

# Identifying Defining Aspects of Chronic Fatigue Syndrome via Unsupervised Machine Learning and Feature Selection

Samuel P. Watson, Amy S. Ruskin, Valerie Simonis, Leonard A. Jason, Madison Sunnquist, and Jacob D. Furst

**Abstract**—In this work we propose an unsupervised machine learning method of predicting chronic fatigue syndrome (CFS) based on the k-means algorithm using self-reported questionnaire responses. We first suggest a method of determining the presence of a symptom based on its frequency and severity using an unsupervised dynamic thresholding approach. This threshold is used to diagnose subjects with 54 symptoms related to CFS. Based on these diagnoses, k-means is used to predict the presence of CFS. We find that k-means does not have significantly worse predictive diagnostic accuracy than commonly used CFS case definitions. After applying supervised feature selection, k-means achieves significantly better diagnostic accuracy than any of the case definitions examined. We use these results to suggest the basis for an empirically founded CFS case definition.

**Index Terms**—Chronic fatigue syndrome, computer-aided diagnosis, k-means clustering, machine learning.

## I. BACKGROUND AND MOTIVATION

Chronic fatigue syndrome (CFS) is a disease characterized by six months of debilitating fatigue, accompanied by post-exertional malaise, pain, and various neurological and autonomic symptoms such as short-term memory loss and imbalance [1]. The causes of CFS are not yet well understood. There is no reliable, widely accepted biomarker for CFS, and because its symptoms are often shared by other illnesses, classifying CFS in a precise way has presented a significant challenge. As a result, there has been interest from the CFS research community in utilizing data mining techniques to better identify key features of the illness. Machine learning and statistical selection techniques can help reveal which symptoms are the most useful in distinguishing between CFS patients and healthy people, and hence which symptoms are most characteristic of the illness. These symptoms may then form the basis for an empirical CFS definition.

Many definitions of CFS have been proposed, but few have gained traction and none have escaped criticism. The Fukuda criteria, developed in 1994, which is currently the

most widely used diagnostic measure. The Fukuda criteria stipulates that chronic fatigue of a definite onset (i.e. not lifelong) must be present, in addition to four of the eight following symptoms: memory/concentration impairment, sore throat, tender lymph nodes, muscle pain, joint pain, headaches, unrefreshing sleep, and post-exertional malaise [2]. However, the Fukuda criteria has been criticized for being overly inclusive, as cardinal CFS symptoms such as post-exertional malaise and memory or concentration impairment do not need to be present for a positive diagnosis [1]. Additionally, the Fukuda criteria places particular emphasis on the construct of fatigue, which is a common complaint among the general population and thus less useful in distinguishing between CFS sufferers and healthy people [1].

Since the introduction of the Fukuda criteria, there have been numerous efforts to formulate a precise, standardized definition for CFS. In 2003, the Canadian case definition was proposed, which requires the presence of fatigue, post-exertional malaise, sleep dysfunction, and pain, in addition to two forms of various neurological or cognitive manifestations, and one or more symptoms from at least two of the following symptom categories: autonomic, neuroendocrine, and immune [3]. The Canadian criteria has been found to select for cases with greater physical functioning and neurological impairment, as well as less psychiatric comorbidity than the Fukuda definition [4]. However, the Canadian criteria has come under scrutiny for not requiring sudden symptom onset, a feature most other definitions agree on [1].

The Myalgic Encephalomyelitis International Consensus Criteria (ME-ICC) was developed in 2011, requiring post-exertional neuroimmune exhaustion and at least 50% reduction in activity level following onset. The ME-ICC also stipulates at least one type of energy production/transportation impairment (e.g. cardiovascular or respiratory dysfunction), as well as at least one symptom from three of the four following categories: neurocognitive impairments, pain, sleep disturbance, and neurosensory, perceptual, and motor disturbances [5]. The ME-ICC has been shown to select for patients with greater functional, physical, and mental impairment than the Fukuda definition, but those meeting the ME-ICC also have higher rates of psychiatric comorbidity [4].

