ORIGINAL ARTICLE

MJS PUBLISHING

SUBPHENOTYPIC CLASSIFICATION OF COVID-19 SURVIVORS AND RESPONSE TO TELEREHABILITATION: A LATENT CLASS ANALYSIS

Yide WANG, PhD[1], Qianqian XUE, PhD[1], Zheng LI, PhD [1,2] and Fengsen LI, PhD [1,2]
From the [1]Department of Integrated Pulmonology, The Fourth Clinical Medical College of Xinjiang Medical University, Urumqi, Xinjiang, and [2]Xinjiang National Clinical Research Base of Traditional Chinese Medicine, Xinjiang Medical University, Urumqi, Xinjiang, China

Objective: **Investigating the role of telerehabilitation in aiding recovery and societal reintegration for COVID-19 survivors, this study aims to identify distinct subphenotypes among survivors and assess their responsiveness to telerehabilitation.**
Design: **A secondary analysis of a multicentre, parallel-group randomized controlled trial from April 2020 through to follow-up in 2021.**
Subjects/Patients: **The study included 377 COVID-19 survivors (47.1% male), with a mean age of 56.4 years.**
Methods: **Data from the Telerehabilitation Programme for COVID-19 (TERECO) were analysed using Latent Class Analysis to identify subphenotypes based on baseline characteristics. Clinical outcomes were compared between subphenotypes and treatment groups.**
Results: **Latent Class Analysis identified 2 phenotypes: Phenotype 1 (52.9%) characterized by impaired lung function and Phenotype 2 (47.1%) with better lung function. Among those receiving corticosteroids, only Phenotype 1 showed significant benefits from the TERECO intervention. Discrimination accuracy using forced expiratory volume in 1 s (FEV1) and peak expiratory flow was high (AUC = 0.936).**
Conclusion: **Two distinct phenotypes were identified in COVID-19 survivors, suggesting potential improvements in clinical trial design and personalized treatment strategies based on initial pulmonary function. This insight can guide more targeted rehabilitation approaches, enhancing recovery outcomes for specific survivor groups.**

Key words: COVID-19 survivors; telerehabilitation; latent class analysis; subphenotypes.

Submitted Dec 14, 2024. Accepted after revision Mar 10, 2025

Published Mar 26, 2025. DOI: 10.2340/jrm.v57.42726

J Rehabil Med 2025; 57: jrm42726.

Correspondence address: Zheng Li, Department of Integrated Pulmonology, The Fourth Clinical Medical College of Xinjiang Medical University, 116 Huanghe Road, Xinjiang, Urumqi, 830000, China. E-mail: benjasz123@163.com

LAY ABSTRACT

Many people who recover from COVID-19 face ongoing breathing problems and health issues. To aid recovery, we explored whether different groups benefit differently from remote rehabilitation programmes delivered via smartphones and wearable devices. Our study included 377 COVID-19 survivors still experiencing breathing difficulties. We found 2 main types of survivors: 1 group with weaker lungs and another with stronger lungs. Those with weaker lungs benefited more from the remote rehabilitation programme, especially if they were taking certain medications. This means that personalized rehabilitation can be more effective for specific groups of survivors. Our findings suggest that doctors should tailor rehabilitation programmes to match each patient's needs. This approach can lead to better recovery outcomes and improved quality of life for people recovering from COVID-19. By understanding these differences, healthcare providers can offer more targeted support to help patients regain their strength and well-being.

A multitude of individuals who have survived COVID-19 continue to grapple with persistent symptoms months after the initial infection. A comprehensive systematic review, encompassing 194 global studies, revealed that 45% of survivors, regardless of their hospitalization status, still exhibited at least 1 unresolved symptom 4 months post-infection (1). Among the enduring symptoms afflicting COVID-19 survivors, dyspnoea stands out as particularly prevalent (2–4). The widespread transmission of COVID-19 has been studied extensively, with research utilizing large-scale human mobility data to analyse SARS-CoV-2 transmission patterns and their implications for public health measures (5). These studies highlight the global reach of the virus and its lasting consequences, which reinforce the need for effective post-infection care. In light of this, the provision of rehabilitative services becomes paramount to alleviate the burden on both patients and the healthcare system (6). Research findings have already suggested the efficacy of rehabilitation measures, including telerehabilitation, in addressing respiratory ailments associated with COVID-19 and other respiratory diseases (6–10). The pandemic has also led to notable shifts in physical activity, pain levels, and psychological well-being, as evidenced by international multicentre studies assessing behaviour changes during lockdowns (11). These findings emphasize the need for rehabilitation strategies that not only

JRM Journal of Rehabilitation Medicine JRM Journal of Rehabilitation Medicine JRM

restore respiratory function but also address broader health consequences. Furthermore, a recent clinical randomized controlled trial has illuminated the benefits of intervening with telerehabilitation programs specifically tailored for COVID-19 survivors experiencing dyspnoea as a predominant symptom. The intervention has proven advantageous in facilitating the recovery of patients' physical endurance and overall mobility (12).

Latent Class Analysis (LCA) is a rigorously validated statistical method that employs objective criteria to identify phenotypes within broader populations (12). Its application in studying disease phenotypes can offer valuable insights for guiding foundational, clinical, and translational research efforts (13–15). Since the discovery of COVID-19, researchers have recognized its clinical and biological heterogeneity. Recent clustering analyses of COVID-19 patients based on respiratory support requirements have further highlighted the variation in clinical presentations and long-term outcomes (16). However, in comparison with respiratory diseases such as acute respiratory distress syndrome, the study of phenotypes associated with COVID-19 has been notably deficient. To date, research employing LCA to identify phenotypes among survivors of COVID-19 experiencing dyspnoea is lacking. Understanding whether different subphenotypes exist among COVID-19 survivors with dyspnoea, and assessing potential variations in the sensitivity to telerehabilitation across these subphenotypes, remains an unexplored area. Moreover, a systematic review of quality of life in COVID-19-affected countries has revealed significant disparities, suggesting the need for tailored rehabilitation approaches to address the diverse impacts of the disease (17). Addressing these inquiries is undeniably crucial, especially in the context of tailoring telerehabilitation interventions for COVID-19 survivors dealing with dyspnoea.

To address these inquiries, we conducted a secondary analysis of a randomized controlled trial assessing the intervention effects of the telerehabilitation programme for COVID-19 (TERECO) on survivors. We hypothesized that COVID-19 survivors with dyspnoea may exhibit distinct phenotypes, each potentially associated with varied prognoses. Furthermore, we postulated that certain phenotypes might display greater sensitivity to interventions like TERECO. In this regard, whole-body vibration exercise has been proposed as a potential rehabilitation strategy for improving clinical conditions in COVID-19 patients, offering an additional avenue for structured recovery interventions (18). This assumption underscores the potential for future therapeutic trials in COVID-19 survivors to benefit from phenotype-based stratification, allowing for more precise targeting of treatment strategies tailored to specific COVID-19 survivor phenotypes.

