated Kaplan–Meier (KM) curves for each of the eight selected symptoms. Survival objects were created using the Surv() function, and KM curves were estimated via the survfit() function in R. Curves were stratified by infection status (infected vs control) and visualized with 95% confidence intervals, log-rank test *P*-values, and risk tables indicating the number of participants at risk over time. Plots were produced using the survminer package (ggsurvplot) with standardized axes (days since first symptom) and line style (solid for control, dashed for infected).

Second, we applied a log-normal Accelerated Failure Time (AFT) model to estimate factors associated with the duration of long-COVID symptoms. The log-normal AFT model was chosen because it provides a flexible and interpretable framework for analyzing right-skewed duration data, accommodating long-tailed distributions and non-monotonic hazard shapes commonly observed in post-viral recovery. The model was specified as: $\log(T_i) = X_i^T\beta + \sigma\varepsilon_i$, where T_i is the symptom duration for the participant i, X_i is the vector of covariates, $\sigma > 0$ is the scale parameter, and ε_i follows a standard normal distribution. Parameters were estimated using maximum likelihood. Each symptom-specific model included the following baseline covariates: 1) Demographics and characteristics: infection (control as reference), age group 20-39 as reference, 40-59, ≥60), sex, BMI category (Normal weight 18.5-22.9 as reference; Underweight <18.5; Overweight >22.9), alcohol use, and smoking status; cumulative number of vaccine doses prior to follow-up and infection indicator; and Charlson Comorbidity Index (CCI). In the infection-only group models, we additionally included the number of prior COVID-19 infections and whether symptoms occurred after COVID-19 onset. Model coefficients were exponentiated and reported as Time Ratios ($TR = e^{\beta}$) with corresponding 95% confidence intervals (CIs). A TR > 1 indicates a longer symptom duration (slower resolution), whereas a TR < 1 indicates a shorter duration (faster resolution), compared with the reference category while holding other factors constant.

Sensitivity analyses. To address baseline differences between the infected and control groups, we performed a series of sensitivity analyses. We estimated propensity scores for infection using a logistic model that included age group, sex, BMI category, alcohol use, smoking, Charlson Comorbidity Index, and vaccination, and derived stabilized inverse-probability-of-treatment weights (truncated at the 1st and 99th percentiles); covariate balance was evaluated using standardized mean differences (SMD; |SMD| < 0.1 considered adequate). We repeated the Kaplan–Meier and AFT analyses with these weights and additionally (i) compared crude, covariate-adjusted, inverse-probability-of-treatment weighting (IPW)-weighted, and 1:1 propensity-score-matched estimates (Figures S1-S3); (ii) administratively censored both groups at a shared 6- and 12-month window from symptom onset to address the longer follow-up of the infected group (Figure S4); (iii) per-