Although there have been numerous other proposed CFS definitions, this paper will focus on the aforementioned three, as they have received the most widespread recognition. However, a commonality between these and nearly all other case definitions is that they were formed via a clinician

Manuscript received August 19, 2013; revised December 10, 2013. This work was supported under NSF Grant #1062909.

S. P. Watson is with the Carleton College, Northfield, MN 55057 USA (e-mail: watsons@carleton.edu).

A. S. Ruskin is with Pomona College, Claremont, CA 91711 USA (e-mail: asr02010@pomona.edu).

J. D. Furst and V. Simonis are with the College of Computing and Digital Media, DePaul University, Chicago, IL 60604 USA.

L.A. Jason and M. Sunnquist are with the College of Science and Health, DePaul University, Chicago, IL 60614 USA.

consensus approach, rather than an empirical one. An empirical definition could help to more accurately diagnose patients by reflecting authentic symptom patterns of CFS sufferers.

Reeves *et al.* developed an empirical definition based on patient self-report data from the Multidimensional Fatigue Inventory (MFI), the Medical Outcomes Short Form-36 (SF-36), and Symptom Inventory (SI), as measures of fatigue, disability, and frequency and severity of accompanying symptoms, respectively [6]. Reeves *et al.* compared patients who either currently met the Fukuda criteria or had met it at one time against a control group. 2-step log-likelihood clustering was used to set thresholds for a CFS definition based on scores from the respective tests. However, since the study exclusively examined patients diagnosed under the Fukuda criteria their results may simply reflect selection biases of the Fukuda definition, rather than providing a truly unique basis for a new classification standard. Moreover, the MFI, SF-36, and SI were created as general-purpose diagnostic tests and are not specifically directed toward CFS. This may result in a definition that is overly inclusive, as the tests do not target specific CFS symptoms. The Reeves definition has been criticized for its broadness, as CFS rates would increase by over ten times previous estimates under the new criteria [1]. Jason *et al.* found that 38% of patients with a Major Depressive Disorder would be misdiagnosed as having CFS under the Reeves definition [7].

## II. RELATED WORK

Numerous past studies have employed supervised machine learning techniques to understand the disease. Hanson, Gause, and Natelson used multilayer neural networks to predict CFS based on 29 immunological genetic transcription factors [8]. The neural network identified 14 relevant factors in distinguishing CFS patients from controls, and achieved an out of sample classification accuracy of 65%. Their results support the hypothesis that CFS stems from immunological dysfunction, and indicate the underlying genetic relationships associated with CFS are complex and nonlinear.

Similarly, Huang, Hsu, and Lin sought to identify Single Nucleotide Polymorphisms (SNPs) associated with CFS [9]. They explored a wide range of classification algorithms and feature selection techniques, and were able to identify 8 SNPs correlated with CFS using a Naïve Bayes classifier with wrapper-based feature selection, with an AUC of 70%.

In an effort to study the disease from a semantic perspective, Bronikowski *et al.* explored ensemble decision tree classifiers as a means of predicting CFS based on patient self-report data. Ensemble methods did not achieve a higher accuracy than standard two-way decision tree classifiers, which were able to distinguish between patients and “others” (controls, exclusions, and unexplained fatigue) with 68.75% accuracy [10].

There has been little application of unsupervised machine in regards to CFS. Supervised machine learning techniques, as used in [8]-[10], are only valid insofar as they reflect the initial diagnosis criteria. In order to develop an empirical definition of CFS, it is imperative to minimize any reliance on pre-existing case definitions so that results do not simply mirror selection biases of the prior definitions.

In this paper, we draw on diverse patient data to show that unsupervised machine learning techniques can accurately predict CFS, even for patients diagnosed under different criteria. We then use these techniques as a basis for identifying the most predictive CFS symptoms, and draw conclusions as to which features an empirical CFS definition should focus on.

