Published in final edited form as:

Physiol Meas. ; 39(11): 115001. doi:10.1088/1361-6579/aae1ed.

Multiscale network dynamics between heart rate and locomotor activity are altered in schizophrenia

Erik Reinertsen¹, Supreeth P. Shashikumar², Amit J. Shah^{3,4}, Shamim Nemati⁵, and Gari D. Clifford^{1,5}

¹Department of Biomedical Engineering, Georgia Institute of Technology, Atlanta, GA, United States of America

²Department of Biomedical Informatics, Emory University School of Medicine, Atlanta, GA, United States of America

³Department of Epidemiology, Rollins School of Public Health at Emory University, Atlanta, GA, United States of America

⁴Division of Cardiology, Department of Medicine, Emory University School of Medicine, Atlanta, GA, United States of America

⁵Department of Biomedical Informatics, Emory University School of Medicine, Atlanta, GA, United States of America

Abstract

Objective.—Changes in heart rate (HR) and locomotor activity reflect changes in autonomic physiology, behavior, and mood. These systems may involve interrelated neural circuits that are altered in psychiatric illness, yet their interactions are poorly understood. We hypothesized interactions between HR and locomotor activity could be used to discriminate patients with schizophrenia from controls, and would be less able to discriminate non-psychiatric patients from controls.

Approach.—HR and locomotor activity were recorded via wearable patches in 16 patients with schizophrenia and 19 healthy controls. Measures of signal complexity and interactions were calculated over multiple time scales, including sample entropy, mutual information, and transfer entropy. A support vector machine was trained on these features to discriminate patients from controls. Additionally, time series were converted into a network with nodes comprised of HR and locomotor activity states, and edges representing state transitions. Graph properties were used as features. Leave-one-out cross validation was performed. To compare against non-psychiatric illness, the same approach was repeated in 41 patients with atrial fibrillation (AFib) and 53 controls.

Main results.—Network features enabled perfect discrimination of schizophrenia patients from controls with an area under the receiver operating characteristic curve (AUC) of 1.00 for training and test data. Other bivariate measures of interaction achieved lower AUCs (train 0.98, test 0.96), and univariate measures of complexity achieved the lowest performance. Conversely, interaction features did not improve discrimination of AFib patients from controls beyond univariate approaches.

Significance.—Interactions between HR and locomotor activity enabled perfect discrimination of subjects with schizophrenia from controls, but these features were less performant in a non-psychiatric illness. This is the first quantitative evaluation of interactions between physiology and behavior in patients with psychiatric illness.

1. Introduction

Schizophrenia is a severely disabling and chronic mental illness which affects over 21 million people worldwide (Saha et al. 2005). Currently schizophrenia is diagnosed and managed by mental health professionals, whose availability is often scarce, particularly in low- and middle-income countries (Saxena et al. 2007). Additionally, in the context of stable schizophrenia treated in the outpatient setting, months to years can pass between clinical visits despite changes in patient status over shorter time scales (NICE guideline (CG178) 2014).

To assess clinical status more frequently and without direct observation, noninvasive technologies such as smartphones and wearable devices that measure locomotor activity via accelerometry, heart rate (HR), and other signals have been investigated. High resolution accelerometry has also been used to measure changes in social routine and circadian rhythms in mental illnesses such as depression, bipolar disorder, and schizophrenia (Van Someren et al. 1999; Reinertsen et al. 2017; Millar et al. 2004; Berle et al. 2010). Pulse sensors and electrocardiography (ECG) are used to assess HR and heart rate variability (HRV) measures, which indicate dysfunction in the autonomic nervous system (ANS). These tools could alert providers to a change in a patient's condition, monitor the effectiveness of interventions, and identify mediators of illness severity (Steinhubl et al. 2015).

While HR and locomotor activity have been assessed in a univariate sense, measures of interaction between these signals have not yet been explored, and may incrementally improve disease classification. Information is transferred between cardiovascular systems and behavioral centers over several time scales. In normal individuals, circadian rhythms mediate the increase in blood pressure and heart rate during the early morning prior to an increase in consciousness, which in turn leads to waking and locomotor activity in the morning. Conversely, the rising of an individual from a chair leads to a rise in blood pressure, heart rate, and sympathetic tone. These responses, interactions, and transitions between physiological and behavioral states vary over time scale, and are partially mediated by the ANS, the baroreflex, and central command, a feed-forward neural mechanism that contributes to motor and cardiovascular function during arousal and exercise (Hall 2010). These dynamics may be altered in patients with schizophrenia, who suffer from dysautonomia and other abnormalities in physiological control systems such as the renin-angiotensin-aldosterone pathway (Chang et al. 2009; Montaquila et al. 2015; Alvares et al. 2016).

Interactions between time series can be quantified via mutual information and transfer entropy. However, such measures are often calculated using an entire time series and fail to capture transitions between physiological and behavioral states that could provide clinically useful information. To better assess these dynamics, a time series can be evaluated over

multiple time scales, and can also be represented as a network. Attributes of this network may provide a more nuanced measure of system complexity. Previously, networks have been constructed using beat-to-beat intervals from ECGs (known as RR intervals) of patients with congestive heart failure, and visually contrasted with networks from healthy controls (Campanharo et al. 2011). Recently we demonstrated the utility of network representations for the early prediction of sepsis, indicating an association between systemic inflammation and a loss of information flow between HR and blood pressure (Shashikumar et al. 2017a). To our knowledge, interactions between different signals such as HR and locomotor activity have never been quantified in patients with mental illness. Measures of these interactions could have clinical utility if found to correlate with disease status or symptom severity.