METHODS

Study design

This study is a multicentre, parallel-group, randomized controlled trial (12). The original protocol for this study can be accessed from http://idmr.scu.edu.cn/info.htm?id=1841614474692833. The objective of the research is to compare the effectiveness of a 6-week unsupervised home exercise programme among COVID-19 recovered patients. The study was conducted at 3 hospitals in the Jiangsu and Hubei provinces of China, involving a total of 1,242 eligible COVID-19 patients, with approximately one-third (377 individuals) meeting the criteria. Participants, aged between 18 and 75, were discharged from the hospitals after undergoing COVID-19 treatment, with modified British Medical Research Council respiratory difficulty grades of 2–3. The research employed a smartphone application named "RehabApp" for a 6-week home exercise programme. Participants engaged in exercise routines 3–4 times a week, monitored through wearable chest-mounted heart rate telemetry devices. Teleconsultations with therapists were conducted weekly. The primary research outcome focused on post-treatment functional exercise capacity, assessed using the 6-Minute Walk Distance (6MWD) in accordance with guidelines from the European Respiratory Society and the American Thoracic Society. Secondary research outcomes included squat time, pulmonary function measurements (such as FEV1, forced vital capacity [FVC], FEV1/FVC, maximum voluntary ventilation [MVV], and peak expiratory flow [PEF]), assessments of health-related quality of life (utilizing the Short Form Health Survey-12 [SF-12] tool, encompassing physical and mental health scores), along with evaluations of perceived respiratory difficulty based on the modified Medical Research Council (mMRC) scale.

Statistical analysis

To determine the optimal number of latent classes within the dataset, we utilized Mplus v8 software (https://www.statmodel.com/), incorporating an extensive range of baseline data. This encompassed demographic features such as age and gender, along with a comprehensive set of baseline clinical data. The clinical data included disease severity, presence of comorbidities, smoking history, days from hospital discharge to baseline, days from baseline to post-treatment assessment, days from baseline to follow-up assessment, and other relevant variables. Furthermore, baseline pulmonary function metrics (FEV1, FVC, FEV1/FVC, MVV, PEF) and baseline quality of life indicators (SF-12 Physical Component Score [PCS] and SF-12 Mental Component Score [MCS]) were incorporated, among various other variables. Models ranging from 1 to 4 classes were assessed to determine the optimal number of classes.

For each model, a probability of belonging to each class was assigned to each subject. Ideally, this probability would approach 1.0 for a single class and trend toward zero for the others. To discern the most suitable number of latent classes, multiple factors were taken into consideration. These factors included (19–21): (*i*) Bayesian Information Criterion (BIC): a decreasing BIC number indicates an improved model fit. (*ii*) Distribution of Subjects Across Classes: a small count was deemed unlikely to represent a clinically significant subgroup. (*iii*) the Vuong-Lo-Mendell-Rubin (VLMR) test, assessing whether k classes fit the data better than k-1 classes. Entropy: values ≥ 0.80 denote commendable class separation. Through a comprehensive evaluation of these factors, we aimed to determine the optimal number of latent classes for robust and clinically meaningful insights.

Once the optimal number of categories was established, research participants were allocated to their most likely category. Subsequently, baseline features were compared using *t*-tests, Pearson χ^2 tests, or Wilcoxon rank-sum tests based on the nature of the variables. Subgroup analyses were conducted based on the use of corticosteroids. Outcome measures, such as 6MWD, were treated as dependent variables, with intervention status as the independent variable. Phenotype was considered as a moderating factor. The analysis aimed to assess the impact of different interventions on outcomes within distinct subgroups and tested for interaction between intervention measures and phenotypes. To enhance statistical power and address potential biases stemming from the exclusion of data with missing values, a technique called multivariate multiple imputation with chained equations was employed to impute the absent values (22). To ensure robustness, we repeated the analysis of the main outcome measures for comparison using data treated with multiple imputation.

Furthermore, diverse methods were employed to assess whether the categories identified by LCA could be defined by a finite set of biomarkers. Initially, least absolute shrinkage and selection operator (LASSO) regression and receiver operating characteristic (ROC) curve analysis to examine variables with the maximum average differences in predicting specified phenotypes were used to filter variables. This process served to effectively discern potential predictive factors for phenotypes. Ultimately, a limited set of biomarkers was utilized to construct clinical prediction models for different phenotypes (23). Evaluation of these prediction models involved the use of ROC curves (24), calibration curves, and clinical decision curves. Analyses other than LCA were carried out using R (version 4.1.0) and R Studio (version1.4.1717) (R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

Population characteristics

The baseline characteristics of patients enrolled in the telerehabilitation programme for the COVID-19 trial were partially described in the original publication. This study further conducted a more comprehensive analysis by stratifying baseline features according to gender, disease severity, and the use of corticosteroids. For detailed information, please refer to Tables SI–III.

Two-class model optimally fits the population

In this investigation, patients underwent stratification into 2, 3, and 4 subtypes, grounded on their medical information. Corresponding analyses of posterior probabilities are presented in Tables SIV–VI. Subsequent to this, a comparative analysis scrutinized the results derived from the aforementioned classification methods. As the number of classifications increased, a discernible decreasing trend in the BIC was observed, indicative of a gradual improvement in model fitting. All identified subtypes exhibited entropy values surpassing 0.9, with all values exceeding 0.8, thereby meeting acceptable criteria. Additionally, it was observed that, among the 3 classification methods, the smallest participant count assigned to a category was noted in the 4-class model, specifically in Class 2, with a precise count of 18 (15.1%). Proportions for all other categories exceeded 20%, with no notably undersized percentages. Vuong-LoMendell-Rubin tests yielded *p*-values < 0.0001, underscoring statistical significance. The aforementioned findings suggest that an excessive number of classes contribute marginally to the improvement of model fit. Instead, opting for 2-class subtypes not only maintains a robust model fit but also facilitates practical clinical application, as delineated in Table I. For simplicity, the 2 classes will be referred to as phenotypes 1 and 2, respectively.

Comparison of phenotypic features between subphenotypes

The characterization of 2 phenotypes is accomplished by assigning included patients to their most likely phenotypes and then examining demographic, clinical, and outcomes (baseline) characteristics of each phenotype. Fig. 1 illustrates a scale of continuous variables used in the analysis, ordered by the degree of separation between phenotypes. The variables with the greatest separation primarily relate to pulmonary function, with patients in Phenotype 1 exhibiting poorer lung function characteristics. Disparities in demographics data between the 2 phenotypes are evident, notably in the significantly higher age of Phenotype 1 compared with Phenotype 2, and a markedly lower proportion of males in Phenotype 1, both differences being statistically significant. No significant differences were observed between the 2 groups in other relevant data such as in-hospital treatment modalities, medical history (baseline), or trial information, as detailed in Table SVIII. The primary outcome measure in this study, the 6MWD, showed no significant difference

Table I. Fit statistics for latent class models from 1 to 4 classes

			Number assigned to each subphenotype				
N classes	BIC	Entropy	Phenotype 1, *n* (%)	Phenotype 2, *n* (%)	Phenotype 3, *n* (%)	Phenotype 4, *n* (%)	*p*-value
1	6089.40						
2	5724.57	0.92	63 (52.91%)	56 (47.13%)			< 0.001
3	5594.55	0.98	26 (21.80%)	40 (33.61%)	53 (44.53%)		< 0.001
4	5576.71	0.97	46 (38.78%)	18 (15.15%)	24 (20.28%)	31 (26.12%)	< 0.001

P-value represents the Vuong-Lo-Mendell-Rubin (VLMR) test, which tests whether k classes confer an improved model fit over k-1 classes.
Entropy: an index ranging between 0 and 1, which indicates separation of classes.
BIC: Bayesian information criteria, a measure of model fit.