## III. METHODS

### A. Datasets

In order to increase generalizability we analyzed data from three separate datasets: BioBank, DePaul I, and Newcastle. Since these datasets used different criteria for recruiting subjects and diagnosing CFS, a cumulative analysis of the datasets can provide a holistic interpretation of the disease.

Each dataset contains subject self-responses to the DePaul Symptom Questionnaire (DSQ). The DSQ is designed to gauge the presence of 54 symptoms most commonly associated with CFS [1]. Each symptom is reported in terms of frequency and severity on a 0-4 Likert scale, so there are effectively 108 DSQ features. A frequency of 0 means a symptom occurs “none of the time,” whereas a frequency of 4 means a symptom occurs “all of the time.” Likewise, a severity of 0 means the symptom is “not present,” and a severity of 4 means a symptom occurs “all of the time.” Relevant background information for the individual datasets is given below, but all analysis was done using the combined data from all three. Table I summarizes relevant database statistics.

#### 1) BioBank

TABLE I: NUMBER OF CFS PATIENTS AND CONTROLS IN BIOBANK, DEPAUL, NEWCASTLE, AND COMBINED DATASETS

Sample set	Number of CFS patients	Number of controls	Total number of subjects
BioBank	233	80	313
DePaul	187	96	283
Newcastle	95	0	95
Combined datasets	515	176	691

CFS patients were recruited for the BioBank sample via the CFIDS Association of America website, internet forums and other social networks, as well as physician referrals. All patients received a diagnosis under the Fukuda or Canada criteria from a licensed physician specializing in the illness. After eliminating subjects missing 10% or more responses to the DSQ, the dataset contains a total 137 missing values (about 0.4% of the data). Methods for replacing missing data are described in detail below. The final dataset consists of 233 CFS patients and 80 controls.

#### 2) DePaul

The DePaul sample recruited CFS sufferers through internet forums, support groups, and from previous DePaul University studies. DePaul undergraduates were recruited as controls. After removing cases missing 10% or more responses, 214 missing values remained (about 0.6% of the data). 27 subjects with exclusionary conditions such as morbid obesity or depression were also removed. The final dataset contains 187 patients and 96 controls.

### 3) Newcastle

CFS patients were referred to the Newcastle study through their primary care physician, who gave them a suspected diagnosis of CFS. It is worth noting that the Newcastle data was collected from subjects in the United Kingdom, whereas BioBank and DePaul were collected in the United States. The dataset contains no controls. Subjects missing 10% or more responses were removed, leaving 76 missing values (about 0.7% of the total data). The cleaned dataset contains 95 patients.

#### B. Methods for Replacing Missing Values

For the cases that had a score of 0 for either frequency or severity of a symptom and were missing the other field, the missing value was set to 0; the rationale was that a symptom should occur “none of the time” (frequency=0) if and only if the symptom is “not present” (severity=0). Otherwise, if a subject was missing data in only one of the two fields (frequency or severity) for a symptom, then the missing value was replaced with the mode value from the cases that had the same score for the non-missing field. When both fields were missing for a symptom, the values were replaced with the overall medians in those fields for that symptom.

#### C. Unsupervised Thresholding

In order to avoid treating frequency and severity scores for a symptom as independent features when applying machine learning algorithms, it was necessary to consolidate them into a single value. A thresholding approach was used to determine the presence of a symptom based on its frequency and severity, thus reducing the two dimensional scores for each symptom to one binary score (symptom present/not present). Rather than choosing an arbitrary threshold, we wanted to dynamically adjust the threshold for each symptom based on observed frequency and severity scores. This was achieved through a k-means clustering approach.

Generally speaking, the k-means algorithm works by iteratively dividing coordinate points into a predetermined number of clusters based on which cluster center the point lies closest to. The cluster center is defined by the average value of all coordinate pairs in the cluster. After all points have been assigned to a cluster, the new cluster centers are calculated. Each point is then reassigned based on the new cluster centers. This process of center calculation and reassignment repeats until equilibrium is reached or a set number of iterations is completed.

