ECRONICON OPEN ACCESS

EC NEUROLOGY
Research Article

Identifying Key Symptoms Differentiating Myalgic Encephalomyelitis and Chronic Fatigue Syndrome from Multiple Sclerosis

Diana Ohanian, Abigail Brown, Madison Sunnquist, Jacob Furst, Laura Nicholson, Lauren Klebek and Leonard A Jason*

DePaul University, Chicago, USA

***Corresponding Author**: Leonard A Jason, DePaul University, Chicago, USA.

Received: December 08, 2016; **Published:** December 19, 2016

Abstract

It is unclear what key symptoms differentiate Myalgic Encephalomyelitis (ME) and Chronic Fatigue syndrome (CFS) from Multiple Sclerosis (MS). The current study compared self-report symptom data of patients with ME or CFS with those with MS. The self-report data is from the DePaul Symptom Questionnaire, and participants were recruited to take the questionnaire online. Data were analyzed using a machine learning technique called decision trees. Five symptoms best differentiated the groups. The best discriminating symptoms were from the immune domain (i.e., flu-like symptoms and tender lymph nodes), and the trees correctly categorized MS from ME or CFS 81.2% of the time, with those with ME or CFS having more severe symptoms. Our findings support the use of machine learning to further explore the unique nature of these different chronic diseases.

Keywords: *Chronic Fatigue Syndrome; Myalgic Encephalomyelitis; Multiple Sclerosis; Data Mining; Decision Trees*

Introduction

Furst, Raicu, and Jason [1] suggest that data mining or machine learning are particularly appropriate for research on social problems and diagnostic challenges, as they can be used to uncover patterns and relationships within large samples of people that would not otherwise be evident because of the size and complexity of the data. Machine learning has been used in the past to examine disease course, diagnostic criteria and therapy response of a number of fatiguing illnesses [2,3]. Data mining has been used to help in the diagnosis of diabetes, stroke, cancer, and heart disease, and it has also been used for research purposes to predict conditions/diseases and to describe symptoms by using algorithms that differentiate different groups of illness [4].

In the field of Myalgic Encephalomyelitis (ME) and Chronic Fatigue Syndrome (CFS), data mining has been used as a way to identify unique aspects of this illness [5]. For example, Huang, Hsu, and Lin [6] explored Single Nucleotide Polymorphisms associated with CFS to predict the CFS diagnosis. Machine learning in the form of decision trees has been used to compare people with CFS with healthy controls [7]. Jason., *et al*. [8] used decision trees to identify which items best classify someone as a patient or control, and the most important predictive symptoms were fatigue, post-exertional malaise, sleep dysfunction and neurocognitive issues. Other symptoms such as pain, autonomic, immune and neuroendocrine symptoms were not as important.

Several studies have attempted to differentiate CFS or ME from other chronic illnesses. For example, Jason., *et al*. [9] examined the symptomatology differences among those with MS, CFS, and Lupus. This study concluded that those with MS were the most similar to ME or CFS in terms of impairment due to fatigue and reduction in activity. But this study had somewhat small sample sizes and did use data mining strategies to differentiate illnesses.

We know little about the symptoms differentiating MS from ME or CFS. In addition, for those in primary care, it would be useful to be able to identify a set of symptoms that might help differentiate these two conditions. The current study used machine learning to establish what symptom features differentiate MS and CFS from ME.

Method

Participants

Participants were recruited through several different outlets to take the online DePaul Symptom Questionnaire (DSQ) [10]. Links and descriptions of the survey were posted to support group websites, national foundations, research forums, and social media outlets. Social media outlets included Facebook groups and pages and twitter pages. 106 people with MS and 354 people with ME or CFS fully completed the questionnaire.

The DePaul Symptom Questionnaire (DSQ) [10] is a self-report measure of symptomatology, demographic, medical, occupational and social history. In the survey, participants are asked to rate their symptom frequency and severity on a scale from 0 - 4. For frequency, the scores are as follows, 0=none of the time, 1=a little of the time, 2=about half the time, 3=most of the time, 4= all of the time. For severity, the scores are as follows, 0=symptom not present, 1=mild, 2=moderate, 3=severe, 4=very severe. Frequency and severity scores were multiplied by 25 to create 100-point scales. The 100-point frequency and severity scores for each symptom were averaged to create one composite score per symptom. The DSQ is available at REDCap's [11] shared library: https://redcap.is.depaul.edu/surveys/?s=tRxytSPVVw

The DSQ has evidenced good test-retest reliability among both patient and control groups [8]. A factor analysis by Brown and Jason [12] found a three-factor solution, with factors evidencing good internal consistency. Murdock., *et al.* [13], an independent group using the DSQ, found that it demonstrated excellent internal reliability, and that among patient-reported symptom measures, it optimally differentiated between patients and controls.

Data Mining

Decision trees were used in this study to analyze the symptom data. Data mining techniques in general provide statistical analyses that identify which items on the questionnaire best predict class membership and are useful for diagnosis and assessment. In this study, decision trees were used to determine what symptoms differentiated those with MS from those with ME or CFS. For this analysis, symptom scores were calculated by combining the scores of frequency and severity into a score on a 100-point scale. The resulting variables were then put into decision tree analysis.