We hypothesized that measures of interactions between HR and locomotor activity over multiple time scales differ in patients with schizophrenia due to dysautonomia and can be used to distinguish them from healthy controls. We also evaluated the additional predictive power of interaction measures for differentiating between patients with schizophrenia and controls. To better understand the generalizability and limitations of this approach, we also tested if such interactions could differentiate atrial fibrillation (AFib) from sinus rhythm when applied to 10-minute wristband pulse and locomotor activity recordings from quietly seated subjects. This contrasted from the schizophrenia analysis because AFib is cardiac-specific, and the recordings occurred in a highly controlled setting, which limits the ability to assess behavior. We thus hypothesized that measures of interaction between HR and locomotor activity are less useful in differentiating disease status in AFib than in schizophrenia.

2. Methods

The overall approach is illustrated in Figure 1. The data and preprocessing are described in the following four subsections, followed by the extracted features in the subsequent seven subsections, and finally the classifier in the last subsection.

2.1. Schizophrenia study: participants and data collection

16 clinically stable and medicated outpatient subjects diagnosed with schizophrenia, and 19 healthy control volunteers without a history of mental illness were recruited for the study (previously described by Osipov et al. (Osipov et al. 2015)). Subjects were unemployed. Age and gender did not significantly differ among the two groups, as assessed via a two-sided Student's *t*-test and Fisher's exact test, respectively. HR and locomotor activity were monitored for 3–4 weeks using a disposable adhesive patch sensor worn on the chest and manufactured by Proteus Biomedical (Redwood City, CA). ECG-derived HR data were collected every 10 min by calculating mean HR over 15 sec intervals. Locomotor activity data were collected every 5 min by calculating mean acceleration over 15 sec intervals. Data were transmitted to a mobile phone via Bluetooth and uploaded to a server for processing.

2.2. Schizophrenia study: data pre-processing

Data points in the time series of HR (measured in beats per minute, or BPM) and locomotor activity (measured in normalized units $\in [0, 1]$) with an interval exceeding $1.5\times$ the average

sampling period were discarded. The sampling period was 10 minutes for HR data and 5 minutes for activity data. Additionally, HR values lower than 20 BPM or higher than 160 BPM were labeled as low-quality and removed. HR and activity data were re-sampled to 10 min intervals via linear interpolation to ensure all data was evenly sampled and at the same sampling frequency. The root mean square (RMS) energy of acceleration for the i^{th} interval was calculated (equation 1).

Matlab R2017a (Mathworks, Natick, MA) was used for data pre-processing, feature extraction, machine learning classification, and data visualization.

2.3. AFib study: participants and data collection

97 subjects recruited for the study were adult patients (18–89 years old) who were hospitalized and undergoing telemetry monitoring at Emory University Hospital, Emory University Hospital Midtown, and Grady Memorial Hospital (previously described by Shashikumar et al. 2017 (Shashikumar et al. 2017b)). The study was approved by the institutional review board of each hospital. Patients were recruited at random with an over-sampling of patients with AFib; rhythms were reviewed by an ECG technician, physician study coordinator, and cardiologist. 44 subjects had AFib and 53 had other rhythms. Three subjects with AFib had insufficient data to generate subsequent features and were excluded from analysis. Eight channel multiwavelength photoplethysmography (PPG) and tri-axial accelerometry (x, y, z) were recorded simultaneously at a sampling frequency fs of 125 Hz for 5 min using a research version of the wrist-worn Simband device (Samsung, Seoul, South Korea).

2.4. AFib study: data pre-processing

PPG data from a green light wavelength (520–535 nm) were selected because the commercially available version of the Simband contains green light sensors. Data were de-trended, outliers greater than the 95th or less than the 5th percentile were removed, and a 41st order bandpass filter was used with passband 0.0008 – 0.04 Hz. RR intervals were estimated from minima of peaks in the cleaned PPG data, non-physiological RR intervals greater than 2 sec or less than 0.375 sec were removed, and RR intervals occurring less than the sampling period ($1/fs$ sec) after the prior data point were removed.

2.5. RMS energy of acceleration

The RMS energy of acceleration during the i_{th} segment of RR intervals is given by

$$RMSenergy = \sqrt{\frac{\bar{x}^2 + \bar{y}^2 + \bar{z}^2}{N}} \quad (1)$$

where $\bar{x}$, $\bar{y}$, and $\bar{z}$ are x , y , and z -axis accelerometry values in the i^{th} segment, and N is the number of accelerometry data within this segment.

2.6. Statistical moments

The mean, median, mode, variance, skewness, and kurtosis of HR and activity were calculated for both schizophrenia and AFib groups.

2.7. Varying time scales via coarse-graining

Interactions between physiological systems manifest on multiple time scales, and these interactions may differ in healthy versus unhealthy individuals (Ivanov et al. 1999). To assess measures of complexity and interaction over multiple time scales, coarse-grained time series were constructed by averaging the data points within non-overlapping windows of increasing length. The number of time scales τ corresponds to the number of coarse-grainings performed. For the τ^{th} time scale, each element of the coarse-grained time series, $y_j^{(\tau)}$, is given by

$$y_j^{(\tau)} = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} x_i \quad (2)$$

where τ represents the scale factor and $1 \leq j \leq N/\tau$. The first time scale corresponds to the original time series, the second time scale corresponds to one coarse-graining, etc. The number of coarse-grainings for each model was determined via Bayesian optimization, using the training AUC as the cost function.