In this case, the k-means clustering algorithm was set to find two clusters, based on the underlying assumption that the data consists of delineable patient and control groups. Frequency and severity scores for each symptom were treated as coordinate pairs for the purpose of cluster assignment, and the Euclidean distance was used to measure closeness to the cluster centers. After equilibrium was reached, the perpendicular bisector of the line between cluster centers was found. This bisecting line was used as the threshold; frequency-severity pairs above the threshold line were considered “symptom present,” whereas scores below the line were considered “symptom not present.”

#### D. Assessing Threshold Validity

In order to determine the validity of unsupervised

thresholding, the diagnostic accuracy of the unsupervised threshold was compared against two static thresholds and a threshold developed from a supervised clustering technique. The static one-one threshold considered a symptom present if the frequency and severity scores for the symptom are both greater than or equal to one. The static two-two threshold was similar, but used a score of two as the minimum for symptom presence. The supervised threshold was developed using a similar cluster-based perpendicular bisector technique as the unsupervised threshold, but subjects were divided into clusters based on the label of patient or control, rather than frequency and severity of symptoms.

Using these thresholds for determining symptom presence, subjects were diagnosed under the Fukuda, Canadian, and ME-ICC case definitions. The sensitivity, specificity, and accuracy of the thresholds under each case definition were calculated based on how well they predicted the diagnostic labels (e.g. the official label of patient or control as given in the dataset). It is important to note that while the DSQ contains nearly all of the symptoms listed in the ME-ICC and Canadian criteria, some symptoms (disorientation, for example) are absent. Hence, minor modifications to the ME-ICC and Canadian definitions had to be made for the purpose of analysis, either through omission or by substituting closely related symptoms. It was possible to apply the Fukuda criteria without modification.

#### E. Guided Feature Selection

A naive feature selection method was used to identify the most predictive CFS symptoms. K-means clustering ( $k=2$ ) was performed on the frequency/severity scores for each of the 54 DSQ symptoms individually. Hence, for each symptom we identified a cluster of high frequency and severity scores (deemed ‘symptom present’) and a cluster of low frequency and severity scores (‘symptom not present’). Sensitivity, specificity, and accuracy were calculated for each symptom by treating the diagnostic label of each case (patient or control) as the actual value, and membership in the symptom present or symptom not present clusters as the predicted value (e.g. assigning a patient to the symptom present cluster is considered a true positive, assigning a control subject to the symptom not present cluster is a true negative, etc.). Thus, the predictive accuracy of a symptom provides a measure of how well it can be used to distinguish between CFS patients and controls. Symptoms were then ranked by descending accuracy, and subsets of the top 1-15 symptoms were evaluated for their predictive diagnostic accuracy.

#### F. Assessing Feature Subset Validity via Unsupervised Symptom Clustering

K-means clustering was used as a diagnostic tool to measure the predictive accuracy of subsets of the top 1-15 symptoms against all 54 DSQ items. For each case, the presence of each symptom was determined using the unsupervised thresholding technique, resulting in a new set of binary (symptom present/not present) data. K-means clustering ( $k=2$ ) was then performed on the binarized data for each of the feature subsets, as well as the set of all 54 DSQ features. Thus, clusters of low symptom presence and high symptom presence were identified for each feature set. The

sensitivity, specificity, and accuracy of each feature set were calculated using the diagnostic label of each case as the actual value and cluster membership as the predicted value, where placement in the low symptom cluster represents a negative diagnosis (control), and placement in the high symptom cluster represents a positive diagnosis (CFS).