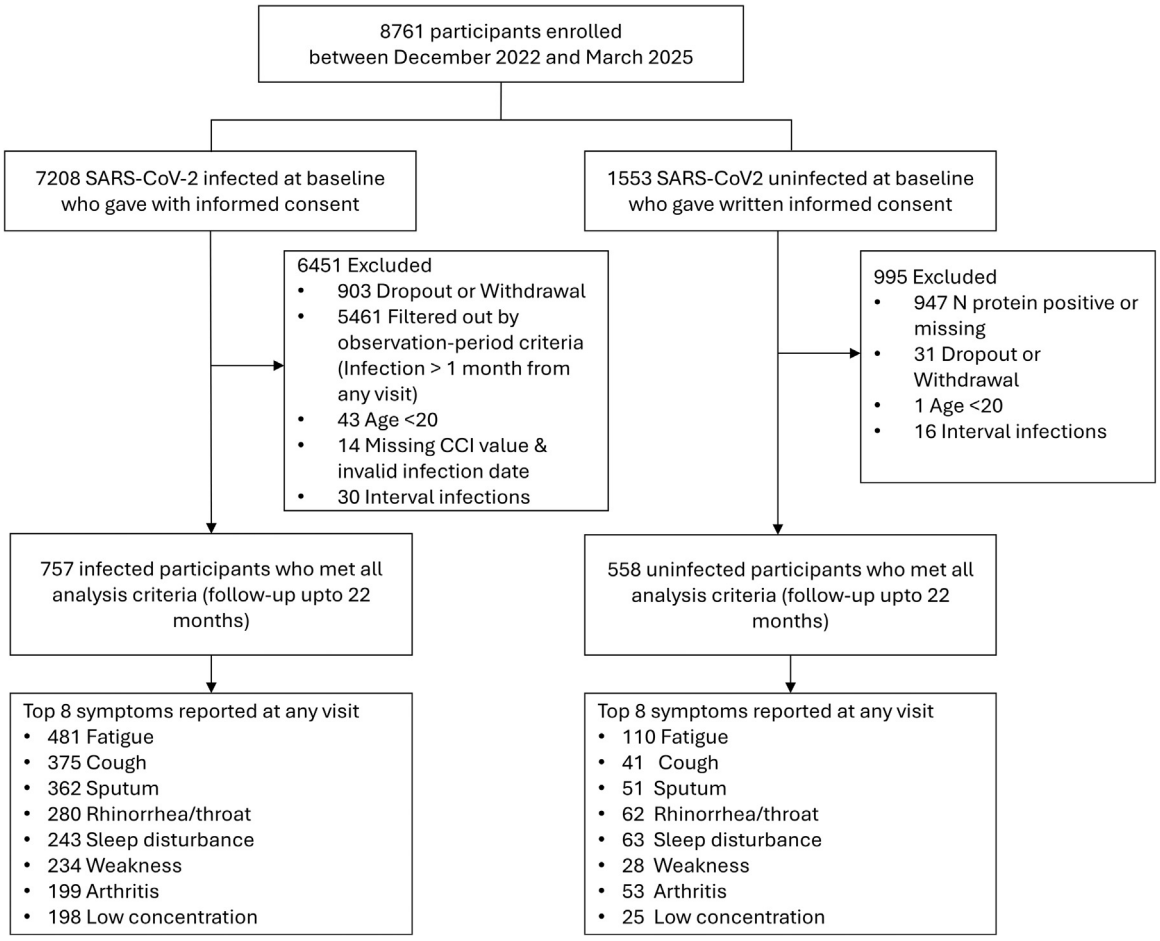


Figure 1. Flow diagram of participant enrollment, exclusions, and final analytic cohorts.

formed an age-stratified analysis of fatigue across age bands (20-39, 40-59, and ≥60 years), reporting covariate-adjusted and IPW-weighted time ratios (Table S1); (iv) examined, within the infected group, whether the period of infection—2024 (JN.1/later) versus 2023 (XBB), used as a proxy for the dominant circulating variant given the absence of individual-level viral sequencing—was associated with symptom duration, identified from a per-visit recent-infection indicator (Table S2). All analyses were performed in R 4.3.1.

Results

Table 1 summarizes baseline demographic, clinical characteristics, and symptom profiles of 757 SARS-CoV-2–infected participants and 558 uninfected controls. The infected group was significantly older, with a higher proportion aged ≥60 years, and had a greater comorbidity burden as reflected by higher CCI scores, while sex distribution and BMI categories were broadly similar between groups. Infected participants were less likely to report alcohol use or smoking and had higher vaccination uptake, particularly receipt of more than three doses. Symptom prevalence differed markedly by infection status: Among infected participants, fatigue was the most frequently reported symptom (63.5% vs 19.7% in controls), followed by cough (49.5% vs 7.3%), sputum production (47.8% vs 9.1%), and rhinorrhea/throat symptoms (37% vs 11.1%). Importantly, a large proportion of symptoms in the infected group represented new-onset manifestations following SARS-CoV-2 infection rather than pre-existing conditions: Among fatigue was the most frequently reported symptom (63.5% vs 19.7% in controls), followed by cough (49.5% vs 7.3%), sputum production (47.8% vs 9.1%), and rhinorrhea/throat symptoms (37.0% vs 11.1%). Importantly, a large proportion of symptoms in the infected group represented new-onset manifestations following SARS-CoV-2 infection rather than pre-existing conditions: Among infected participants, 40.7% reported new-onset fatigue (accounting for approximately 64% of all fatigue cases), 42.1% reported cough (85% of all cough cases), 37.5% reported sputum (78% of all sputum cases), and 25.4% reported rhinorrhea or throat symptoms (69% of all such cases). Across all symptom categories, differences between infected and control participants were highly statistically significant ($P < 0.001$), indicating a pronounced excess burden of both incident and persistent symptoms attributable to SARS-CoV-2 infection rather than background population symptom prevalence.

Figure 2 shows Kaplan–Meier estimates of time to symptom resolution for eight major COVID-19–related symptoms, comparing SARS-CoV-2–infected participants with uninfected controls. The crude estimates of symptom duration revealed a clear divergence across symptoms: while respiratory and physical symptoms such as cough, sputum, and weakness showed rapid resolution rates indistinguishable from controls, systemic sequelae exhibited a significantly protracted course. Specifically, fatigue ($P = 0.0072$) and sleep disturbances ($P = 0.048$) persisted at elevated rates in the infected group. At 1 year, approximately 20-25% of infected participants still experienced unresolved fatigue, compared with roughly 10% of controls, and approximately 15-20% reported persistent sleep disturbances, compared with only ~5% of controls. These trends highlight symptom-specific heterogeneity in post-COVID re-

Table 1
Baseline demographic, clinical characteristics and symptom profiles of study participants by infection status.