Decision trees consist of a series of binary choices (or branches) that end with a classification of participants. At each branch the computer decides what symptom would best predict classifications, in this case whether someone has MS or ME or CFS. The process continues and more symptoms continue to separate the two groups until the tree reaches a balance between classification accuracy and generalizing to new data.

SPSS statistics software was used to create the decision trees. To build the models we used the Classification and Regression Tree (CRT) algorithm. Due to the unbalanced samples, we created a random sub sample of the ME or CFS group each time we ran a tree that would be roughly equal to the amount of MS responses we had. We ran 100 trees with all of the symptoms of the DSQ and then ran specific trees combinations that were informed by those initial results 100 times each. While the samples are disproportionate, the number of responses from both groups are large enough to produce reliable results.

We examined the number of times symptoms from the DSQ on the 100-point scale appeared in the decision trees. If a symptom appeared in the first level, it means it was the most important in that tree for separating the people with MS from those with ME or CFS. Those that appeared in level 2 were the second most important in the tree. The level of importance in separating the samples into the two

distinct groups lessens with each level. Due to the fact that the trees were ran 100 times, almost every symptom in the DSQ appeared in at least the 4th level once. It is noteworthy if a symptom appears multiple times. Accuracy assess how well the trees classified MS and ME or CFS responses. Sensitivity in this case assess how well the trees identified those who have ME or CFS correctly. Specificity in this case assess how well the trees classified those with have MS correctly.

Results

Table 1 displays the top 5 symptoms that differentiated MS from ME or CFS. The first two symptoms listed appeared the most in the first level, indicating that flu-like symptoms and tender lymph nodes are the best at differentiating these illnesses. Furthermore, it should be noted that the two best symptoms that differentiated the greatest number of people with MS from ME or CFS in the decision trees were from the immune domain. In Table 2 the means and standard deviations are displayed for the top five symptoms that appeared in the decision trees, and it should be noted that the means for these symptoms are significantly higher for those with ME or CFS than for those with MS.

Symptom	Level 1	Level 2	Level 3	Level 4	Level 5	Total
Flu-like symptoms	60	13	6	2	0	81
Tender lymph nodes	27	20	3	0	0	50
Alcohol intolerance	4	29	5	1	0	39
Inability to tolerate upright position	6	18	5	1	1	31
Next-day soreness after non-strenuous activities	3	14	11	0	0	28

Table 1: *Number of Times Symptoms Appeared in Decision Trees with 100-Point Scale (MS vs. ME or CFS).*

Means and Standard Deviations	(*n* = 106)		(*n* = 354)		
Symptom	M	(SD)	M	(SD)	Sig.
Flu-like symptoms	17.0	(22.0)	51.3	(27.1)	***
Tender lymph nodes	8.1	(18.3)	36.7	(29.4)	***
Alcohol intolerance	9.4	(23.9)	38.3	(37.0)	***
Inability to tolerate upright position	16.1	(27.6)	50.7	(34.6)	***
Next-day soreness after non-strenuous activities	46.1	(28.4)	75.0	(20.6)	***

Table 2: *Means and Standard Deviations.*

Finally, two symptoms were chosen to build a tree due to the fact that flu-like symptoms appeared the most as a symptom of primary importance (level 1) and tender lymph nodes appeared the most as a symptom of secondary importance (level 2). With only two symptoms, flu-like symptoms and tender lymph nodes, the trees correctly categorized MS and ME or CFS 81.2% of the time. Furthermore, ME or CFS was correctly categorized 84.0% of the time and MS was correctly categorized 79.2% of the time. In an attempt to increase those percentages, we ran two other tree formations which attempted to address symptoms that were of secondary (level 2) and tertiary importance (level 3). The accuracy, sensitivity, and specificity did not increase significantly with these alternate combinations. Therefore, it appears that flu-like symptoms and tender lymph nodes are the most important symptoms when differentiating these two illnesses.

Discussion

This is the first study to use machine learning to examine the differences between MS versus ME or CFS, which have a number of overlapping symptoms including fatigue. Among the best 5 discriminating symptoms, those with ME or CFS had significantly worse scores than those with MS. The most important two symptoms that differentiated MS versus ME or CFS existed within the immune domain.

Our results support the use of machine learning for developing diagnostic criteria and assessment tools for ME or CFS. Furthermore, the findings regarding immune differences between people with MS and ME or CFS encourage further research into immune dysfunction in ME or CFS. The immune symptoms are generally not seen as the most important symptoms of ME or CFS, as they were not specified in the recent IOM (2015) clinical case definition of Systemic Exertion Intolerance Disease. However, we found that with just two immune symptoms decision trees could accurately predict more than 80% of the time whether a person had MS or ME or CFS. This should encourage future research into immune cell activity as explored by Brenu., *et al*. [14] or plasma abnormalities as explained by Hornig., *et al* [15].