IV. DISCUSSION

In order to assess threshold validity, subjects' frequency and severity scores for each of the 54 DSQ symptoms were converted into binary symptom present/not present scores using the unsupervised clustering threshold, the supervised clustering threshold, the two-two threshold, and the one-one threshold. Next, subjects were given diagnoses under the Fukuda, Canadian, and ME-ICC definitions for each threshold. Sensitivity, specificity, and accuracy were determined by comparing the predicted diagnosis of CFS against the diagnostic label associated with each case. Statistical significance was calculated using paired t-tests with  $p = 0.01$ .

As Table II shows, unsupervised thresholding generally provides significantly better or comparable accuracy to the static and supervised thresholds. Unsupervised and supervised thresholding are not significantly different in terms of sensitivity, specificity, or accuracy for the Fukuda and Canada definitions. For the ME-ICC case definition, supervised thresholding achieves significantly better sensitivity and accuracy, but significantly worse specificity. Supervised and unsupervised thresholds achieve high scores overall, indicating that they are reliable measures relative to the static thresholds.

TABLE II: COMPARATIVE SENSITIVITY, SPECIFICITY, AND ACCURACY OF THRESHOLDING TECHNIQUES FOR FUKUDA, CANADA, AND ME-ICC CASE DEFINITIONS

Technique	Definition	Sensitivity	Specificity	Accuracy
U.T. <sup>a</sup>	Fukuda	83.1	85.8	83.8
	Canada	82.9	87.5	84.1
	ME-ICC	74.4	91.5	78.7
S.T. <sup>b</sup>	Fukuda	80.8	86.4	82.2
	Canada	85.8	87.5	86.3
	ME-ICC	89.9*	81.3*	87.7*
Two-two static threshold	Fukuda	80.8	85.8	82.1
	Canada	77.9*	89.8	80.9*
	ME-ICC	67.4*	91.5	73.5*
One-one static threshold	Fukuda	98.1*	42.0*	83.8
	Canada	97.3*	50.0*	85.2
	ME-ICC	93.4*	52.8*	83.1

<sup>a</sup> Unsupervised thresholding

<sup>b</sup> Supervised thresholding

\* Significantly different from corresponding value using unsupervised thresholding for  $p = 0.01$

With regard to statistical significance, unsupervised thresholding performs as well or better than the static two-two threshold across all definitions. Unsupervised thresholding achieves significantly higher sensitivity and accuracy than the two-two threshold for the Canadian and ME-ICC definitions but otherwise does not perform significantly differently than the two-two threshold. Scores for the unsupervised threshold tend to be higher overall.

The one-one static threshold provides significantly higher sensitivity than unsupervised thresholding for each case definition. However, the one-one threshold also has significantly lower specificity than the unsupervised threshold for each definition, dipping as low as 42% for the Fukuda criteria. Accuracies for the two thresholds are not significantly different for any case definition, but the particularly low specificity of the one-one threshold makes unsupervised thresholding a more reliable measure.

Hence, we conclude that unsupervised thresholding is a reliable metric for symptom diagnosis. It is superior to a static two-two threshold, substantially more specific than the one-one threshold, and generally comparable to the supervised threshold. Supervised thresholding seems to provide the best scores overall, but inflated performance is to be expected, as it was trained based on the original patient and control labels.

Using the unsupervised threshold, the presence of each of the 54 DSQ symptoms was determined for each subject. K-means clustering ( $k=2$ ) was applied to this dataset, resulting in one cluster of subjects experiencing few symptoms, and another cluster of subjects with relatively many symptoms. Hence, placement in the low-symptom cluster was treated as a negative CFS diagnosis, whereas placement in the high-symptom cluster was considered a positive diagnosis. Sensitivity, specificity, and accuracy were calculated by comparing this predicted diagnosis against the diagnostic label. As Table III shows, this approach provides significantly better predictive accuracy than the ME-ICC definition, and does not perform significantly differently than the Fukuda or Canadian criteria.

In order to improve diagnostic accuracy and suggest a more precise definition, a naive method of feature selection was performed upon the 54 DSQ symptoms. The 15 symptoms most characteristic of CFS were identified by ranking symptoms according to greatest predictive accuracy for the disease (see Table V). K-means clustering was then applied as a method of predicted diagnosis (described above) using subsets of the top 1-15 symptoms (see Fig. 1).