Variable		Infected (n = 757)		Control (n = 558)	Infected IPW-weighted (n = 757)	Control IPW-weighted (n = 555)	Chi-square test
Sex	Female	500 (66.1%)		385 (69.0%)	510 (67.3%)	376 (67.7%)	0.344
	Male	257 (33.9%)		173 (31.0%)	247 (32.7%)	179 (32.3%)	
Age	20-39	325 (42.9%)		248 (44.4%)	331 (43.7%)	243 (43.8%)	<0.001***
	40-59	230 (30.4%)		246 (44.1%)	274 (36.1%)	203 (36.5%)	
	60+	202 (26.7%)		64 (11.5%)	153 (20.2%)	110 (19.7%)	
BMI	Underweight (<18.5)	329 (43.5%)		273 (48.9%)	348 (45.9%)	257 (46.2%)	0.120
	Normal weight (18.5-22.9)	403 (53.2%)		264 (47.3%)	383 (50.6%)	280 (50.4%)	
	Overweight (>22.9)	25 (3.3%)		21 (3.8%)	26 (3.5%)	19 (3.4%)	
Alcohol (Yes)		293 (38.7%)		272 (48.7%)	430 (56.8%)	313 (56.4%)	<0.001***
Smoking (Yes)		68 (9.0%)		79 (14.2%)	673 (88.9%)	493 (88.8%)	0.004**
Vaccination	0	24 (3.2%)		16 (2.9%)	25 (3.2%)	15 (2.7%)	0.004**
	1-2	140 (18.5%)		79 (14.2%)	146 (19.3%)	76 (13.7%)	
	3	391 (51.7%)		346 (62.0%)	405 (53.5%)	334 (60.1%)	
	>3	202 (26.7%)		117 (21.0%)	182 (24.0%)	131 (23.5%)	
Charlson Comorbidity Index (CCI)	None (0)	225 (29.7%)		153 (27.4%)	218 (28.8%)	159 (28.6%)	<0.001***
	Moderate (1-2)	438 (57.9%)		371 (66.5%)	467 (61.7%)	343 (61.7%)	
	Severe/High (≥3)	94 (12.4%)		34 (6.1%)	72 (9.5%)	54 (9.7%)	
Number of prior infections*	0	0 (0.0%)		558 (100.0%)	NA	555 (100.0%)	NA
	1	485 (64.1%)		0 (0.0%)	489 (64.6%)	NA	
	2	262 (34.6%)		0 (0.0%)	258 (34.1%)	NA	
	3	10 (1.3%)		0 (0.0%)	9 (1.2%)	NA	
Symptoms		Total (preexisting)	Total (new onset)	Total	Total (preexisting)	Total	Chi-square test
	Fatigue	481 (63.5%)	308 (40.7%)	110 (19.7%)	484 (63.9%)	110 (19.8%)	<0.001***
	Cough	375 (49.5%)	319 (42.1%)	41 (7.3%)	370 (48.9%)	45 (8.1%)	<0.001***
	Sputum	362 (47.8%)	284 (37.5%)	51 (9.1%)	356 (47.0%)	53 (9.6%)	<0.001***
	Rhinorrhea/throat	280 (37.0%)	192 (25.4%)	62 (11.1%)	277 (36.6%)	65 (11.7%)	<0.001***
	Sleep disturbance	243 (32.1%)	102 (13.5%)	63 (11.3%)	239 (31.6%)	64 (11.5%)	<0.001***
	Weakness	234 (30.9%)	168 (22.2%)	28 (5.0%)	231 (30.5%)	32 (5.8%)	<0.001***
	Arthritis	199 (26.3%)	83 (11.0%)	53 (9.5%)	195 (25.7%)	59 (10.6%)	<0.001***
	Low concentration	198 (26.2%)	140 (18.5%)	25 (4.5%)	197 (26.0%)	29 (5.2%)	<0.001***
	Others†	580 (76.6%)	312 (41.2%)	152 (27.2%)	577 (76.2%)	156 (28.1%)	<0.001***

Note: †Others in symptoms refers to participants who reported at least one of the following symptoms: fever, chill, myalgia, rash, sore throat, chest pain, dyspnea, palpitation, cognition disorder, hearing disturbance, loss of appetite, anosmia, ageusia, headache, dizziness, other neurological symptoms, diarrhea, nausea, pain, constipation, other gastrointestinal symptoms, or alopecia. The number of prior infections and post-COVID-19 symptom-onset variables were included only in an infection group model. Statistical significance: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