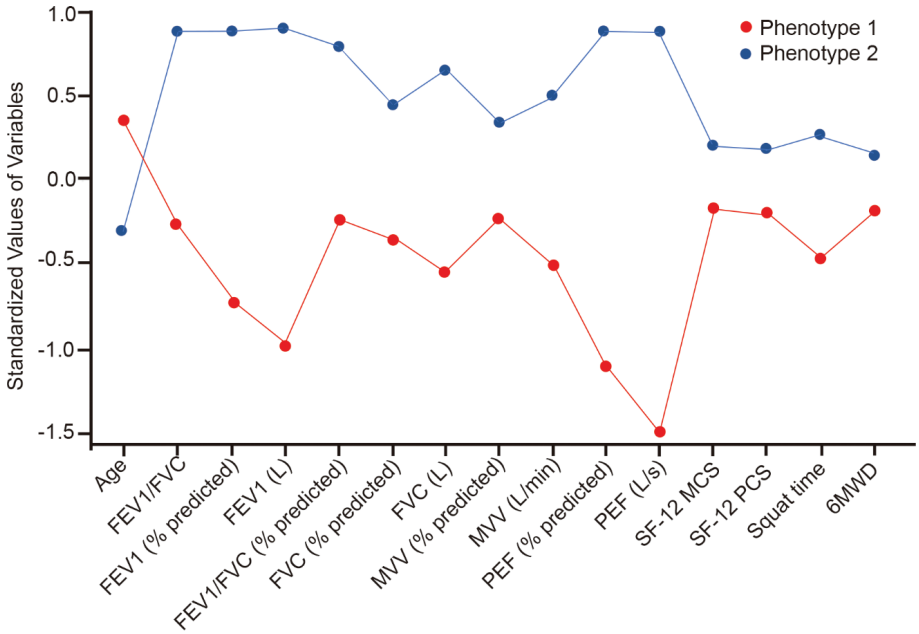

Fig. 1. Differences between phenotypes in the standardized values of continuous variables. The individual continuous variables underwent standardization by transforming them onto a z scale, characterized by a mean of zero and a standard deviation of 1. Standardized values for each phenotype denote their deviation from the mean within the entire cohort. The arrangement of variables from left to right corresponds to the extent of differentiation between phenotypes 1 and 2. Notably, higher standardized values in Phenotype 1 are situated on the left side of the graph, while lower standardized values are observed on the right side, highlighting the distinctive features between the 2 phenotypes. FVC: forced vital capacity; FEV1: forced expiratory volume in first second; MVV: maximum voluntary ventilation; PEF: peak expiratory flow; SF-12: Short Form Health Survey-12; MCS: mental component score; PCS: physical component score; 6MWD: 6-minute walk distance.

between Phenotype 1 and Phenotype 2 (p-value 0.078). Among other outcomes (baseline)-related variables, squat time, and pulmonary function indicators in Phenotype 1 were notably inferior to those in Phenotype 2, while remaining variables including quality of life and mMRC perceived dyspnoea showed no significant statistical differences between the 2 phenotypes. Refer to Table SVIII for detailed information.

Treatment effect in subphenotypes

Among the entire cohort under consideration, the therapeutic efficacy of the TERECO group, exemplified by the primary outcome (6MWD), notably surpasses that of the control group. This phenomenon persists even after meticulous adjustments for potential confounding factors. Employing phenotype as a modifier of the treatment effect, an interactive analysis between various intervention measures and distinct phenotypes reveals no discernible interaction between the 2. Furthermore, this study conducted a subgroup analysis based on the use of corticosteroids, revealing that within the subset of patients undergoing remote rehabilitation intervention while using corticosteroids, Phenotype 1 patients experience significantly greater benefits. Following adjustments for confounding factors, the improvement in the 6MWD for Phenotype 1 patients undergoing remote rehabilitation intervention in the corticosteroid subgroup is markedly higher than that of Phenotype 2. Specifically, the improvement in distance during the "Post-treatment (6 weeks)" period is 99.79 m (95% CI 38.08 to 161.49; $p = 0.0036$), and during the "Follow-up (~28 weeks)" period, it is 93.57 m (95% CI 22.85 to 164.30; $p = 0.0154$). Additionally, a significant interaction was identified between different intervention measures and distinct phenotypes among patients using corticosteroids. However, in Phenotype 2 patients, whether remote rehabilitation intervention was administered did not exhibit a significant difference. For detailed results of the comparative analysis of therapeutic outcomes among different groups 6 weeks post-baseline, refer to Table II, and for results 28 weeks post-baseline, consult Table III. This study exhibits some missing values, as indicated in Table SIX. To mitigate the potential reduction in statistical power and the introduction of bias resulting from the direct exclusion of missing values, we employed the method of multiple imputation to estimate these gaps, generating a total of 5 datasets. Upon replicating the aforementioned analytical procedures and conducting sensitivity analyses, no significant alterations in the research findings were observed. This suggests that the study results are relatively robust, as detailed in Tables SX and SXI.

Additionally, this study conducted an analysis of secondary outcome measures, such as squat time, pulmonary function, and quality of life. It was

Table II. Comparative analysis of primary outcome (6MWD) among different groups 6 weeks post-baseline