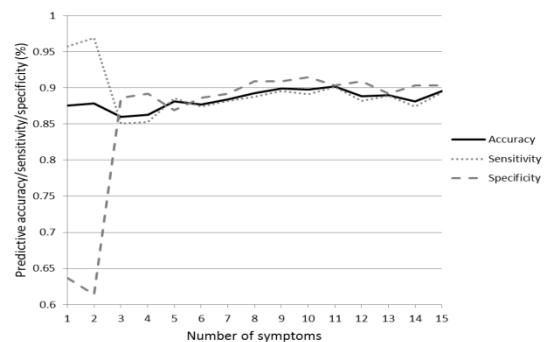


Fig. 1. Diagnostic sensitivity, specificity, and accuracy of subsets of 1-15 most predictive CFS symptoms.

TABLE III: DIAGNOSTIC SENSITIVITY, SPECIFICITY, AND ACCURACY OF SELECTIONS OF DSQ ITEMS

Number of symptoms or case definition used	Sensitivity	Specificity	Accuracy
Top 11	90.1	90.3	90.2
All 54	79.4	90.9	82.3
Fukuda	83.1	85.8	83.8
Canadian	82.9	87.5	84.1
ME-ICC	74.4	91.5	78.7

Using the top 11 symptoms achieves the highest accuracy at 90.2%, with a corresponding sensitivity of 90.1% and specificity of 90.3%. This is significantly more accurate ( $p=0.01$ ) than using all 54 symptoms, as well as any of the case definitions (see Table III). Using the top 11 symptoms does not provide significantly different sensitivity, specificity,

or accuracy than the top 5, 7-10, 13, or 15 symptoms. This suggests that an empiric CFS definition based on these selections of symptoms may provide greater diagnostic accuracy than consensus-based criteria.

Perhaps unsurprisingly, the most predictive symptom of CFS is fatigue. Symptoms 2-5 and 13 are related to post-exertional malaise, symptom 6 is a facet of sleep dysfunction, symptoms 7-10 and 12, 14, and 15 are neurocognitive issues, and symptom 11 is an aspect of general pain. This indicates that fatigue, post-exertional malaise, and neurocognitive disorders are the most predictive symptom categories of CFS. As such, a CFS case definition should place particular emphasis on these factors.

TABLE IV: SENSITIVITY, SPECIFICITY, AND ACCURACY OF TOP 15 DSQ SYMPTOMS MOST PREDICTIVE OF CFS

Ranking*	Symptom	Sensitivity	Specificity	Accuracy
1	Fatigue/extreme tiredness	95.7	63.6	87.6
2	Next day soreness or fatigue after non-strenuous, everyday activities	86.2	84.7	85.8
3	Minimum exercise makes you physically tired	85.6	85.2	85.5
4	Physically drained or sick after mild activity	82.7	90.3	84.7
5	Dead, heavy feeling after starting to exercise	82.1	88.6	83.8
6	Feeling unrefreshed after waking up in the morning	85.4	76.7	83.2
7	Problems remembering things	82.9	81.3	82.5
8	Muscle weakness	74.2	91.5	78.6
9	Difficulty finding the right word to say or expressing thoughts	77.7	79.5	78.1
10	Only able to focus on one thing at a time	76.7	82.4	78.1
11	Pain or aching in your muscles	75.1	84.1	77.4
12	Difficulty paying attention for a long period of time	78.8	72.7	77.3
13	Mentally tired after the slightest effort	72.8	89.2	77.0
14	Absent-mindedness or forgetfulness	72.0	87.5	76.0
15	Sensitivity to noise	69.7	91.5	75.3