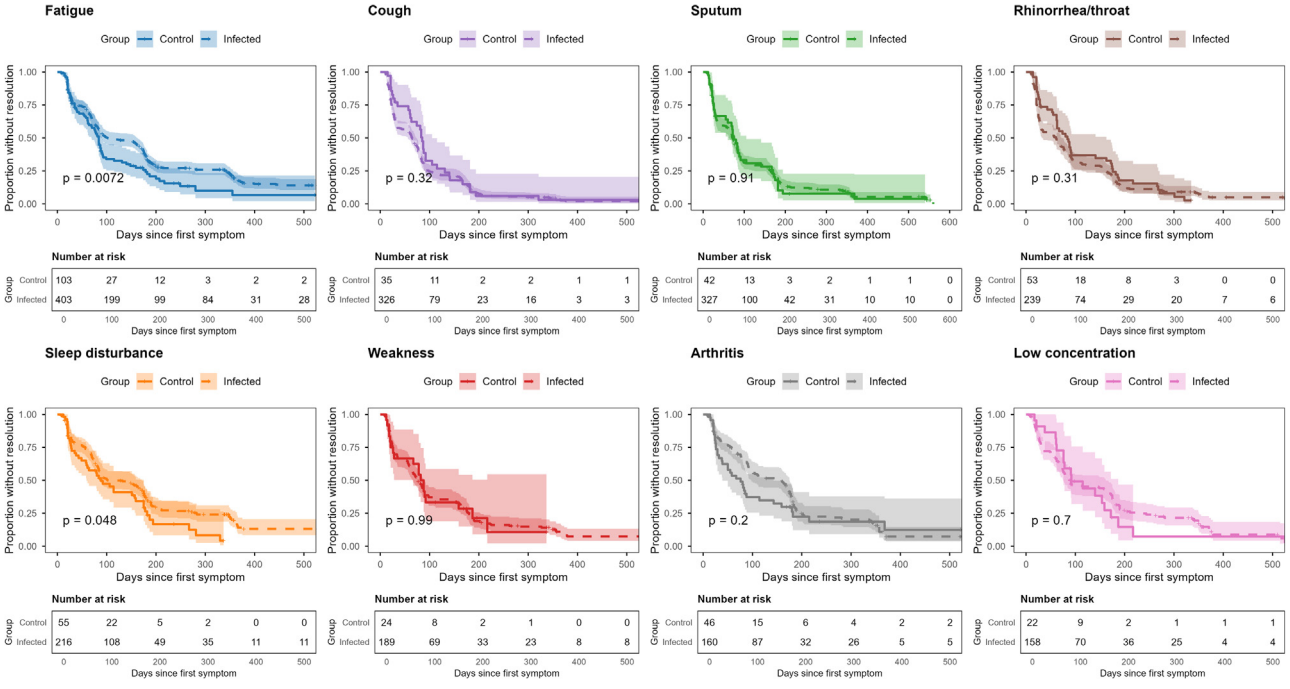


Figure 2. Time to symptom resolution by infection status for major COVID-19–related symptoms.

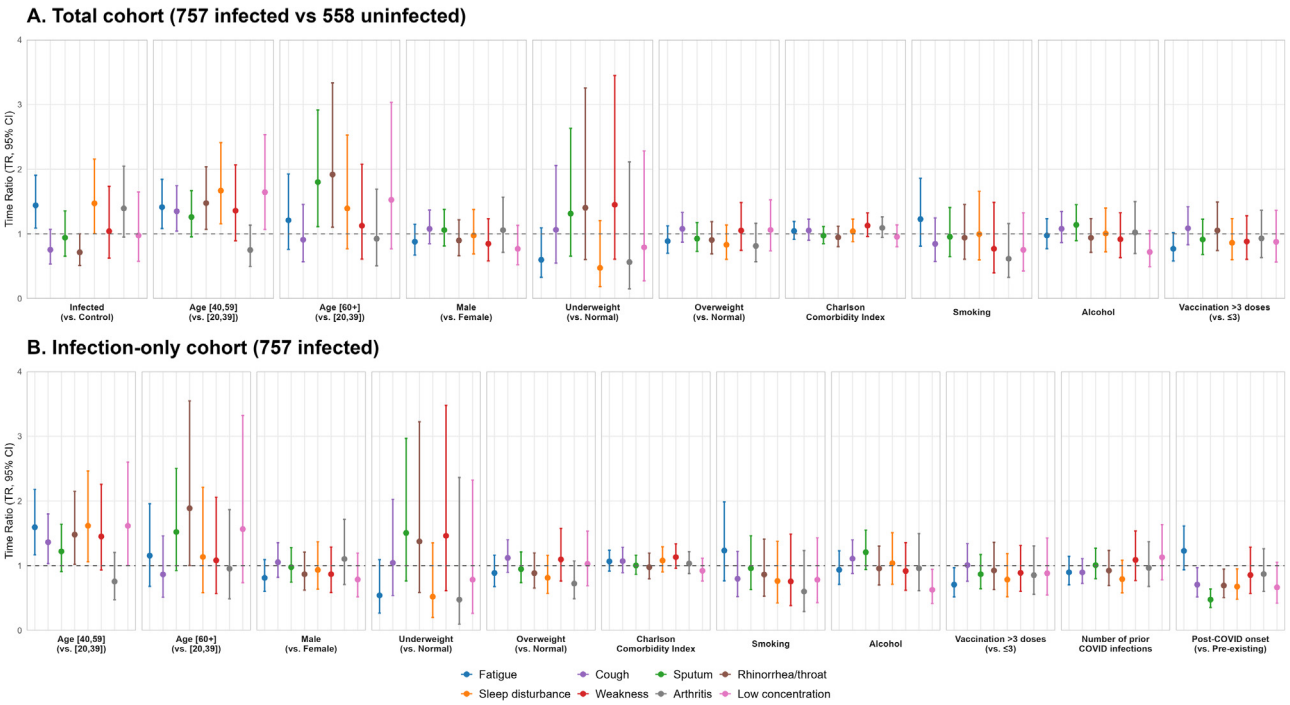


Figure 3. Adjusted associations between COVID-19 infection and symptom persistence across demographic and clinical subgroups.

covery and underscore that neuro-systemic burdens can linger well beyond the acute phase.

