Latent class growth analysis of COVID-19 pneumonia patients with longitudinal radiologic data for predicting severity and prognosis

Yeonju Seo[1*], Dongwon Ahn[2*], Teawan Goo[1], Jibong Jeong[2**], and Taesung Park[3**]

[1] *Interdisciplinary Program in Bioinformatics, Seoul National University, Seoul, Republic of Korea*

[2] *Department of Internal Medicine, Seoul Metropolitan Government Seoul National University Boramae Medical Center, Seoul National University College of Medicine, Seoul, Republic of Korea*

[3] *Department of Statistics, Seoul National University, Seoul, Republic of Korea*

*These authors have contributed equally to this work

**Corresponding author: jibjeong@snu.ac.kr, tspark@stats.snu.ac.kr

The COVID-19 pandemic originated in Wuhan, China, in December 2019 and quickly spread worldwide. Previous studies on severe disease risk factors mostly used one-time assessments, which may not fully capture the changing symptom of the disease. Repeated measurements of the extent of pneumonia can effectively capture the changing symptom in COVID-19. Understanding the relationship between these changes and clinical variables could help assess risk, guide treatment, and allocate healthcare resources efficiently. This study aims to identify the class of COVID-19 patients with different changes of pneumonia extent and understand the clinical characteristics of each class by comparing clinical outcomes, initial clinical variables, and laboratory variables. We identified six optimal classes of patients through latent class growth modeling, a type of group-based trajectory model in terms of Bayesian Information Criteria. After identifying the latent classes, significant differences for each class were analyzed using the Kruskal-Wallis test for continuous variables and Chi-squared/Fisher exact tests for categorical variables. As a result, significant differences were observed in clinical outcomes, including ICU admission (p-value: 1.30E-97) and death (p-value: 2.01E-27), as well as treatment variables such as High Flow Nasal Cannula (p-value: 1.29E-62). Within each class, notable changes in clinical and laboratory variables were found, such as Respiration Rate (p-value: 6.79E-14), LDHL (p-value: 6.60E-46), D-dimer (p-value: 1.23E-33), and Troponin-I (p-value: 9.72E-32). These findings underscore the critical role of repeated measurements, advocating for personalized interventions based on unique clinical trajectories. This approach shows promise in tailoring treatment strategies and early identification.