Estimates of treatment effects from different scenarios with 95% CI									
	Total			Treatment without corticosteroids			Treatment with corticosteroids		
Subgroup	β (95% CI)	*p*-value	*p*-value*	β (95% CI)	*p*-value	*p*-value*	β (95% CI)	*p*-value	*p* value*
Not adjusted			0.539			0.145			0.016
T and P 2	Ref.			Ref.			Ref.		
C and P 2	−55.30 (−93.01, −17.59)	0.005		−74.91 (−120.92, −28.89)	0.002		−31.27 (−87.81, 25.26)	0.285	
T and P 1	27.24 (−10.46, 64.94)	0.160		−33.72 (−80.44, 13.01)	0.162		104.00 (48.65, 159.35)	< 0.001	
C and P 1	−43.90 (−79.97, −7.83)	0.019		−63.15 (−106.54, −19.77)	0.006		−22.00 (−77.35, 33.35)	0.440	
Adjusted for age, sex			0.522			0.157			0.019
T and P 2	Ref.			Ref.			Ref.		
C and P 2	−54.78 (−93.20, −16.36)	0.006		−75.73 (−122.50, −28.97)	0.002		−32.35 (−90.59, 25.90)	0.283	
T and P 1	29.47 (−11.23, 70.18)	0.159		−28.59 (−78.18, 21.00)	0.263		98.83 (36.91, 160.75)	0.003	
C and P 1	−41.87 (−81.42, −2.32)	0.040		−60.45 (−106.88, −14.03)	0.013		−25.74 (−89.74, 38.26)	0.435	
Plus smoking history			0.571			0.079			0.008
T and P 2	Ref.			Ref.			Ref.		
C and P 2	−56.07 (−95.44, −16.69)	0.006		−86.79 (−133.90, −39.68)	0.001		−25.08 (−83.74, 33.59)	0.407	
T and P 1	28.45 (−12.90, 69.80)	0.180		−34.04 (−82.87, 14.78)	0.177		107.71 (45.01, 170.41)	0.002	
C and P 1	−42.52 (−82.44, −2.61)	0.039		−67.05 (−112.96, −21.15)	0.006		−23.27 (−86.76, 40.22)	0.477	
Plus disease severity			0.625			0.064			0.009
T and P 2	Ref.			Ref.			Ref.		
C and P 2	−57.57 (−97.20, −17.94)	0.005		−88.22 (−135.80, −40.64)	< 0.001		−26.73 (−86.36, 32.90)	0.385	
T and P 1	27.33 (−14.20, 68.86)	0.200		−35.85 (−85.27, 13.58)	0.161		106.93 (43.53, 170.34)	0.002	
C and P 1	−43.13 (−83.15, −3.11)	0.037		−66.63 (−112.80, −20.46)	0.006		−24.39 (−88.68, 39.89)	0.462	
Plus days from onset to admission			0.627			0.059			0.002
T and P 2	Ref.			Ref.			Ref.		
C and P 2	−59.00 (−98.76, −19.24)	0.005		−89.50 (−137.93, −41.07)	< 0.001		−20.79 (−77.39, 35.81)	0.476	
T and P 1	27.01 (−14.98, 69.00)	0.210		−37.79 (−89.00, 13.42)	0.154		112.98 (52.82, 173.13)	< 0.001	
C and P 1	−44.82 (−84.98, −4.65)	0.031		−67.55 (−114.47, −20.63)	0.007		−28.16 (−89.02, 32.70)	0.370	
Plus length of inpatient stay			0.709			0.059			0.004
T and P 2	Ref.			Ref.			Ref.		
C and P 2	−57.74 (−97.63, −17.86)	0.006		−88.88 (−137.96, −39.80)	< 0.001		−19.20 (−74.99, 36.60)	0.504	
T and P 1	27.07 (−14.96, 69.09)	0.210		−37.38 (−89.12, 14.37)	0.163		112.23 (52.96, 171.50)	< 0.001	
C and P 1	−40.57 (−81.81, 0.66)	0.057		−66.42 (−114.54, −18.30)	0.009		−15.75 (−77.98, 46.47)	0.623	
Plus days from hospital discharge to baseline			0.363			0.189			0.005
T and P 2	Ref.			Ref.			Ref.		
C and P 2	−54.58 (−92.00, −17.15)	0.005		−86.61 (−133.59, −39.63)	< 0.001		−19.34 (−73.77, 35.09)	0.491	
T and P 1	29.24 (−10.17, 68.64)	0.149		−27.90 (−77.96, 22.16)	0.280		102.37 (43.41, 161.33)	0.002	
C and P 1	−48.01 (−86.84, −9.17)	0.017		−74.28 (−120.72, −27.84)	0.003		−18.25 (−79.03, 42.53)	0.560	
Plus diabetes			0.468			0.175			0.006
T and P 2	Ref.			Ref.			Ref.		
C and P 2	−58.10 (−95.63, −20.58)	0.003		−87.28 (−134.54, −40.02)	< 0.001		−18.54 (−74.95, 37.87)	0.524	
T and P 1	27.21 (−12.08, 66.50)	0.178		−28.65 (−79.01, 21.71)	0.270		102.80 (42.67, 162.93)	0.002	
C and P 1	−48.92 (−87.57, −10.28)	0.015		−74.44 (−121.13, −27.76)	0.003		−18.19 (−79.84, 43.47)	0.567	
Plus heart disease			0.455			0.174			0.002
T and P 2	Ref.			Ref.			Ref.		
C and P 2	−59.85 (−97.47, −22.23)	0.002		−87.38 (−135.13, −39.63)	< 0.001		−29.27 (−82.18, 23.64)	0.286	
T and P 1	26.22 (−13.08, 65.51)	0.194		−28.84 (−79.81, 22.14)	0.273		99.81 (44.05, 155.57)	0.001	
C and P 1	−52.05 (−91.07, −13.03)	0.010		−74.67 (−122.00, −27.34)	0.003		−36.04 (−94.77, 22.69)	0.238	
Plus lung disease (including inactive TB)			0.418			0.276			0.002
T and P 2	Ref.			Ref.			Ref.		
C and P 2	−59.75 (−97.49, −22.00)	0.003		−83.21 (−130.90, −35.53)	0.001		−26.32 (−80.19, 27.55)	0.345	
T and P 1	26.09 (−13.33, 65.51)	0.198		−27.36 (−77.92, 23.20)	0.294		100.96 (44.72, 157.21)	0.001	
C and P 1	−53.72 (−93.21, −14.24)	0.009		−77.42 (−124.49, −30.35)	0.002		−31.48 (−91.87, 28.92)	0.315	
Plus obesity			0.469			0.277			0.002
T and P 2	Ref.			Ref.			Ref.		
C and P 2	−60.51 (−98.20, −22.83)	0.002		−86.48 (−133.55, −39.40)	< 0.001		−26.90 (−82.11, 28.30)	0.347	
T and P 1	24.18 (−15.28, 63.64)	0.233		−26.71 (−76.46, 23.04)	0.298		101.49 (43.98, 159.00)	0.002	
C and P 1	−54.18 (−93.59, −14.77)	0.008		−80.95 (−127.45, −34.44)	0.001		−32.03 (−93.77, 29.72)	0.317	
Plus hypertension			0.429			0.375			0.002
T and P 2	Ref.			Ref.			Ref.		
C and P 2	−58.61 (−96.72, −20.50	0.003		−79.96 (−127.47, −32.44)	0.002		−26.43 (−82.88, 30.02)	0.366	
T and P 1	25.48 (−14.23, 65.18)	0.212		−21.00 (−70.92, 28.93)	0.414		101.83 (43.21, 160.45)	0.002	
C and P 1	−52.63 (−92.35, −12.92)	0.011		−74.98 (−121.79, −28.17)	0.003		−32.12 (−94.88, 30.63)	0.324	
Plus other comorbidity			0.393			0.428			0.003
T and P 2	Ref.			Ref.			Ref.		
C and P 2	−56.45 (−94.75, −18.15)	0.005		−79.49 (−127.33, −31.65)	0.002		−25.12 (−83.40, 33.15)	0.405	
T and P 1	23.84 (−15.96, 63.63)	0.243		−21.37 (−71.63, 28.88)	0.409		99.79 (38.08, 161.49)	0.004	
C and P 1	−53.57 (−93.30, −13.84)	0.010		−77.56 (−125.36, −29.77)	0.003		−31.05 (−95.37, 33.28)	0.352	

**P*-value for interaction test: examining the interaction between treatment groups (TERECO vs Control) and subphenotypes.
CI: confidence interval; Ref: reference; T: TERECO; P: phenotype; C: control; TB: tuberculosis; 6MWD: 6-minute walk distance.