After adjusting for potential confounders—including age, sex, BMI, comorbidity index, and vaccination status—SARS-CoV-2 infection remained a significant independent predictor of prolonged symptom duration, particularly for systemic and neurological sequelae. In the total cohort analysis (Figure 3, Panel A), infected participants exhibited Time Ratios (TR) significantly greater than 1 for fatigue (TR: 1.44; 95% CI: 1.09-1.91) and sleep disturbance (TR: 1.47; 95% CI: 1.01-2.16). The infected-versus-control time ratios were consistent across crude, covariate-adjusted, IPW-weighted, and propensity-score-matched analyses (Figure S2). Beyond infection status, age emerged as a notable determinant, with participants aged 40-59 demonstrating consistently higher Time Ratios compared to the 20-39 reference group in most symptoms, including fatigue (TR: 1.41; 95% CI: 1.08-1.84), sleep disturbance (TR: 1.67; 95% CI: 1.16-2.41), and low concentration (TR: 1.65; 95% CI: 1.07-2.53). An age-stratified analysis (Table S1) confirmed that the prolonged fatigue was strongest among working-age adults, especially those aged 40-59 years (TR: 1.88, 95% CI: 1.18-2.99), compared with those aged 20-39 (TR: 1.32, 95% CI: 0.88-1.97) or ≥60 years (TR: 0.91, 95% CI: 0.45-1.85). In contrast, other covariates, such as sex (Male vs Female), the CCI, and the number of prior COVID infections, showed Time Ratios consistently clustering near 1.0 with narrow confidence intervals, suggesting these variables did not significantly alter the duration of established symptoms.

The associations observed in the infection-only models were broadly consistent with those estimated in the full cohort, with similar directions and magnitudes across most demographic and clinical covariates (Figure 3, Panel B). However, the timing of symptom onset (Post-COVID onset vs Pre-existing) showed divergent effects: new-onset fatigue tended to be associated with longer symptom duration than pre-existing fatigue, although the confidence interval included unity (TR: 1.23; 95% CI: 0.94-1.61), whereas several new-onset respiratory and related symptoms—including sleep disturbance (TR: 0.68; 95% CI: 0.48-0.95), sputum production (TR: 0.48; 95% CI: 0.35-0.64), cough (TR: 0.71; 95% CI: 0.52-0.97), and rhinorrhea/throat symptoms (TR: 0.69; 95% CI: 0.51-0.95)—were associated with shorter duration compared with symptoms already present before infection. This pattern suggests that newly emerging fatigue may reflect a more persistent post-COVID condition, whereas many respiratory symptoms occurring after infection may represent transient manifestations that resolve over time. The number of prior infections showed no consistent association with symptom persistence across symptoms, with most estimates close to the null.

Discussion

In this longitudinal cohort study of 1315 participants followed for up to 22 months, we used AFT models to characterize symptom resolution trajectories following SARS-CoV-2 infection. Distinct recovery patterns emerged: respiratory and physical symptoms such as cough, sputum, and weakness resolved at rates comparable to uninfected controls, whereas systemic and neurological sequelae—particularly fatigue and sleep disturbance—persisted substantially longer. These associations remained robust after adjustment for demographic and clinical covariates, with infection status independently predicting prolonged symptom duration in long-term follow-up. Specifically, infected participants experienced significantly slower resolution of these systemic symptoms, with time ratios (TR) indicating approximately 50% longer persistence for fatigue and sleep disturbance. Together, these findings indicate that Long COVID chronicity is driven primarily by systemic neuroimmunological dysregulation rather than isolated respiratory injury.

The divergence between respiratory and systemic symptom trajectories is a key finding that aligns with emerging understanding of long COVID pathophysiology [14–16]. The Kaplan-Meier analysis and AFT models concordantly showed that respiratory symptoms (cough, sputum, rhinorrhea) in infected participants rapidly converged with baseline rates observed in uninfected controls. This suggests that for the majority of patients, direct mucosal or pulmonary injury from SARS-CoV-2 heals within a definable time-

frame. In contrast, fatigue and sleep disturbance demonstrated protracted trajectories, with 20-25% of infected participants reporting unresolved fatigue and 15-20% persistent sleep issues at 1 year—rates two- to fourfold higher than controls. These findings are consistent with global estimates from systematic reviews, where fatigue persists in 30-50% of survivors at 6-12 months, and sleep disturbances affect 20-40%, often clustering with cognitive and neuropsychiatric features [17,18]. Neuroinflammation has emerged as a central mechanism underlying the persistent neurological and systemic features of long COVID, rather than residual organ damage alone [19–21]. Additional proposed mechanisms include viral reservoir persistence [22], immune dysregulation [23], or autonomic dysfunction [24], rather than residual organ damage alone [19].

Our multivariate analysis provides nuanced insights into risk factors for symptom chronicity. Consistent with prior literature, older age (40-59 years) was a significant determinant of prolonged recovery, particularly for fatigue and sputum [25,26]. However, sex, BMI, CCI, and reinfection history clustered near null effects (TR ≈1.0), contrasting some prior reports linking female sex and comorbidities to higher long COVID risk [19]. This suggests that while certain demographics may increase susceptibility to the initial post-acute condition, the biological drivers of symptom persistence (how long it lasts) may be independent of pre-existing metabolic or comorbidity profiles. The divergent effects of new-onset versus pre-existing symptoms—faster resolution for incident sputum and rhinorrhea/throat but prolonged new fatigue—highlight infection-triggered de novo pathophysiology, possibly involving immune dysregulation or neuroinflammation [27,28].

