

# Development of a simple image evaluation algorithm for swallowing disorders in the elderly using Vision Transformer

21mp211 Kaoru Masuoka

## Abstract

**Background:** Sarcopenic dysphagia is a clinical condition associated with loss of swallowing-related muscle mass and its strength in elderly. In the midst of ageing society, the objective of the study was to develop a simple, rapid and safe screening test using a deep learning-based image recognition to identify patients with sarcopenic pharyngeal dysphagia. With rising concern over novel infectious diseases and seasonal influenza, the goal of this study was to create an effective screening test that would be low-risk for droplet transmission as well as to propose automated image recognition using a deep learning-based analysis requiring fewer steps of assessment.

**Methods:** The study compared the results of a Vision Transformer (ViT) model to the FAST method in a previous study by Sakai *et al.* (2021) using elderly patients' photographs of the neck. The prediction performance of the ViT model was statistically evaluated using a confusion matrix on 62 test samples based on the fine-tuned model obtained on Google Colaboratory.

**Results:** A total of 308 patients admitted to a post-acute care hospital in Japan were included in the study with 89 (28.90%) patients diagnosed with sarcopenic pharyngeal dysphagia and 219 (71.10%) without sarcopenic pharyngeal dysphagia. The area under the receiver operating characteristic curve (ROC-AUC), sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) using the ViT model were 0.827, 89.13% 81.25% 93.18% and 72.22%, respectively. These results were compared to the results from a previous image-only model by Sakai *et al.* (2021) which found ROC-AUC, sensitivity, specificity, PPV and NPV values of 0.814, 71.88% 80.00%, 65.71%. and 84.21%, respectively. The comparative study found that the ViT model had higher performance measures than the previous study in all measures except for NPV.

**Conclusion:** The deep learning-based screening test for sarcopenic pharyngeal dysphagia using image recognition of neck appearance had a good prediction performance.

**Keywords:** dysphagia; sarcopenia; elderly; aging; swallowing; pharyngeal; screening; swallow test; image recognition; image classification; deep learning; ViT