\*by predictive accuracy of CFS

## V. CONCLUSION

Our results indicate that the DSQ, when combined with unsupervised thresholding and feature selection techniques, can provide an accurate basis for diagnosing CFS. These findings hold promise for the development of an empirical CFS case definition as an accurate diagnostic tool. We compared unsupervised, supervised, and naive thresholding techniques and found unsupervised thresholding to be a reliable means of determining symptom presence based on DSQ responses. We went on to show that applying k-means clustering to the DSQ data can provide an accurate diagnosis of CFS, as compared to the Fukuda, Canadian, and ME-ICC case definitions. Applying feature selection techniques to the DSQ symptoms can further increase the DSQ's diagnostic accuracy. In particular, we identify eleven symptoms that provide the highest predictive capability of CFS. Future work should focus on extending these findings to standardized criteria that can be easily implemented in a clinical setting. Further exploration of unsupervised thresholding could yield standardized thresholds for each DSQ symptom, reducing its ambiguity as a diagnostic tool. Additionally, our findings suggest subsets of symptoms that are most important in diagnosing CFS. Future studies should identify some criteria for the number or types of symptoms that must be expressed in order to develop a definition that can be applied in a clinical setting.

## REFERENCES

- [1] L. A. Jason *et al.*, "Myalgic encephalomyelitis case definitions," *Automatic Control of Physiological State and Function*, vol. 1, 2012.
- [2] K. Fukuda *et al.*, "The chronic fatigue syndrome: a comprehensive approach to its definition and study," *Annals of Internal Medicine*, vol. 121, no. 12, Dec. 1994.
- [3] B. M. Carruthers *et al.*, "Myalgic encephalomyelitis/chronic fatigue syndrome clinical working case definition, diagnostic and treatment protocols," *Journal of Chronic Fatigue Syndrome*, vol. 11, no. 1, pp. 7-115, 2003.
- [4] L. A. Jason, S. R. Torres-Harding, A. Jurgens, and J. Helgerson, "Comparing the Fukuda *et al.* criteria and the Canadian case definition for chronic fatigue syndrome," *Journal of Chronic Fatigue Syndrome*, vol. 12, no. 1, pp. 37-52, 2004.
- [5] B. M. Carruthers *et al.*, "Myalgic encephalomyelitis: international consensus criteria," *Journal of Internal Medicine*, vol. 270, no. 4, pp. 327-338, Aug. 2011.
- [6] W. C. Reeves *et al.*, "Chronic fatigue syndrome – a clinically empirical approach to its definition and study," *BioMed Central*, vol. 3, no. 19, Dec. 2005.
- [7] L. A. Jason, N. Najjar, N. Porter, and C. Reh, "Evaluating the centers for disease control's empirical chronic fatigue syndrome case definition," *Journal of Disability Policy Studies*, vol. 20, pp. 93-100, Oct. 2009.
- [8] S. J. Hanson, W. Gause, and B. Natelson, "Detection of immunologically significant factors for chronic fatigue syndrome using neural-network classifiers," *Clinical and Diagnostic Laboratory Immunology*, vol. 8, no. 3, pp. 658-662, May 2001.
- [9] L. C. Huang, S. Y. Hsu, and E. Lin, "A comparison of classification methods for predicting chronic fatigue syndrome based on genetic data," *Journal of Translational Medicine*, vol. 7, no. 81, Sept. 2009.
- [10] C. M. Bronikowski *et al.*, "Prediction of chronic fatigue syndrome using decision tree-based ensemble methods," presented at the

International Conference on Artificial Intelligence, Las Vegas, NV, July 18-21, 2011.



**Samuel P. Watson** was born in Colorado in 1992. He currently attends Carleton College and is pursuing a degree in computer science, with a concentration in cognitive science.

During the summer of 2013 he participated in the Medical Informatics REU program at DePaul University. He went on to study computational biology at the Aquincum Institute of Technology in Budapest during the fall of that year.



**Amy S. Ruskin** is a senior at Pomona College in Claremont, California. She is currently pursuing a computer science major. She participated in the Medical Informatics REU program at DePaul University in Chicago.