2.8. Sample entropy

Sample entropy H is a metric of signal complexity, derived from the negative logarithm of the conditional probability of the appearance of longer patterns in a signal, considering the presence of a shorter pattern.

$$H(m, r, N) = -\ln \frac{A^m(r)}{B^m(r)} \quad (3)$$

where m is the template length, r is the radius of similarity or distance threshold between patterns, $A^m(r)$ is a probability of matching a template of length $m+1$, $B^m(r)$ is the probability of matching a template of length m , and N is the number of elements in the time series. Two patterns of length m are considered similar if each point of a pattern in one part of the signal is within a normalized distance r from the respective point in the other part of the signal.

We also calculated multiscale sample entropy (MSE), or H_τ for varying time scales, using the coarse-graining approach described earlier (Costa et al. 2002). Ranges of parameter values ($m = 1, 2, \dots, 4$, $r = 0.01, 0.02, \dots, 1.00$, and $\tau_{\text{max}} = 1, 2, 3$, or 4) were explored via Bayesian optimization, using training AUC as the objective function (Ghassemi et al. 2014; Shahriari et al. 2016). $m_{\text{max}} = 4$ was selected because longer template lengths were computationally expensive to assess, and previous studies of entropy have reported excellent

classification of mental and cardiovascular illness using $m = 2$. $r_{\max} = 1$ was selected because a vector's similarity to an identical vector cannot exceed unity. $\tau_{\max} = 4$ was chosen because preliminary work by our group and others have found diminishing returns after that number of coarse-grainings, as well as increased computational time. Optimal values were $m = 2$, $r = 0.15$, and $\tau_{\max} = 4$. The same parameter values were used for calculating sample entropy for both HR and activity, i.e. $m_{\text{HR}} = m_{\text{act}}$, and $r_{\text{HR}} = r_{\text{act}}$.

2.9. Mutual information

The mutual information of two discrete random variables X and Y , given by $\mathcal{I}(X; Y)$, measures how much knowing one of the two variables reduces uncertainty about the other. Mutual information was calculated over multiple time scales, e.g. multiscale mutual information (MMI). Significant instances of mutual information were determined by Monte Carlo surrogates, i.e., data randomly shuffled in time. For each subject and time scale, mutual information were computed for 100 surrogates. Transfer entropy from the original source time series was deemed to be statistically significant if it was greater than the 95th percentile of the surrogate results (Kantz et al. 2004).

$$\mathcal{I}(X; Y) = \sum \sum p(x, y) \log \left(\frac{p(x, y)}{p(x)p(y)} \right) \quad (4)$$

where $p(x, y)$ is the joint probability function of X and Y , and $p(x)$ and $p(y)$ are the marginal probability density functions of X and Y respectively.

For example, if X and Y are independent, then knowing X does not give any information about Y and vice versa, so their mutual information is zero.

2.10. Darbellay-Vajda (D-V) adaptive partitioning

The computation of transfer entropy and the transformation of time series into a network representation requires estimating joint probability density functions (PDFs). PDFs were estimated via the D-V adaptive partitioning algorithm, in which two time series X and Y are substituted with their ranks ranging from 1 (smallest value) to N (largest value) in sorted X and Y , in a manner similar to some non-parametric statistical tests. The transformed time series of X and Y are $U = \{u_1, u_2, \dots, u_N\}$ and $V = \{v_1, v_2, \dots, v_N\}$. The two-dimensional space defined by u_j - t and v_j - w is then recursively partitioned into squares of varying sizes. Initially the space is divided into four equal quadrants where boundaries are at the mid-points. The null hypothesis that data points are evenly distributed across the four quadrants is tested via the χ^2 statistic (Hudson 2006; Lee et al. 2012):

$$s_{\chi^2} = \sum_{i=1}^4 (M_i - \mu_M^2) \quad (5)$$

where M_i is the number of data points in the i^{th} square and μ_M is the mean number of data points per square. If s_{χ^2} is greater than the critical chi-square statistic value for $p = 0.05$,

$\chi^2_{95\%}$ with $n^2 - 1$ degrees of freedom where n is the number of dimensions or time series being partitioned, the null hypothesis is rejected, the distribution of the data is not uniform, and the partitioning continues such that the quadrant is split into four sub-quadrants. The partitioning process continues recursively until all partitions satisfy the χ^2 test for containing equal proportions of data. If $\chi^2 \leq \chi^2_{95\%}$, the null hypothesis is not rejected, the partitions in the current iteration are discarded, and the current four quadrants are considered to be one partition. Squares that do not contain data do not contribute to the estimation of transfer entropy or network representations of time series.

2.11. Transfer entropy

Transfer entropy given by $\mathcal{T}_{X \rightarrow Y}$ is a measure of directional coupling between two concurrently sampled time series $X = \{x_1, x_2, \dots, x_N\}$ and $Y = \{y_1, y_2, \dots, y_N\}$. Formally, $\mathcal{T}_{X \rightarrow Y}$ is a reduction in uncertainty, given by the conditional entropy of y_i given its past values minus the conditional entropy of y_i given both its past values and past values of the other variable $y_{i-w}^{(l)}$:

$$\begin{aligned} \mathcal{T}_{X \rightarrow Y} &= H(y_i | y_{i-w}^{(l)}) - H(y_i | y_{i-w}^{(l)}, x_{i-t}^{(k)}) \\ &= \sum_{y_i, y_{i-w}^{(l)}, x_{i-t}^{(k)}} p(y_i, y_{i-w}^{(l)}, x_{i-t}^{(k)}) \log \frac{p(y_i | y_{i-w}^{(l)}, x_{i-t}^{(k)})}{p(y_i | y_{i-w}^{(l)})} \\ &= \sum_{y_i, y_{i-w}^{(l)}, x_{i-t}^{(k)}} p(y_i, y_{i-w}^{(l)}, x_{i-t}^{(k)}) \\ &\quad \log \frac{p(y_i, y_{i-w}^{(l)}, x_{i-t}^{(k)}) p(y_{i-w}^{(l)})}{p(y_{i-w}^{(l)}, x_{i-t}^{(k)}) p(y_i, y_{i-w}^{(l)})} \end{aligned} \quad (6)$$