Table III. Comparative analysis of primary outcome (6MWD) among different groups 28 weeks post-baseline

Estimates of treatment effects from different scenarios with 95% CIs

	Total			Treatment without corticosteroids			Treatment with corticosteroids		
Subgroup	β (95% CI)	*p*-value	*p*-value*	β (95% CI)	*p*-value	*p*-value*	β (95% CI)	*p*-value	*p*-value*
Not adjusted			0.810			0.125			0.030
T and P 2	Ref.			Ref.			Ref.		
Control and P 2	−66.77 (−109.37, −24.17)	0.003		−91.67 (−147.20, −36.14)	0.002		−40.65 (−100.56, 19.26)	0.191	
T and P 1	15.45 (−26.74, 57.63)	0.475		−45.69 (−100.32, 8.94)	0.107		95.65 (35.74, 155.56)	0.003	
Control and P 1	−58.32 (−98.52, −18.12)	0.005		−80.62 (−131.99, −29.26)	0.003		−38.07 (−96.47, 20.34)	0.209	
Adjusted for age, sex			0.824			0.124			0.033
T and P 2	Ref.			Ref.			Ref.		
Control and P 2	−65.89 (−108.89, −22.90)	0.003		−91.98 (−148.34, −35.63)	0.002		−39.57 (−100.52, 21.38)	0.211	
T and P 1	18.62 (−26.89, 64.13)	0.425		−40.71 (−97.88, 16.46)	0.168		101.77 (32.09, 171.46)	0.007	
Control and P 1	−53.75 (−97.37, −10.14)	0.018		−76.12 (−130.52, −21.71)	0.008		−28.24 (−95.45, 38.98)	0.416	
Plus smoking history			0.807			0.091			0.018
T and P 2	Ref.			Ref.			Ref.		
Control and P 2	−65.29 (−108.94, −21.64)	0.004		−98.31 (−155.22, −41.39)	0.001		−33.49 (−94.46, 27.49)	0.289	
T and P 1	19.09 (−26.89, 65.07)	0.418		−43.49 (−100.53, 13.55)	0.141		109.02 (39.24, 178.80)	0.004	
Control and P 1	−53.38 (−97.37, −9.38)	0.019		−79.89 (−134.32, −25.45)	0.006		−24.35 (−91.11, 42.42)	0.479	
Plus disease severity			0.793			0.090			0.018
T and P 2	Ref.			Ref.			Ref.		
Control and P 2	−64.79 (−108.66, −20.93)	0.005		−98.62 (−156.27, −40.98)	0.002		−33.24 (−94.66, 28.18)	0.296	
T and P 1	18.98 (−27.19, 65.15)	0.422		−43.54 (−101.10, 14.02)	0.144		107.71 (37.33, 178.10)	0.005	
Control and P 1	−53.48 (−97.66, −9.31)	0.020		−79.78 (−134.74, −24.83)	0.006		−24.62 (−91.87, 42.64)	0.478	
Plus days from onset to admission			0.820			0.066			0.006
T and P 2	Ref.			Ref.			Ref.		
Control and P 2	−65.84 (−109.15, −22.53)	0.004		−100.62 (−157.75, −43.49)	0.001		−27.41 (−86.55, 31.72)	0.370	
T and P 1	15.53 (−30.67, 61.73)	0.512		−49.11 (−107.36, 9.14)	0.105		105.82 (38.34, 173.30)	0.004	
Control and P 1	−56.91 (−100.63, −13.19)	0.012		−82.45 (−136.96, −27.95)	0.005		−31.81 (−96.64, 33.01)	0.343	
Plus length of inpatient stay			0.909			0.068			0.009
T and P 2	Ref.			Ref.			Ref.		
Control and P 2	−65.08 (−108.43, −21.74)	0.004		−97.68 (−155.05, −40.31)	0.002		−27.48 (−87.48, 32.53)	0.376	
T and P 1	16.06 (−30.17, 62.28)	0.498		−45.82 (−104.37, 12.73)	0.131		105.52 (36.96, 174.08)	0.005	
Control and P 1	−52.35 (−97.02, −7.68)	0.024		−77.51 (−132.79, −22.23)	0.008		−30.33 (−98.46, 37.80)	0.389	
Plus days from hospital discharge to baseline			0.613			0.150			0.011
T and P 2	Ref.			Ref.			Ref.		
Control and P 2	−60.20 (−101.04, −19.36)	0.005		−94.84 (−149.65, −40.03)	0.001		−26.09 (−84.69, 32.51)	0.389	
T and P 1	15.29 (−28.17, 58.76)	0.492		−40.54 (−96.58, 15.51)	0.163		94.79 (26.61, 162.97)	0.010	
Control and P 1	−58.66 (−100.80, −16.53)	0.008		−85.68 (−138.85, −32.51)	0.003		−32.78 (−99.36, 33.80)	0.342	
Plus diabetes			0.678			0.136			0.012
T and P 2	Ref.			Ref.			Ref.		
Control and P 2	−62.37 (−103.46, −21.27)	0.004		−95.87 (−151.03, −40.70)	0.001		−26.78 (−86.86, 33.30)	0.389	
T and P 1	14.53 (−28.98, 58.04)	0.514		−41.55 (−97.96, 14.86)	0.155		94.96 (25.68, 164.23)	0.012	
Control and P 1	−59.13 (−101.30, −16.97)	0.007		−85.96 (−139.41, −32.51)	0.003		−32.62 (−100.26, 35.02)	0.352	
Plus heart disease			0.739			0.111			0.007
T and P 2	Ref.			Ref.			Ref.		
Control and P 2	−66.53 (−107.57, −25.49)	0.002		−98.68 (−154.26, −43.10)	0.001		−39.91 (−95.51, 15.68)	0.170	
T and P 1	12.63 (−30.55, 55.82)	0.568		−43.87 (−100.58, 12.84)	0.136		89.30 (26.01, 152.59)	0.001	
Control and P 1	−62.83 (−104.86, −20.79)	0.004		−87.86 (−141.54, −34.17)	0.002		−47.86 (−110.52, 14.79)	0.145	
Plus lung disease (including inactive TB)			0.715			0.143			0.007
T and P 2	Ref.			Ref.			Ref.		
Control and P 2	−67.19 (−108.38, −25.99)	0.002		−96.68 (−151.96, −41.40)	0.001		−38.59 (−95.75, 18.58)	0.196	
T and P 1	12.26 (−31.06, 55.58)	0.581		−42.94 (−99.28, 13.40)	0.142		89.98 (25.54, 154.42)	0.011	
Control and P 1	−64.71 (−107.17, −22.24)	0.004		−90.00 (−143.42, −36.58)	0.002		−45.81 (−110.93, 19.32)	0.179	
Plus obesity			0.770			0.167			0.007
T and P 2	Ref.			Ref.			Ref.		
Control and P 2	−68.84 (−109.78, −27.90)	0.001		−99.83 (−154.31, −45.36)	< 0.001		−38.83 (−97.05, 19.38)	0.202	
T and P 1	9.83 (−33.28, 52.94)	0.656		−40.40 (−95.87, 15.07)	0.160		91.30 (24.01, 158.59)	0.013	
Control and P 1	−66.75 (−108.98, −24.53)	0.003		−94.60 (−147.41, −41.80)	0.001		−45.72 (−111.97, 20.53)	0.187	
Plus hypertension			0.754			0.240			0.011
T and P 2	Ref.			Ref.			Ref.		
Control and P 2	−68.02 (−110.00, −26.05)	0.002		−90.99 (−146.29, −35.69)	0.002		−43.80 (−103.73, 16.13)	0.164	
T and P 1	10.09 (−33.33, 53.52)	0.650		−34.87 (−90.31, 20.56)	0.224		92.31 (24.50, 160.11)	0.013	
Control and P 1	−66.25 (−109.01, −23.49)	0.003		−87.61 (−140.77, −34.45)	0.002		−44.41 (−111.21, 22.38)	0.204	
Plus other comorbidity			0.667			0.375			0.011
T and P 2	Ref.			Ref.			Ref.		
Control and P 2	−64.52 (−106.44, −22.60)	0.003		−89.19 (−143.16, −35.22)	0.002		−44.99 (−107.69, 17.71)	0.171	
T and P 1	8.58 (−34.57, 51.74)	0.698		−34.18 (−88.25, 19.89)	0.222		93.57 (22.85, 164.30)	0.015	
Control and P 1	−67.27 (−109.73, −24.80)	0.003		−95.29 (−147.79, −42.79)	< 0.001		−45.52 (−114.84, 23.80)	0.209	