This interpretation is supported by our finding that higher vaccination exposure (>3 doses) may accelerate the resolution of systemic symptoms such as fatigue (TR: 0.71; 95% CI: 0.52-0.97) among the infected cohort, but has limited associations with most symptom durations. This pattern may reflect the systemic nature of fatigue, which has been linked to immune dysregulation and neuroinflammatory mechanisms in long COVID. One possible explanation is that booster vaccination may reduce persistent antigenic stimulation or low-level viral reservoirs, thereby attenuating chronic immune activation and neuroinflammatory or autonomic pathways implicated in fatigue and sleep dysregulation [29]. In contrast, respiratory symptoms such as cough, sputum, and rhinorrhea/throat irritation are more likely driven by localized airway inflammation and may therefore be less influenced by vaccination status. Prior studies report mixed evidence on vaccination and long COVID resolution, showing a 23-46% reduction in long COVID risk [30–32] but variable symptom improvement after infection (20% improvement vs 54% no change) [32–34]. Together, these findings suggest that vaccination may preferentially influence systemic post-COVID manifestations rather than uniformly affecting all symptom trajectories. Because vaccine type/brand was incompletely recorded (>60% missing), a type-based analysis was not feasible; future studies should examine whether vaccine platform (e.g., mRNA vs viral-vector) affects symptom resolution.

Several limitations warrant consideration. First, our study did not fully capture time-varying phenotype transitions or phenotype-specific recovery dynamics across infected and uninfected cohorts. In addition, symptom recovery was not strictly monotonic, and relapsing–remitting patterns may have been underrepresented. To preserve interpretability of symptom-specific recovery times, we modeled individual symptom trajectories using AFT models and adopted a conservative approach that focused on time to first resolution, as relapse or symptom re-emergence can be difficult to attribute specifically to long COVID; Future research should employ dynamic modeling frameworks, such as multi-state or latent transition analysis, to explicitly characterize the temporal evolution of symptom phenotypes and the non-monotonic dynamics of relapse and remission. Second, heterogeneity in follow-up duration and participant attrition may introduce survival or selection bias, as individuals with persistent symptoms may be more likely to remain under observation, while those who recover quickly may disengage from follow-up. However, stratified analyses in a > 12-month follow-up sub-cohort showed patterns similar to those in the full cohort (Figure S4), suggesting minimal influence on the estimated recovery patterns. Third, reliance on discrete follow-up visits introduces interval censoring and potential misclassification of symptom onset and resolution times, as exact transition dates are not observed. These limitations highlight the need for cautious interpretation of absolute duration estimates, while supporting the validity of relative comparisons across infection status and follow-up strata that are central to our conclusions [35]. Fourth, our symptom-specific analyses were restricted to participants who ever reported each symptom, which is necessary for defining time-to-resolution but may introduce selection bias by excluding individuals with milder or transient symptoms that were resolved between visits or were never reported. Thus, our estimates reflect the recovery course of *symptomatic* individuals rather than the average burden across the entire infected population. In addition, although interval reinfection was ascertained by self-report (from PCR and rapid antigen tests) at each visit and is likely underestimated as fully asymptomatic or unrecorded infections cannot be determined. Such misclassification of controls would bias the infected-versus-control contrast toward the null, rendering our estimates conservative. We did not measure the burden of seasonal co-infections (e.g., RSV, influenza, rhinovirus). However, because controls were enrolled contemporaneously, this burden is likely shared by both arms and partially differenced out in the infected-versus-control contrast. Finally, we focused on the eight most prevalent symptoms reported among infected participants; less common manifestations were not examined, which may limit the generalizability of our findings to populations with different symptom profiles.

Nonetheless, a distinct strength of this study lies in its granular examination of the natural history of eight individual symptoms, stratified by full and infection-only cohorts, which reveal critical divergences often masked by aggregate “symptom burden” scores or “long COVID indices.” The trajectory data from the NIH RECOVER Adult Cohort [36] identified that 5% of patients had persistently high symptom burden, 12% had intermittently high burden, and notably, 14% did not meet criteria initially but developed increasing symptoms by 15 months—underscoring the heterogeneity and chronicity of the condition. By modeling time-to-resolution for specific manifestations rather than composite outcomes, we identified a definitive bifurcation in recovery trajectories: while respiratory and physical symptoms—such as cough, sputum, and weakness—resolved at rates statistically indistinguishable from uninfected controls, systemic and neurological sequelae like fatigue and sleep disturbance exhibited significant, protracted persistence. Our findings indicate that the “long tail” of recovery is not a generalized prolongation of acute illness, but rather the specific entrenchment of neuro-systemic burdens that become increasingly distinct from background risks over time. These results highlight the importance of symptom-specific monitoring and management in long COVID, particularly focusing on systemic manifestations such as fatigue and sleep disturbance rather than respiratory symptoms that may reflect background population prevalence.