where i indicates a given point in time, t and w are the time lags in X and Y respectively, and k and l are the block lengths of past values in X and Y respectively. k and l were both set to 1 to improve computational speed, and t and w were both set to 1 under the assumption that the maximum auto-transfer of information occurs from the data point in X immediately before the target value in Y , and vice-versa. These choices of $k = l = t = 1$ are appropriate in biomedical experiments where time series length is usually short and the absolute values of auto-correlation functions tend to decrease monotonically as time lag increases (Lee et al. 2012).

Multiscale transfer entropy (MTE) was calculated by coarse-graining HR and activity time series τ times, estimating joint probability distribution functions via D-V partitioning, and

calculating $\mathcal{T}_{HR \rightarrow act}(\tau)$ and $\mathcal{T}_{act \rightarrow HR}(\tau)$ for $\tau = 1, 2, \dots, \tau_{\max}$ time scales. Optimal parameter values of t (time lag in X), w (time lag in Y), and $\tau_{\max}$ were selected via Bayesian optimization (Ghassemi et al. 2014; Shahriari et al. 2016).

Significant information flows were determined by Monte Carlo surrogates, i.e., temporally shuffled time series were created and evaluated for larger values of $\mathcal{T}_{X \rightarrow Y}$ than the original. For each subject, time scale, and parameter value, transfer entropies were computed for 100 surrogates generated by randomizing the order of HR and activity time series. Transfer entropy from the original source time series was considered statistically significant if it was greater than the 95th percentile of the surrogate results.

2.12. Multiscale network representations of time series

A time series partitioned using the D-V algorithm can be mapped to a network consisting of a set of nodes (each corresponding to a unique partition) and an adjacency matrix describing transitions between nodes. Network representations were constructed over varying time scales, e.g. after coarse-graining the input time series data several times.

Topological attributes of the network of heart rate and activity, i.e. multiscale network representation (MSNR) features, were used to classify of illness (Shashikumar et al. 2017a). A map M from time series domain $X \in T$ to a network $g \in G$ can be given by $M: T \Rightarrow G$, where $X = \{X_1, X_2, \dots, X_k\}$, k is the total number of time series being considered, and $X_i \in \mathbb{R}^L$, with L being the length of the time series, and $g = \{S, A\}$ consisting of a set of nodes S and adjacency matrix A . The total number of nodes N correspond to the total number of partitions obtained from the DV partitioning algorithm. Each partition p_i ($i = 1, \dots, N$) is assigned to a node $n_i \in N$ in the graph g . Every data point in X is assigned to one of the partitions. The adjacency matrix A is a $N \times N$ matrix where a_{ij} corresponds to the transition from node n_i to node n_j . Transitions from node n_i and n_j are represented by the weight a_{ij} .

The following network attributes were computed: **number of nodes** (total number of nodes in the network), **number of edges** (total number of edges in the network), **link density** (the total number of edges divided by the maximum possible edges in the network), **average degree** (the average value of the degree of all nodes in the network, where the degree of a node is defined as the total number of its neighboring edges), **number of loops** (the total number of independent loops in the network, also known as the “cyclomatic number” or the number of edges that need to be removed so that the network cannot have cycles), **loop3** (the total number of loops of size 3 in the network), **loop4** (the total number of loops of size 4 in the network), **average clustering coefficient** (the clustering coefficient $c(u)$ for node u can be defined as the ratio of the number of actual edges between the neighbors of u to the number of possible edges between them, and the average clustering coefficient $C(G)$ of a network is the average of $c(u)$ taken over all the nodes in the network), **Pearson coefficient** (the Pearson degree correlation of a network), **algebraic connectivity** (the second smallest Eigenvalue of the Laplacian matrix of a network, where the Laplacian matrix is the difference between the sum of degrees of the diagonal elements in adjacency matrix and the adjacency matrix), **closeness** (the closeness centrality, $cc(u)$ for node u is the inverse of sum of distance from node u to all other nodes in the network, where the closeness centrality of a

graph is the average mean of the above is the average of $cc(u)$ taken over all the nodes in the network, **average eccentricity** (eccentricity of a node u is defined as $e(u) = \max\{d(u, v) : v \in V\}$, where the distance $d(u, v)$ is the length of the shortest path from u to v , and V is the set of all nodes. The average effective eccentricity is the average of effective eccentricities over all nodes in the network), **maximum effective eccentricity** (Also known as the effective diameter, is defined as the maximum value of effective eccentricity over all nodes in the graph), **spectral radius** (defined as the largest magnitude eigenvalue of the adjacency matrix of the network), **trace** (sum of the eigenvalues of the adjacency matrix, i.e., λ , and **energy** (squared sum of the eigenvalues of the adjacency matrix A , i.e. $E(G) = \sum_i \lambda_i^2$).