**P*-value for interaction test: examining the interaction between treatment groups (TERECO vs Control) and subphenotypes.
CI: confidence interval; Ref: reference; T: TERECO; P: phenotype; C: control; TB: tuberculosis; 6MWD: 6-minute walk distance.

observed that among patients with Phenotype 1 who underwent remote rehabilitation intervention using corticosteroids, there was a significant improvement in pulmonary function indicators, including FEV1 (L) and FEV1/FVC, at the 28-week follow-up, compared with Phenotype 2 patients. Furthermore, a significant interaction effect was found between different phenotypes and the utilization of remote rehabilitation intervention. No such patterns were detected in other secondary outcome measures. For further details, please refer to Tables SXII–XXIII.

Construction of clinical predictive models of subphenotypes

This study endeavours to establish a predictive model discerning between 2 phenotypes. Given the multitude of independent variables in this research, the LASSO regression method, incorporating a LASSO penalty term, facilitates the transformation of smaller-weight coefficients to zero, thereby achieving feature selection and model sparsity. Consequently, we initially employ the LASSO regression model to screen for predictive factors, resulting in the inclusion of a total of 18 predictive factors, as depicted in Fig. S1. Although the use of the aforementioned variables demonstrates excellent diagnostic performance in predicting different phenotypes, with an AUC of 0.989 (95% CI 0.975 to 1.00), as illustrated in Fig. S2, the method involves too many variables, rendering it less conducive to clinical application.

To assess whether these phenotypes can robustly be identified by 2 or 3 variables, we utilize ROC curve analysis to examine variables with the maximum average differences in predicting specified phenotypes. The results incorporate FEV1, FVC, and PEF – 3 pulmonary function indicators – to construct the predictive model. The ROC curve and calibration curve both exhibit excellent diagnostic capability, with an AUC of 0.971 (95% CI 0.948 to 0.994), as shown in Fig. S3. Furthermore, employing FEV1 and PEF as 2 pulmonary function indicators for predicting different phenotypes also demonstrates robust diagnostic performance, with an AUC of 0.936 (95% CI 0.936 to 0.991), detailed in Fig. S4. The final predictive model and clinical decision curve are illustrated comprehensively in Fig. S5.

DISCUSSION

This study, employing the LCA method, has yielded 2 pivotal discoveries that hold substantial implications for future clinical trials related to the rehabilitation of dyspnoea COVID-19 survivors. Primarily, within the subset of post-discharge COVID-19 survivors experiencing persistent respiratory difficulties, we have identified, for the first time, 2 distinct clinical phenotypes. Each phenotype presents unique clinical characteristics, with the most salient differentiating variables, post-screening and refinement of relevant factors, predominantly manifesting in pulmonary function. Specifically, Phenotype 1 represents a subset of COVID-19 survivors with compromised baseline pulmonary function, while Phenotype 2 comprises individuals exhibiting superior baseline pulmonary function. Moreover, of greater significance is the disparate response of these 2 subtypes of COVID-19 survivors to the intervention measures of randomly assigned TERECO. Notably, within the cohort without corticosteroid intervention, there was no statistically significant difference in the response to TERECO intervention between Phenotype 1 and Phenotype 2. However, within the corticosteroid-intervened cohort, a notable divergence in sensitivity to TERECO intervention was observed between Phenotype 1 and Phenotype 2, with Phenotype 1 demonstrating greater benefits. Although TERECO intervention in Phenotype 2 showed improvements, it did not achieve statistical significance. These findings underscore the critical importance of identifying subtypes among COVID-19 survivors for future clinical trials. The results of this study emphasize the necessity of implementing remote rehabilitation training, such as TERECO, for COVID-19 survivors experiencing moderate respiratory difficulties post-discharge. This is particularly relevant for individuals exhibiting the specific phenotype of compromised baseline lung function, as they stand to derive greater benefits from such interventions. The interaction between phenotypes and the efficacy of rehabilitation interventions necessitates careful consideration in the design and implementation of future clinical trials in this context.

Numerous diseases present a complex spectrum of symptoms, leading to significant heterogeneity among affected individuals. This intricacy not only complicates disease diagnosis but also poses challenges in tailoring personalized treatment strategies. By employing LCA, we can attain a comprehensive understanding of the characteristics of distinct patient cohorts, thereby laying solid groundwork for precision medicine and individualized treatment approaches (25). The concept of tailoring therapies based on biomarker-defined subgroups within heterogeneous syndromes, a fundamental aspect of precision medicine, has notably revolutionized patient care within the field of oncology. For instance, this methodology extends from considering the oestrogen receptor status in breast cancer to assessing the BRAF mutation status in various malignancies, notably melanoma (26). Recently, there has been a growing body of research on pheno-

types related to COVID-19 (27–29). However, studies on the phenotypes of individuals, particularly those experiencing respiratory difficulties post-discharge from COVID-19, remain deficient. Over 3 years have elapsed since the onset of the COVID-19 pandemic, and there is mounting concern among patients, clinicians, and scientists regarding the long-term sequelae of hospitalized COVID-19 patients. Many survivors continue to manifest persistent symptoms in the months following infection. Hence, there is a pressing need to investigate whether subtypes exist among this distinctive population of COVID-19 survivors. If such phenotypes are discernible, it becomes crucial to ascertain whether there are prognostic differences among them and whether these phenotypes exhibit variations in sensitivity to treatment.

Previously, there has been research utilizing respiratory measurement parameters to conduct a phenotypic study on patients with COVID-19-related acute respiratory distress syndrome (ARDS) (30). It is important to note that a practical limitation of the aforementioned subtyping method is its reliance on CT scans, introducing certain constraints. Due to resource limitations or difficulties in transporting hypoxic patients, obtaining CT scans may prove challenging. However, in terms of outcomes, the discrimination of the mentioned phenotypes requires fewer variables to achieve higher accuracy, underscoring its significant guiding implications. In contrast, our study relies primarily on discerning variables related to pulmonary function, wherein we employed a portable pulmonary function device for practical assessments. In practical application, this portable tool eliminates the complexities associated with transportation issues. Furthermore, for ease of practical application, our study ultimately narrowed down the variables to FEV1 and PEF, 2 indicators that easily identify subphenotypes with high accuracy. An additional study employed longitudinal LCA to conduct a phenotypic investigation of COVID-19-related ARDS (31). This research identified 2 distinct phenotypes with significant differences in prognostic indicators. In our study, we did not employ the longitudinal LCA method. Instead, we assert that utilizing baseline data to promptly identify distinct phenotypes and subsequently implementing personalized interventions for different phenotypic groups is of utmost importance.