Conclusions

In this longitudinal cohort study with follow-up extending to 22 months, we found that recovery trajectories following SARS-CoV-2 infection differ markedly across symptom types. While respiratory and physical symptoms such as cough, sputum, and weakness resolved at rates comparable to uninfected controls, systemic and neurological manifestations—particularly fatigue and

sleep disturbance—persisted substantially longer and remained independently associated with infection after adjustment for demographic and clinical factors. These findings suggest that the long-term burden of post-COVID conditions is driven primarily by persistent neuro-systemic sequelae rather than generalized prolongation of acute respiratory illness. Recognizing this symptom-specific heterogeneity has important clinical implications, highlighting the need for targeted monitoring and management strategies that prioritize systemic manifestations such as fatigue and sleep disturbance in long COVID care.

CRediT authorship contribution statement

Youngji Jo: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Supervision, Visualization, Writing – original draft, Writing – review & editing. **Zeyu Hu:** Formal analysis, Visualization, Writing – original draft. **Hyejin Joo:** Data curation, Project administration. **Jaehun Jung:** Data curation, Funding acquisition, Investigation, Project administration, Resources, Supervision, Writing – review & editing. **Jacob Lee:** Funding acquisition, Investigation, Project administration, Supervision, Writing – review & editing.

Data and code availability

All analysis code and processed data used in this study are publicly available at: https://github.com/UConn-Health-Disease-Modeling/longcovid-aft

Declaration of generative AI and AI-assisted technologies in the writing process

Generative AI tools used to correct grammar and spelling errors. After using this tool, the authors carefully reviewed and edited the content as needed and take full responsibility for the final version of the publication.

Declarations of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

Korea Health Technology R&D project through the Korea Health Industry Development Institute (KHIDI), funded by the National Institute of Infectious Diseases, National Institute of Health, Republic of Korea (grant number: HD22C2045). The funder of this study had no involvement in the direction, content, or conclusions of the analysis.

Ethical approval

All participants submitted written informed consent, and the protocol was reviewed by the institutional review boards of each participating center.

Acknowledgments

The data used for this study are not publicly available at present in accordance with KDCA policies. However, the data may become available in the future, subject to the authority of the KDCA.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ijid.2026.108963.

