Can Patterns of Habitual Activity Stratify Primary Sjogren's Syndrome Participants with Persistent Fatigue?

Chloe Hinchliffe[1], Bing Zhai[2], Victoria Macrae[3], Jade Walton[3], Silvia Del Din[1,3], and Wan-Fai Ng[3,4]
[1]Translational and Clinical Research Institute, Newcastle University, Newcastle Upon Tyne, UK
[2]Computer and Information Science, Northumbria University, Newcastle upon Tyne, UK
[3]NIHR Newcastle Biomedical Research Centre, Newcastle upon Tyne Hospitals NHS Trust, Newcastle Upon Tyne, UK
[4]HRB Clinical Research Facility, University College Cork, Cork, Ireland
Email: chloe.hinchliffe@newcastle.ac.uk

Summary

Wearable devices may be able to provide objective, continuous assessments of fatigue in chronic disease. 97 participants with primary Sjogren's syndrome wore sensors on the chest and wrist whilst self-reporting their fatigue. Measures of activity derived from these devices were used to stratify participants with persistent fatigue using statistical and machine learning methods. Those with persistent fatigue were shown to be less active but this difference was insufficient for reliable machine learning classification.

Introduction

Many individuals with chronic diseases, such as primary Sjogren's syndrome (PSS), experience debilitating fatigue that substantially impacts their quality of life. Currently, assessments of fatigue rely on patient-reported outcomes, which are subjective and prone to recall bias. Wearable devices, however, can provide continuous assessments of activity and may provide objective evidence of fatigue, which can aid in therapeutic development. Currently, the literature on this subject is very limited. Therefore, this study aims to separate PSS participants with different levels of fatigue persistency using real-world measures of physical activity.

Methods

For the Newcastle Biomedical Research Centre Tools Study [1], 97 participants with PSS wore two devices containing accelerometers at home for up to two 7-day continuous periods: VitalPatch on the chest and Actigraph GT9X Link on the non-dominant wrist. Meanwhile, participants reported their fatigue up to 4 times a day. The total durations engaging in different activities (standing, walking, or sitting) were estimated by the VitalPatch and total durations for different activity intensities (inactive, light, moderate, and vigorous) were estimated from the Actigraph using GGIR (version 2.1) [2,3]. GGIR also extracted activity duration within several bout lengths: 1-10 min, 10-30 min, and >30 min for inactivity, 1-10 min and >10 min for light, and >1 min for moderate-to-vigorous. Participants were categorised as "persistently fatigued", otherwise "non-persistently fatigued", if their minimum reported fatigue score was above a threshold: 1, 2, 3, and 4. Each digital measure was tested for each threshold. Welch's t-test was used if both classes were normal (tested with Shapiro); otherwise, Mann-Whitney U was used.

The statistically significant features ($p<0.05$) for each threshold were classified with four machine learning models: support vector machine (SVM), k-nearest neighbours (kNN), random forest (RF), and Gaussian naive Bayesian network (NB). The data were split into training and testing sets using 10-fold leave-subjects-out cross-validation. Missing values were imputed with the median of the training set, the classes were equalised with oversampling, and then the inputs were scaled. The hyperparameters of the classifiers were optimised using a 5-fold grid search within each fold.

Results and Discussion

None of the features were statistically significant ($p\geq0.05$) for thresholds 1 and 2. For threshold 3, those with persistent fatigue spent more time inactive (total ($p=0.034$) and >30 min bouts ($p=0.033$)), and non-persistent fatigue spent more time in light activity (total ($p=0.008$) and 1-10 min bouts ($p=0.004$)). For threshold 4, non-persistent fatigue spent more time in light activity (total ($p=0.023$) and 1-10 min ($p=0.024$)). This indicates that those with persistent fatigue had reduced duration in shorter bouts of light activity, instead they spent more time in longer bouts of inactivity. Thus, highlighting a decrease in habitual activity with increased fatigue.

The machine learning classifiers, however, show that the statistically significant features of each threshold were unable to predict the classes substantially better than random chance (50%). Despite using only two features, threshold 4 outperformed threshold 3, with the highest mean balanced accuracy of 63.2%.

Table: Balanced accuracies of the machine learning classifiers, averaged across all folds.

Threshold	SVM	kNN	RF	NB	Mean
3	46.2%	50.1%	49.9%	57.2%	50.9%
4	62.2%	57.8%	63.2%	60.4%	60.9%

Conclusions

The activities derived from the VitalPatch were insufficient to separate persistently from non-persistently fatigued PSS participants, but activity intensity analysis from the Actigraph indicated to a decline in light activity and increased inactivity in persistent fatigue. However, machine learning analysis determined that these features alone are inadequate to reliably identify persistent fatigue. Therefore, future research should explore additional measures to support real-world activity and lead to an objective, continuous assessment of fatigue with digital wearables.

References

[1] https://doi.org/10.1186/ISRCTN12062123.
[2] Migueles JH et al. (2019). *JMPB*, **2**: 188-196.
[3] van Hees VT et al. (2013) *PLoS ONE*, **8**: e61591.