Our research, following subcomponent analysis, revealed significant variations in treatment outcomes within the subgroup of whether corticosteroids were used. Specifically, in the subgroup without corticosteroids, TERECO demonstrated a significant intervention effect on COVID-19 survivors, unaffected by phenotype. However, contrasting results emerged in the corticosteroid-use subgroup: TERECO intervention showed no statistically significant impact on Phenotype 2 individuals, yet significant disparities surfaced among different phenotypes within the TERECO intervention group, with Phenotype 1 individuals benefiting notably more than Phenotype 2. We posit that the observed phenomenon may be attributed to differences in the composition of populations with distinct phenotypes: in the non-corticosteroid group, "Not severe" accounted for a higher proportion at 92.86%, whereas in the corticosteroid group, "Severe" had a higher percentage at 67.35%. Could the occurrence of such phenomena be linked to the severity of the disease? Even after adjusting for disease severity as a covariate to address confounding factors, however, the aforementioned patterns persisted. In this clinical trial, the lack of detailed descriptions of corticosteroid types and doses hinders a more in-depth subcomponent analysis. Currently, there are apparent differences in the therapeutic effects of corticosteroids on COVID-19, where low-dose corticosteroids can reduce the mortality rate of COVID-19 patients requiring oxygen or respiratory support (non-invasive mechanical ventilation, invasive mechanical ventilation, or extracorporeal membrane oxygenation), while higher doses significantly increase mortality (32). In summary, further clinical trials are necessary to design interventions for COVID-19 survivors, considering different types and doses of corticosteroids in conjunction with TERECO, specific phenotypes, and varying degrees of disease severity.

Strengths and limitations

However, the study is not without limitations. First, in terms of participant inclusion, it focused exclusively on COVID-19 survivors with moderate dyspnoea who had undergone prior inpatient treatment. Consequently, the findings may not be applicable to individuals with mild or severe dyspnoea, or to those infected with SARS-CoV-2 who did not require hospitalization. The exclusion of patients using β-blockers or bronchodilators further narrows the generalizability of specific conclusions. Moreover, due to constraints in the sample size, the trial did not conduct more extensive subgroup analyses, leaving the magnitude of effects within specific subgroups unclear. Additionally, despite efforts to adjust for as many confounding factors as possible, the presence of unmeasured confounders could potentially impact the study's conclusions. Lastly, the follow-up duration in this study was limited to 24 weeks, which may be insufficient to fully capture the long-term recovery outcomes, especially given the potential prolonged effects of long COVID. Despite these limitations, our study also possesses several strengths that enhance its scientific and clinical relevance. First, the use of Latent Class Analysis enabled the identification of distinct COVID-19 survivor phenotypes, allowing for a nuan-

ced understanding of heterogeneity in post-recovery responses. Second, the application of a multicentre, randomized controlled trial design ensured robust data collection and generalizability. Third, the development of a simplified predictive model using key pulmonary function indicators enhances the clinical feasibility of phenotype-based stratification. These strengths collectively reinforce the study's contribution to personalized rehabilitation strategies for COVID-19 survivors.

Conclusion

This study unveils 2 distinct phenotypes among survivors of COVID-19 experiencing dyspnoea. Crucially, these subtypes may exhibit significant interactions with TERECO intervention measures. Importantly for clinical applications, these differences are poised to inform personalized treatment decisions for dyspnoea COVID-19 survivors. Furthermore, the identification of diverse phenotypes through pulmonary function variables provides potentially valuable information for selecting patients for future clinical trials. Nevertheless, additional research is imperative to validate these phenotypes and delve deeper into their mechanisms of response to various therapeutic interventions, thus offering enhanced guidance for clinical practices.

ACKNOWLEDGEMENTS

The authors express gratitude to all patients who participated in this research. Their appreciation extends to Jian an Li and the entire research team for their contributions to the design and implementation of the experiment. Additionally, the authors highly appreciate their transparency in sharing the data from this study.

Funding/ financial support: This work was supported by Research on Key Technologies for Prevention and Control of Novel Coronavirus Pneumonia (2021B03003-2).

Data availability statement: Anonymised patient-level data on which the analysis, results, and conclusions reported in this paper are based are available as a Dryad dataset. Reinhardt, Jan D. (2021), Telerehabilitation program for COVID-19 survivors (TERECO) – randomized controlled trial, Dryad, Dataset, https://doi.org/10.5061/dryad.59zw3r27n.

Ethics approval and informed consent: This study was registered at the Chinese Clinical Trial Registry on 11 April 2020 (ChiCTR2000031834). Ethical approval was first received from the Institutional Review Board (IRB) of the First Affiliated Hospital of Nanjing Medical University/Jiangsu Province Hospital (2020-SR-171, 9 April 2020) and then subsequently from the IRBs of Hubei Province Hospital of Integrated Chinese and Western Medicine (2020016, 14 April 2020), and Huangshi Hospital of Chinese Medicine (HSZYPJ-2020-026-01, 20 April 2020). Before filling out the questionnaire, the study participants were informed of the purpose and process of the study. They were informed of the right to withdraw from the study at any time during the survey process. Consent was obtained from the study participants.

The authors have no conflict of interest to declare.