2.13. Binary classification of illness status

Features were used to train a support vector machine (SVM) algorithm with a linear kernel to classify subjects into the schizophrenia or healthy control class, i.e. to perform a binary discrimination task. Classifier performance was assessed via subject-wise leave-one-out crossfold validation (LOOCV). Given N patients, $N - 1$ patients are used to train the classifier and the remaining patient is used as the test set. Features in the training set were transformed to have Gaussian distributions using either the identity, square root, or logarithmic transformations. The transformation resulting in the most normal data, e.g. the lowest k-statistic using the Lilliefors test, was determined from the training set and then applied to the test set to prevent leakage of information. Data in both training and test sets were normalized by subtracting the training mean and dividing by the training standard deviation. Predictions for each subject – defined as the probability of having a diagnosis of schizophrenia – were pooled across crossfolds to report a single pooled area under the receiver operating curve (AUC; Airola et al. 2009). AUCs were calculated and reported for both training and test sets, and different models (i.e. the set of features used to train the SVM) were compared by calculating the integrated discrimination improvement (IDI; Pencina et al. 2008), given by:

$$IDI = (IS_{\text{new}} - IS_{\text{old}}) - (IP_{\text{new}} - IP_{\text{old}}) \quad (7)$$

where IS is the integral of sensitivity over all possible cut-off values over (0,1) interval, IP is the integral of 1-specificity, and “new” and “old” refer to the two models being compared (Pencina et al. 2008). An asymptotic test for the null hypothesis of $IDI = 0$ was performed, and the P-value reported:

$$z = \frac{\widehat{IDI}}{\sqrt{(\widehat{se}_{\text{events}})^2 + (\widehat{se}_{\text{nonevents}})^2}} \quad (8)$$

3. Results

3.1. Mutual information

Mutual information between HR and activity ($\mathcal{I}(HR; act)$) was calculated over four time scales τ_{1-4} , with the maximum timescale determined via Bayesian optimization. $\mathcal{I}(HR; act)$ for patients with schizophrenia and controls and compared via the two-sided Wilcoxon rank-sum test. $\mathcal{I}(HR; act)$ significantly different for the first three timescales ($P < 0.05$), whereas the difference in medians was not significant for $\tau = 4$ (Figure 2A). The opposite trend was observed in the AFib group; for all time scales, patients exhibited significantly lower values of $\mathcal{I}(HR; act)$ compared to controls ($P < 0.05$; Figure 2B).

Significance of mutual information between HR and activity was estimated using surrogates whereby each time series was shuffled 100 times and $\mathcal{I}_{surrogate}$ was calculated for each instance. A mutual information ratio was calculated. The numerator was the 95th percentile of $\mathcal{I}_{surrogate}$ and the denominator was $\mathcal{I}$ of the original data. For all time scales, this ratio was significantly less than the red dashed line of unity for patients with schizophrenia and controls, demonstrating significant mutual information between HR and activity (supplemental Figure S1). In contrast, patients with AFib displayed mutual information ratio metrics near or greater than unity, suggesting the observed values of $\mathcal{I}$ could have been due to random chance. However, for control subjects the mutual information ratio was close to or slightly below unity for all time scales except $\tau = 4$, for which the ratio was slightly above unity.

3.2. Transfer entropy

MTE from HR to activity ($\mathcal{T}_{HR \rightarrow act}$) and from activity to HR ($\mathcal{T}_{act \rightarrow HR}$) were calculated for patients and controls in both the schizophrenia and AFib groups. $\mathcal{T}_{HR \rightarrow act}$ was higher in patients than in controls for the first three time scales ($\tau = 1, 2, 3$), but did not differ for $\tau = 4$ (Figure 3A). Similarly, $\mathcal{T}_{act \rightarrow HR}$ was higher in patients than in controls but for all time scales (Figure 3B). In the AFib group, both $\mathcal{T}_{HR \rightarrow act}$ and $\mathcal{T}_{act \rightarrow HR}$ were lower in patients with AFib than in controls for all time scales (Figure 3C & D).

Following a similar approach as described earlier, a transfer entropy ratio was calculated, with the numerator being $\mathcal{T}_{HR \rightarrow act}$ of the original data and the denominator being the 95th percentile of $\mathcal{T}_{HR \rightarrow act, surrogates}$ (supplemental Figure S2). Patients with schizophrenia had ratios significantly less than unity, suggesting the observed values of $\mathcal{T}_{HR \rightarrow act}$ were significant. Controls in the schizophrenia group had slightly lower ratios, with 95% confidence interval bounds crossing 1 for several time scales, suggesting less directed information transfer from HR to activity in healthy people. In contrast, $\mathcal{T}_{HR \rightarrow act}$ from both AFib patients and controls were greater than 1, suggesting $\mathcal{T}_{HR \rightarrow act}$ was not significant in that group. Finally, the ratio for $\mathcal{T}_{act \rightarrow HR}$ was much less than 1 for both patients and controls in both schizophrenia and AFib groups, indicating changes in activity consistently preceded changes in HR regardless of illness type or status.

3.3. Network representations of time series

MSNR were constructed from time series data. The number of time scales that maximized classifier performance was selected via Bayesian optimization (Ghassemi et al. 2014; Shahriari et al. 2016). For the schizophrenia group, three time scales were optimal (Figure 4A), whereas for the AFib group, the first time scale was optimal (Figure 4B). Gross differences in network structure, measured by complexity, node count and edge count, varied both by patient type and time scales.

3.4. Classifier performance

Nine feature groups were used to train a support vector machine: 1) statistical moments, 2) MSE, 3) MMI, 4) MTE, 5) MSNR, 6) MSE and MTE, 7) MSE and MSNR, 8) MTE and MSNR, and 9) MTE, MSE, and MSNR. LOOCV was performed to assess classifier performance. Receiver operating characteristic curves (ROCs) were plotted for schizophrenia (Figure 5A) and AFib groups (Figure 5B), and areas under the ROCs (AUCs) were reported for training and test sets (Table 1).