Our findings could also inform assessment criteria. Because ME or CFS and MS have similar presentations recognizing the importance of immune dysfunction for ME or CFS might be important for healthcare providers. Therefore, while not necessarily applicable in case definitions for the illness, researchers should be aware of the importance of these symptoms when performing assessment and examining subtypes. Our results also suggest that people with ME or CFS have more severe symptom expression on the five best discriminating symptoms than people with MS [16-23].

Conflict of Interest

The authors of this article have no conflict of interest to report.

Acknowledgements

The authors would like to thank all of the support groups and organizations who made this project possible.

Funds

Funding was provided by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (Grant No. HD072 208).

Bibliography

1. Furst J., *et al*. "Data mining". In L. A. Jason & D. S. Glenwick (Eds). Handbook of methodological approaches to community-based research: Qualitative, quantitative, and mixed methods. New York, NY: Oxford University Press (2016): 187-196.

2. Fiorini S., *et al*. "A machine learning pipeline for multiple sclerosis course detection from clinical scales and patient reported outcomes". *Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (2015): 4443-4446.

3. Kourou K., *et al*. "Machine learning applications in cancer prognosis and prediction". *Computational and Structural Biotechnology Journal* 13 (2014): 8-17.

4. Shouman M., *et al*. "Using Decision Tree for Diagnosing Heart Disease Patients". *Australasian Data Mining Conference, Eds., ACS* (2011): 23-30.

5. Hanson S J., *et al*. "Detection of immunologically significant factors for chronic fatigue syndrome using neural-network classifiers". *Clinical and Diagnostic Laboratory Immunology* 8.3 (2001): 658-662.

6. Huang LC., *et al.* "A comparison of classification methods for predicting Chronic Fatigue Syndrome based on genetic data". *Journal of Translational Medicine* 7 (2009): 81.

7. Watson S., *et al*. "Identifying defining aspects of chronic fatigue syndrome via unsupervised machine learning and feature selection". *International Journal of Machine Learning and Computing* 4.2 (2014): 133-138.

8. Jason L A., *et al*. "Examining case definition criteria for chronic fatigue syndrome and myalgic encephalomyelitis". *Fatigue: Biomedicine, Health and Behavior* 2.1 (2014): 40-56.

9. Jason L A., *et al.* "A Screening Instrument for Chronic Fatigue Syndrome". *Journal of Chronic Fatigue Syndrome* 3.1 (1997): 39-59.

10. Jason L A., *et al*. "The development of a revised Canadian myalgic encephalomyelitis-chronic fatigue syndrome case definition". *American Journal of Biochemistry and Biotechnology* 6.2 (2010): 120-135.

11. Harris PA., *et al*. "Research electronic data capture (REDCap) - A metadata-driven methodology and workflow process for providing translational research informatics support". *Journal of Biomedical Informatics* 42.2 (2009): 377-381.

12. Brown AA and Jason LA. "Validating a measure of myalgic encephalomyelitis/chronic fatigue syndrome symptomatology". Fatigue: *Biomedicine Health and Behavior* 2.3 (2014): 132-152.

13. Murdock KW., *et al*. "The utility of patient-reported outcome measures among patients with myalgic encephalomyelitis/chronic fatigue syndrome". *Quality of Life Research* (2016).

14. Brenu E W., *et al*. "Immunological abnormalities as potential biomarkers in Chronic Fatigue Syndrome/Myalgic Encephalomyelitis". *Journal of Translational Medicine* 9 (2011): 81.

15. Hornig M., *et al*. "Distinct plasma immune signatures in ME/CFS are present early in the course of illness". *Science Advances* 1.1 (2015): e1400121-e1400121.

16. Anderson JS and Ferrans CE. "The quality of life of persons with chronic fatigue syndrome". *Journal of Nervous and Mental Disease* 185.6 (1997): 359-367.

17. Green J., *et al*. "Stigma and chronic fatigue syndrome". *Journal of Chronic Fatigue Syndrome* 5.2 (1999): 63-95.

18. Richman JA., *et al*. "Feminist perspectives on the social construction of illness states". *Health Care for Women International* 21.3 (2000): 173-185.

19. Richman JA and Jason LA. "Gender biases underlying the social construction of illness states: The case of chronic fatigue syndrome". *Current Sociology* 49.3 (2001): 15-29.

20. Sorenson M., *et al*. "Disruption of cytokine pathways in Chronic Fatigue Syndrome" (2016).

21. Sorenson M., *et al*. "The production of Interleukin-8 is increased in plasma and peripheral blood mononuclear cells of patients with fatigue". *Neuroscience and Medicine* 3 (2012): 47-53.

22. Tidmore T., *et al*. "Lack of knowledgeable healthcare access for patients with neuro-endocrine-immune diseases". *Frontiers in Clinical Medicine* 2 (2015): 46-54.

23. Twemlow SW., *et al*. "Patterns of utilization of medical care and perceptions of the relationship between doctor and patient with chronic illness including chronic fatigue syndrome". *Psychological Reports* 80.2 (1997): 643-659.

Volume 4 Issue 2 December 2016