

Predicting patients with high insurance expenditures for home care services using machine learning

Background: The number of persons using home care services increased in Switzerland in the past years and is expected to increase further. According to Swiss home care service organisations especially the number of cases with high health care use increased and they assert that the current Swiss tariff system poorly reflects high users. Moreover, it remains poorly understood which patient characteristics are linked to higher service use or high insurance expenditures and home care service organisations have a poor basis to predict which patient might be part of this group. The objective of this study was to identify patient characteristics that predict insurance expenditures above 1000 CHF per month and to determine which patient characteristics are most important to predict expenditures.

Methods: This observational study was based on data from eight Swiss home care service organisations. All patients undergoing routine assessments with the interRAI Home-Care Switzerland (HCS) questionnaire during the inclusion phase were included in the study. Besides data from the interRAI-HCS further characteristics covering mostly the case environment were assessed. Administrative data on service use in the three months following the assessment was the basis to calculate insurance expenditures. 168 patient features served as input variables and insurance expenditures above/below 1000 CHF per month as endpoint in our predictive classification models. The aim was to find the best prediction model by evaluating six machine learning methods (including random forests, neural networks and XGBoost). Explainable-AI approach SHAP was used to explain the contribution of each feature to the model prediction and to assess global feature importance.

Results: 1035 patients were included in our study (mean age: 80 years, SD: 11.6). Best predictions were found with the XGBoost model that resulted in a ROC AUC value of 0.857 in the holdout set. The three patient characteristics most influential for expenditure prediction were: strength of limitations for dressing below the waist, use of meal services and needs for help with medication management.

Discussion: Machine learning is useful to predict insurance expenditures in the home care services setting. However, the generalisability of our models needs to be further evaluated in broader contexts. Our study can provide guidance on which patient characteristics should be routinely collected by organisations to predict insurance expenditures.