In the schizophrenia group, MSE achieved high test AUCs, consistent with our previous work (Osipov et al. 2015). However, MSNR features – alone or in combination with any other feature – resulted in perfect classifier performance in both training and test data (Table 1). The performance of various feature types was compared using the IDI. MSNR was significantly better than MMI as well as MTE ($P < 0.05$; Table 2).

In the AFib group, MSE or any feature combinations including MSE achieved the maximum test AUC. MSE even outperformed MTE, indicating signal complexity was more predictive of disease compared to measures of interaction between HR and activity ($P < 0.05$; Table 2). Finally, MSNR was significantly better than MMI, but not MSE or MTE.

4. Discussion

We assessed interactions between HR and activity by calculating mutual information, transfer entropy, and network representations of time series over multiple time scales. Constructing a network from HR and activity time series is a novel approach that utilizes the D-V partitioning algorithm, which is computationally fast and does not require the specification of as many hyperparameters as other partitioning methods (Hudson 2006). Measures of interactions between HR and activity as well as attributes of each signal were calculated and used to train a machine learning algorithm to classify schizophrenia subjects from controls. Perfect classification accuracy was achieved in the schizophrenia group using MSNR features, whereas combined univariate analyses on separate HR and activity resulted in lower AUCs. On the other hand, network features did not add significant differentiating power to the classifier when evaluating a non-mental population with AFib. To our knowledge, this is the first use of interactions between HR and activity to distinguish patients from controls.

Univariate features such as statistical moments and MSE were less predictive compared to interaction features, yet enabled classification of schizophrenia significantly better than

chance. MSE outperformed statistical moments in both schizophrenia and AFib groups. Measures of complexity of physiological and behavioral time series may better capture loss in system autoregulation compared to simpler features such as statistical moments (Berle et al. 2010; Montaquila et al. 2015). These results are consistent with our previous work on classifying patients with schizophrenia from healthy controls by learning univariate features of complexity where the test AUC did not reach 1.00 (Reinertsen et al. 2017).

Mutual information $\mathcal{I}$ quantifies linear and nonlinear dependence between two variables (Duncan 1970). If HR contains information about activity, or vice-versa, $\mathcal{I}$ will be > 0 . $\mathcal{I}$ is likely to be nonzero in the groups studied here because activity can lead to an increase in HR due to a rise in peripheral oxygen demand, and increased HR can precede a rise in activity due to a behavioral response to external cues. We assessed if $\mathcal{I}$ across several time scales contributed to classifier performance. MMI resulted in lower classification performance compared to other types of features, with a test AUC of 0.81 in the schizophrenia group and 0.77 in the AFib group (Figure 2).

To assess if mutual information between HR and activity was due to random chance, we shuffled each time series 100 times, calculated $\mathcal{I}_{surrogate}$ each time, and calculated the ratio of $\mathcal{I}$ of the original data to the 95th percentile of $\mathcal{I}_{surrogate}$. This ratio was far less than unity for both patients and controls in the schizophrenia group, suggesting significance. Interestingly, this was not the case for the AFib group; the median ratio was close to and slightly above unity for patients whereas the median was below or close to one for controls. These data suggest a modest reduction in coupling between HR and activity in patients with AFib compared to controls within the same group, but more broadly demonstrate a difference in mutual information by group that may be due to measurement method (Proteus patch for schizophrenia subjects versus SimBand smartwatch for AFib subjects) rather than illness class. However, differences between the two groups did not introduce bias, as the classification task was to dichotomize patients from controls within the same group, rather than to distinguish schizophrenia, AFib, and healthy controls in a trinary classification task.

$\mathcal{I}$ is symmetric in that it does not capture directionality of information transfer. Directed information transfer may be a more predictive feature. To assess the directional flow of information between HR and activity, we calculated transfer entropy $\mathcal{T}_{HR \rightarrow act}$ and $\mathcal{T}_{act \rightarrow HR}$ (Schreiber 2000). Both $\mathcal{T}_{HR \rightarrow act}$ and $\mathcal{T}_{act \rightarrow HR}$ were higher in patients with schizophrenia than controls for all time scales (Figure 3).

We assessed significance of transfer entropy using the same surrogate shuffling method. The median ratio of the 95th percentile of $\mathcal{T}_{surrogate}$ to $\mathcal{T}_{HR \rightarrow act, original}$ was less than unity for all time scales and for both patients with schizophrenia and controls, although ratios were slightly lower in patients with schizophrenia versus controls (supplemental Figure S2). Similar to the surrogate analysis of $\mathcal{I}$, this ratio was slightly greater than unity for patients and controls in the AFib group. In contrast, $\mathcal{T}_{act \rightarrow HR}$ was equal to or very close to zero for all subjects in both the schizophrenia and AFib groups, demonstrating significant directed transfer of information from activity to HR regardless of illness group or control status. In

summary, information theoretical measures by illness status were significant as determined via univariate Wilcoxon rank sum testing. Furthermore, the observed values of these measures were not due to random chance, as determined via surrogate time series shuffling.

In both the schizophrenia and AFib groups, MTE outperformed statistical moments and MMI (Table 1), indicating directed information transfer between HR and activity was a more predictive feature than simple attributes of distributions of variables, or the mere presence of asymmetric information flow. A change in HR following a concordant change in activity – e.g. both increase or both decrease – is partially mediated by the ANS, which suffers dysfunction in people with schizophrenia. The physiological underpinnings of the converse scenario whereby a change in activity follows a change in HR remains challenging to understand, and our data did not provide sufficiently granular details of the patients' activities to conjecture further.