REFERENCES

1. O'Mahoney LL, Routen A, Gillies C, Ekezie W, Welford A, Zhang A, et al. The prevalence and long-term health effects of Long Covid among hospitalised and non-hospitalised populations: a systematic review and meta-analysis. EClinicalMedicine 2023; 55: 101762. https://doi.org/10.1016/j.eclinm.2022.101762
2. Visan I. Long COVID. Nat Immunol 2021; 22: 934–935. https://doi.org/10.1038/s41590-021-00992-4
3. Nehme M, Braillard O, Chappuis F, Courvoisier DS, Guessous I. Prevalence of symptoms more than seven months after diagnosis of symptomatic COVID-19 in an outpatient setting. Ann Intern Med 2021; 174: 1252–1260. https://doi.org/10.7326/M21-0878
4. Regmi B, Friedrich J, Jörn B, Senol M, Giannoni A, Boentert M, et al. Diaphragm muscle weakness might explain exertional dyspnea 15 months after hospitalization for COVID-19. Am J Respir Crit Care Med 2023; 207: 1012–1021. https://doi.org/10.1164/rccm.202206-1243OC
5. Kostandova N, Schluth C, Arambepola R, Atuhaire F, Bérubé S, Chin T, et al. A systematic review of using population-level human mobility data to understand SARS-CoV-2 transmission. Nat Commun 2024; 15: 10504. https://doi.org/10.1038/s41467-024-54895-7
6. Hill NS. In persistent dyspnea after COVID-19 ARDS, exercise training rehabilitation vs. usual PT reduced dyspnea at 90 d. Ann Intern Med 2023; 176: Jc117. https://doi.org/10.7326/J23-0073
7. McCarthy B, Casey D, Devane D, Murphy K, Murphy E, Lacasse Y. Pulmonary rehabilitation for chronic obstructive pulmonary disease. Cochrane Database Syst Rev 2015; 2015: Cd003793. https://doi.org/10.1002/14651858.CD003793.pub3
8. Lau HM, Ng GY, Jones AY, Lee EW, Siu EH, Hui DS. A randomised controlled trial of the effectiveness of an exercise training program in patients recovering from severe acute respiratory syndrome. Aust J Physiother 2005; 51: 213–219. https://doi.org/10.1016/S0004-9514(05)70002-7
9. Singh SJ, Baldwin MM, Daynes E, Evans RA, Greening NJ, Jenkins RG, et al. Respiratory sequelae of COVID-19: pulmonary and extrapulmonary origins, and approaches to clinical care and rehabilitation. Lancet Respir Med 2023; 11: 709–725. https://doi.org/10.1016/S2213-2600(23)00159-5
10. Vieira A, Pinto A, Garcia B, Eid RAC, Mól CG, Nawa RK. Telerehabilitation improves physical function and reduces dyspnoea in people with COVID-19 and post-COVID-19 conditions: a systematic review. J Physiother 2022; 68: 90–98. https://doi.org/10.1016/j.jphys.2022.03.011
11. Sonza A, da Cunha de Sá-Caputo D, Sartorio A, Tamini S, Seixas A, Sanudo B, et al. COVID-19 lockdown and the behavior change on physical exercise, pain and psychological well-being: an international multicentric study. Int J Environ Res Public Health 2021; 18: 3810. https://doi.org/10.3390/ijerph18073810
12. Li J, Xia W, Zhan C, Liu S, Yin Z, Wang J, et al. A telerehabilitation programme in post-discharge COVID-19 patients (TERECO): a randomised controlled trial. Thorax 2022; 77: 697–706. https://doi.org/10.1136/thoraxjnl-2021-217382
13. Mori M, Krumholz HM, Allore HG. Using Latent Class Analysis to identify hidden clinical phenotypes. JAMA 2020; 324: 700–701. https://doi.org/10.1001/jama.2020.2278
14. Dahmer MK, Yang G, Zhang M, Quasney MW, Sapru A, Weeks HM, et al. Identification of phenotypes in paediatric patients with acute respiratory distress syndrome: a latent class analysis. Lancet Respir Med 2022; 10: 289–297. https://doi.org/10.1016/S2213-2600(21)00382-9
15. Sinha P, Delucchi KL, Thompson BT, McAuley DF, Matthay MA, Calfee CS. Latent class analysis of ARDS subphenotypes: a secondary analysis of the statins for acutely injured lungs from sepsis (SAILS) study. Intensive Care Med 2018; 44: 1859–1869. https://doi.org/10.1007/s00134-018-5378-3

16. Menéndez R, Méndez R, Latorre A, González-Jiménez P, Peces-Barba G, Molina-Molina M, et al. Clustering patients with COVID-19 according to respiratory support requirements, and its impact on short- and long-term outcome (RECOVID study). Pulmonology 2025; 31: 2442175. https://doi.org/10.1080/25310429.2024.2442175
17. Melo-Oliveira ME, Sá-Caputo D, Bachur JA, Paineiras-Domingos LL, Sonza A, Lacerda AC, et al. Reported quality of life in countries with cases of COVID19: a systematic review. Expert Rev Respir Med 2021; 15: 213–220. https://doi.org/10.1080/17476348.2021.1826315
18. Sañudo B, Seixas A, Gloeckl R, Rittweger J, Rawer R, Taiar R, et al. Potential application of whole body vibration exercise for improving the clinical conditions of COVID-19 infected individuals: a narrative review from the World Association of Vibration Exercise Experts (WAVex) Panel. Int J Environ Res Public Health 2020; 17: 3650. https://doi.org/10.3390/ijerph17103650
19. Lo YT, Mendell NR, Rubin DB. Testing the number of components in a normal mixture. Biometrika 2001: 88: 767–778. https://doi.org/10.1093/biomet/88.3.767
20. Schwarz GE. Estimating the dimension of a model. Ann Stat 1978; 6. https://doi.org/10.1214/aos/1176344136
21. Nylund KL, Asparouhov T, Muthén BO. Deciding on the number of classes in latent class analysis and growth mixture modeling: a Monte Carlo simulation study. Struct Equ Modeling 2007; 14: 535–569. https://doi.org/10.1080/10705510701575396
22. Li P, Stuart EA, Allison DB. Multiple imputation: a flexible tool for handling missing data. JAMA 2015; 314: 1966–1967. https://doi.org/10.1001/jama.2015.15281
23. Balachandran VP, Gonen M, Smith JJ, DeMatteo RP. Nomograms in oncology: more than meets the eye. Lancet Oncol 2015; 16: e173–180. https://doi.org/10.1016/S1470-2045(14)71116-7
24. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. J Thorac Oncol 2010; 5: 1315–1316. https://doi.org/10.1097/JTO.0b013e3181ec173d
25. Sinha P, Calfee CS, Delucchi KL. Practitioner's guide to latent class analysis: methodological considerations and common pitfalls. Crit Care Med 2021; 49: e63–e79. https://doi.org/10.1097/CCM.0000000000004710
26. Rugo HS, Olopade OI, DeMichele A, Yau C, van 't Veer LJ, Buxton MB, et al. Adaptive randomization of veliparib–carboplatin treatment in breast cancer. N Engl J Med 2016; 375: 23–34. https://doi.org/10.1056/NEJMoa1513749
27. Gutiérrez-Gutiérrez B, Del Toro MD, Borobia AM, Carcas A, Jarrín I, Yllescas M, et al. Identification and validation of clinical phenotypes with prognostic implications in patients admitted to hospital with COVID-19: a multicentre cohort study. Lancet Infect Dis 2021; 21: 783–792. https://doi.org/10.1016/S1473-3099(21)00019-0
28. Azoulay E, Zafrani L, Mirouse A, Lengliné E, Darmon M, Chevret S. Clinical phenotypes of critically ill COVID-19 patients. Intensive Care Med 2020; 46: 1651–1652. https://doi.org/10.1007/s00134-020-06120-4
29. Nakamura H, Miyagi K, Otsuki M, Higure Y, Nishiyama N, Kinjo T, et al. Serum KL-6 can distinguish between different phenotypes of severe COVID-19. J Med Virol 2021; 93: 158–160. https://doi.org/10.1002/jmv.26268
30. Filippini DFL, Di Gennaro E, van Amstel RBE, Beenen LFM, Grasso S, Pisani L, et al. Latent class analysis of imaging and clinical respiratory parameters from patients with COVID-19-related ARDS identifies recruitment subphenotypes. Crit Care 2022; 26: 363. https://doi.org/10.1186/s13054-022-04251-2
31. Bos LDJ, Sjoding M, Sinha P, Bhavani SV, Lyons PG, Bewley AF, et al. Longitudinal respiratory subphenotypes in patients with COVID-19-related acute respiratory distress syndrome: results from three observational cohorts. Lancet Respir Med 2021; 9: 1377–1386. https://doi.org/10.1016/S2213-2600(21)00365-9
32. RECOVERY Collaborative Group. Higher dose corticosteroids in patients admitted to hospital with COVID-19 who are hypoxic but not requiring ventilatory support (RECOVERY): a randomised, controlled, open-label, platform trial. Lancet 2023; 401: 1499–1507. https://doi.org/10.1016/S0140-6736(23)00510-X