References

[1] Subramanian A, Nirantharakumar K, Hughes S, Myles P, Williams T, Gokhale KM, et al. Symptoms and risk factors for long COVID in non-hospitalized adults. *Nat Med* 2022;**28**:1706–14.
[2] Wynberg E, van Willigen HDG, Dijkstra M, Boyd A, Kootstra NA, van den Aardweg JG, et al. Evolution of coronavirus disease 2019 symptoms during the first 12 months after illness onset. *Clin Infect Dis* 2022;**75**:e482-90.
[3] Thompson EJ, Williams DM, Walker AJ, Mitchell RE, Niedzwiedz CL, Yang TC, et al. Long COVID burden and risk factors in 10 UK longitudinal studies and electronic health records. *Nat Commun* 2022;**13**:3528.
[4] Wulf Hanson S, Abbafati C, Aerts JG, Al-Aly Z, Ashbaugh C, et al., Global Burden of Disease Long COVID Collaborators Estimated global proportions of individuals with persistent fatigue, cognitive, and respiratory symptom clusters following symptomatic COVID-19 in 2020 and 2021. *JAMA* 2022;**328**:1604–15.
[5] Wu Q, Ailshire JA, Crimmins EM. Long COVID and symptom trajectory in a representative sample of Americans in the first year of the pandemic. *Sci Rep* 2022;**12**:11647.
[6] Alkodaymi MS, Omrani OA, Ashraf N, Shaar BA, Almamlouk R, Riaz M, et al. Prevalence of post-acute COVID-19 syndrome symptoms at different follow-up periods: a systematic review and meta-analysis. *Clin Microbiol Infect* 2022;**28**:657–66.
[7] Tsampasian V, Elghazaly H, Chattopadhyay R, Debski M, Naing TKP, Garg P, et al. Risk factors associated with post-COVID-19 condition: a systematic review and meta-analysis. *JAMA Intern Med* 2023;**183**:566–80.
[8] Arjun MC, Singh AK, Pal D, Das K, G A, Venkateshan M, et al. Characteristics and predictors of long COVID among diagnosed cases of COVID-19. *PLoS One* 2022;**17**:e0278825.
[9] Whitaker M, Elliott J, Chadeau-Hyam M, Riley S, Darzi A, Cooke G, et al. Persistent COVID-19 symptoms in a community study of 606,434 people in England. *Nat Commun* 2022;**13**:1957.
[10] Sudre CH, Murray B, Varsavsky T, Graham MS, Penfold RS, Bowyer RC, et al. Attributes and predictors of long COVID. *Nat Med* 2021;**27**:626–31.
[11] Robineau O, Zins M, Touvier M, Wiernik E, Lemogne C, de Lamballerie X, et al. Long-lasting symptoms after an acute COVID-19 infection and factors associated with their resolution. *JAMA Netw Open* 2022;**5**:e2240985.
[12] Taquet M, Dercon Q, Luciano S, Geddes JR, Husain M, Harrison PJ. Incidence, co-occurrence, and evolution of long-COVID features: a 6-month retrospective cohort study. *PLoS Med* 2021;**18**:e1003773.
[13] Seo G, Joo H, Song K, Kim M, Jung EJ, Kim ES, et al. Static and dynamic scoring systems for post-acute sequelae of SARS-CoV-2 in a Korean cohort. *Int J Infect Dis* 2026;**164**:108378.
[14] Ballouz T, Menges D, Anagnostopoulos A, Domenghino A, Aschmann HE, Frei A, et al. Recovery and symptom trajectories up to two years after SARS-CoV-2 infection. *BMJ* 2023;**381**:e074425.
[15] Tran VT, Porcher R, Pane I, Ravaud P. Course of post COVID-19 disease symptoms over time in the ComPaRe cohort. *Nat Commun* 2022;**13**:1812.
[16] Thant TM, Khandai AC, Gillan A, Peace M, Quinn D, Levenson J. Neuropsychiatric symptoms of subacute and chronic long COVID. *Am J Psychiatry* 2025;**182**:498–9.
[17] Hou Y, Gu T, Ni Z, Shi X, Ranney ML, Mukherjee B, et al. Global prevalence of long COVID and risk factors: a systematic review and meta-analysis. *Open Forum Infect Dis* 2025;**12**:ofaf533.
[18] Batool-Anwar S, Fashanu OS, Quan SF. Long-term effects of COVID-19 on sleep patterns. *Thorac Res Pract* 2025;**26**:9–16.
[19] Hein ZM, Thazin KS, Che Ramli MD, Che Mohd Nassir CMN. Immunomodulatory mechanisms underlying neurological manifestations in long COVID. *Int J Mol Sci* 2025;**26**:6214.
[20] Leng A, Shah M, Ahmad SA, Premraj L, Wildi K, Li Bassi G, et al. Pathogenesis underlying neurological manifestations of long COVID syndrome. *Cells* 2023;**12**:816.
[21] Sherif ZA, Gomez CR, Connors TJ, Henrich TJ, Reeves WB, Force RMPT. Pathogenic mechanisms of post-acute sequelae of SARS-CoV-2 infection. *Elife* 2023;**12**:e86002.
[22] Spatola M, Nziza N, Jung W, Deng Y, Yuan D, Dinoto A, et al. Neurologic sequelae of COVID-19 determined by immunologic imprinting. *Brain* 2023;**146**:4292–305.
[23] Frontera JA, Simon NM. Bridging knowledge gaps in neuropsychiatric sequelae of COVID-19. *JAMA Psychiatry* 2022;**79**:811–17.
[24] Cheng AL, Herman E, Abramoff B, Anderson JR, Azola A, Baratta JM, et al. Multidisciplinary guidance on long COVID management. *PM R* 2025;**17**:684–708.
[25] Estrada-Codecido J, Chan AK, Andany N, Lam PW, Nguyen M, Pinto R, et al. Predictors of persistent post-COVID-19 symptoms. *J Assoc Med Microbiol Infect Dis Can* 2022;**7**:208–19.
[26] Zeng N, Zhao YM, Yan W, Li C, Lu QD, Liu L, et al. Long-term sequelae of COVID-19: systematic review and meta-analysis. *Mol Psychiatry* 2023;**28**:423–33.
[27] Davis HE, McCorkell L, Vogel JM, Topol EJ. Long COVID: major findings and mechanisms. *Nat Rev Microbiol* 2023;**21**:133–46.

[28] VanElzakker MB, Bues HF, Brusaferri L, Kim M, Saadi D, Ratai EM, et al. Neuroinflammation in long COVID assessed by PET. *Brain Behav Immun* 2024;**119**:713–23.
[29] Bohmwald K, Diethelm-Varela B, Rodriguez-Guilarte L, Rivera T, Riedel CA, Gonzalez PA, et al. Immunological features of long COVID. *Front Immunol* 2024;**15**:1341600.
[30] Carazo S, Phimmasone J, Skowronski DM, Giguere K, Ouakki M, Talbot D, et al. Effectiveness of vaccination to reduce long COVID risk. *Clin Infect Dis* 2026;**82**(3):e436–45.
[31] Green R, Marjenberg Z, Lip GYH, Banerjee A, Wisnivesky J, Delaney BC, et al. Impact of vaccination on long COVID: meta-analysis. *Nat Commun* 2025;**16**:10326.
[32] Watanabe A, Iwagami M, Yasuhara J, Takagi H, Kuno T. Vaccination and long COVID prevention: meta-analysis. *Vaccine* 2023;**41**:1783–90.
[33] Ayoubkhani D, Bermingham C, Pouwels KB, Glickman M, Nafilyan V, Zaccardi F, et al. Trajectory of long COVID symptoms after vaccination. *BMJ* 2022;**377**:e069676.
[34] Malden DE, Liu IA, Qian L, Sy LS, Lewin BJ, Asamura DT, et al. Post-COVID conditions following vaccination. *Nat Commun* 2024;**15**:4101.
[35] Greenhalgh T, Sivan M, Perlowski A, Nikolich JZ. Long COVID: a clinical update. *Lancet* 2024;**404**:707–24.
[36] Thaweethai T, Donohue SE, Martin JN, Hornig M, Mosier JM, Shinnick DJ, et al. Long COVID trajectories in the RECOVER cohort. *Nat Commun* 2025;**16**:9557.