Surprisingly, MMI or MTE features enabled classification of AFib better than random chance. The AFib group served as a control; patients did not have psychiatric illness and were assessed in a seated position in a clinical lab that ostensibly reduces HR-activity interactions. However, AFib can cause symptoms such as palpitations and dyspnea (Lip et al. 2016), which could be detectable by a wristband although this has never been rigorously studied in the literature.

Directed transfer of information between two signals likely occurs at specific regions in time rather than constantly throughout. The distribution of these regions could be a more nuanced and predictive feature of illness than global measures of information transfer. However, such distributions are not captured by mutual information or transfer entropy of entire time series. To better assess underlying system dynamics and interactions between signals, we constructed network representations of HR and activity using the approach described by Shashikumar et al. (Shashikumar et al. 2017a). Each node in the network represents a physiological and behavioral state, and was formed from a partition in a six-dimensional space comprised of lagged forms of HR and activity: 1) $HR(t)$, 2) $HR(t-1)$, 3) $HR(t-2)$, 4) $act(t)$, 5) $act(t-1)$, and 6) $act(t-2)$. Our approach exploits Takens' theorem, which describes how a dynamical system can be reconstructed from a sequence of lagged observations, given a sufficient embedding dimension D (Takens 1981). Although the optimal D is unknown, a lagged embedding approach can yield more information about the properties of a dynamical system compared to analyzing only the observed time series.

Networks from a representative subject with schizophrenia demonstrated 60, 57, and 21 nodes for the first, second, and third timescales respectively (Figure 4). In contrast, networks from a healthy control demonstrated 80, 41, and 36 nodes. Although the number and connectivity of nodes exhibit variance across subjects within an illness group, networks from patients with schizophrenia visually differed from networks derived from controls. Because the maximum number of edges in a directed graph with n nodes is $n(n-1)$, even slight differences in node count are amplified in other network properties that correlate with edge count, connectivity, and complexity. These data suggest a decreased number of physiological states and transitions in schizophrenia, consistent with previous reports of more structured

behavioral patterns and decreased markers of HRV compared to healthy controls (Chang et al. 2009; Berle et al. 2010; Montaquila et al. 2015).

On the other hand, networks from patients with AFib featured similar numbers of nodes but greater connectivity between nodes compared to controls. The irregular rhythm of severe AFib may increase the number of physiological states. However, actual events of AFib can be rare even in diagnosed patients, thus not significantly adding to the number of HR-activity states. Despite a similar number of nodes, transitions between nodes were sufficiently altered in AFib to enable discrimination of patients from controls.

Corroborating the visual differences in networks, MSNR features enabled the classifier to perfectly discriminate patients with schizophrenia from healthy controls, with train and test AUCs of 1.00 (Table 1 and Figure 5). However, MSNR features did not outperform MTE features for AFib patients. These results indicate a difference in interactions and temporal structure between HR and activity in patients with mental versus cardiovascular illness.

Finally, a comparison of significant differences in classifier performance via IDI in the schizophrenia group demonstrated statistically significant improvements of MTE versus MMI, and MSNR versus MMI, with P-values less than 0.05 (Table 2). In the AFib group, MMI versus MSE and MSNR versus MMI were significantly different. MSNR outperformed MTE in the schizophrenia group, but the two models were equivalent in the AFib group. These results indicate MTE and MSNR capture different aspects of coupling between HR and activity. Furthermore, these results are consistent with our initial hypothesis that interactions between HR and activity are different in patients with mental versus cardiovascular illness.

The probability of a statistically significant difference will increase in a larger group, but effect sizes of each type of feature may also differ in mental versus cardiovascular patients, so these results should not be over-interpreted. Moreover, although nonintuitive, a lower P-value of an individual feature does not necessarily correspond to a more predictive feature (Lo et al. 2015). It should be noted that AUCs from different models can be definitively compared without significance tests when performing LOOCV – a model achieving an AUC of 1.00 perfectly classifies all subjects, and this performance is deterministic and repeatable insofar as the features used to train the model are non-stochastic and do not vary across experiments.

Network representations appeared to capture more illness-related information concerning underlying system dynamics compared to other information theoretical measures of complexity and interaction. Campanharo et al. qualitatively evaluated network representations of chaotic Lorenz and Rossler equations, but the relationship between network attributes and properties seen in physiological data such as noise, autocorrelation, periodicity, and non-stationarity remain unknown (Campanharo et al. 2011). Simple dynamical systems can be generated with known and varying levels of these properties, and attributes of the resulting network representations can be studied. Understanding this mapping could yield insight about the physiological meaning of altered interactions between HR and activity time series in mental illness.

A change in HR following a concordant change in activity – e.g. both increase or both decrease – is partially mediated by the ANS, which has been shown to be dysregulated in people with schizophrenia. The physiological underpinnings of the converse scenario whereby a change in activity follows a change in HR remains challenging to understand, and our data did not provide sufficiently granular details of the patients' activities to conjecture further.

We note several limitations of this study. Each group was small, consisting of 16 patients with schizophrenia and 19 controls, and 41 AFib patients and 53 controls. We used LOOCV to estimate generalizability of our classifier, but performance metrics may differ for a new group. Comprehensive feature selection was not performed; rather, features from one or several feature groups in combination were compared. An SVM classifier was used, which although can often demonstrate strong results, may not be the best choice. Although a different machine learning algorithm could achieve better classification performance, more complex algorithms are likely to overfit and could not give better results than those reported in this study (AUC=1.0). Estimated joint probability distributions were calculated over all data despite known non-stationarity in human HR and locomotor activity. Entropy and interaction measures calculated from these pooled estimates will thus deviate from the actual measures. Entropy estimates have been shown to vary in the presence of non-stationarities such as slow trends, spikes, local variance changes, and long-range correlations (Xiong et al. 2017). Nonstationarities in time series could be minimized by removing slow trends via a high-pass filter and normalizing data to zero mean and unit variance; however, the effect of doing so on machine learning classification performance has yet to be evaluated. The fact that the data have been heavily averaged every 10 minutes creates a band pass effect and mitigates for the non-stationarities to some extent. We also note that the nonstationarities are similar in prevalence and type among patients of all disease severity as well as controls, classifier performance would be minimally impacted. Lastly, because network representations of a time series capture temporal information, graph theoretical attributes are likely to also be sensitive to non-stationarities, but this has yet to be rigorously evaluated and is an important topic of future work in the broader discipline of network physiology (Bashan et al. 2012). Nevertheless, we note that an earlier study showed the superiority of the D-V approach in estimating joint probabilities even in non-stationary time series (Lee et al. 2012).

Other methods have been used to assess complexity and interaction between signals. Refined MSE has been proposed to address aliasing introduced by coarse-graining and to update the radius of similarity using the new standard deviation of each coarse-grained time series (Valencia et al. 2009). Estimated MSE values in our study probably contain contributions from alias artifact. However, in our cohort of patients with schizophrenia, MSE features alone resulted in a high AUC of 0.96. Any aliasing artifact must either be equally present in both patients and controls and thus does not hinder separability of the two classes, or is only present in one class and thus contributes to excellent separability based on MSE features. Predictability-based methods are useful for predicting future states or to comparing methods of assessing causal interactions in a dynamical system (Porta et al. 2016). Frequency-based methods assess spectral information and can be used to explore a specific physiological hypothesis (Porta et al. 2014). We did not evaluate spectral HRV features for several reasons.

The patches used in the schizophrenia study recorded a mean HR every 10 minutes, a frequency that precludes the estimation of spectral HRV measures. Beat-to-beat measurements via non-Holter ECG wearables such as patches and watches can be problematic and of lower quality. Finally, entropy and MSNR outperformed simpler features such as statistical moments, and kurtosis of HR directly correlates with frequency-domain HRV measures (Clifford 2002). In summary, we emphasize the objectives of this study were not to compare the performance of various analytical methods, examine physiological mechanisms that generate oscillations, nor to compare model-free versus model-based approaches. Rather, we sought to identify if there was some interaction between HR and locomotor activity over multiple long time scales that was manifest in a psychiatric population and added value to conventional univariate approaches.

Alterations of locomotor activity in schizophrenia have been attributed to cognitive dysmetria, or functional disconnectivity between neural centers of cognition, motor function, and coordination (Honey et al. 2005). However, most studies of schizophrenia do not involve data collection at a sufficiently high frequency to discern more nuanced measures such as directed transfer of information. Further research is needed to elucidate the mechanisms governing interactions between physiology and behavior in schizophrenia and other mental illnesses. Better understanding of these interactions could enable clinically useful applications such as predicting relapse or remotely monitoring efficacy of new therapeutics.

5. Conclusion

We demonstrated measures of multiscale interactions – mutual information, transfer entropy, and novel network representations of states – between HR and locomotor activity enable discrimination of patients with schizophrenia from controls. We repeated this approach using data from patients with AFib, and found network and interaction features did not improve prediction over complexity measures that ignored interactions between HR and activity. To our knowledge this is the first evaluation of interactions between physiological and behavioral states in a mental health population using data measured via objective, affordable, and non-invasive wearable devices.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

6. Acknowledgments

We acknowledge the support of the National Science Foundation Award 1636933, National Institutes of Health (Grants K23 HL127251, P50 HL117929, K01 ES025445, and R01 HL136205). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation or the National Institutes of Health. We thank John M. Kane, Georgios Petrides, and Yashar Behzadi for collecting schizophrenia data, and Qiao Li for the pulse detection and signal quality code used in this work.

7. References

Airola A and Pahikkala T (2009). “A comparison of AUC estimators in small-sample studies”. *Journal of Machine Learning Research* 8, pp. 3–13.

- Alvares GA, Quintana DS, Hickie IB, et al. (2016). "Autonomic nervous system dysfunction in psychiatric disorders and the impact of psychotropic medications: a systematic review and meta-analysis". *Journal of Psychiatry & Neuroscience* 412, pp. 89–104. [PubMed: 26447819]
- Bashan A, Bartsch RP, Kantelhardt JW, et al. (2012). "Network physiology reveals relations between network topology and physiological function". *Nature Communications* 3, pp. 702–709.
- Berle JO, Hauge ER, Oedegaard KJ, et al. (2010). "Actigraphic registration of motor activity reveals a more structured behavioural pattern in schizophrenia than in major depression". *BMC Research Notes* 3, p. 149. [PubMed: 20507606]
- Campanharo ASLO, Sirer MI, De Malmgren RD, et al. (2011). "Duality between time series and networks". *PLoS ONE* 68, pp. 1–12.
- Chang JS, Yoo CS, Yi SH, et al. (2009). "Differential pattern of heart rate variability in patients with schizophrenia". *Progress in Neuro-Psychopharmacology and Biological Psychiatry* 336, pp. 991–995. [PubMed: 19427888]
- Clifford GD (2002). "Signal processing methods for heart rate variability". PhD thesis University of Oxford.
- Costa M, Goldberger A, and Peng C-K (2002). "Multiscale entropy analysis of complex physiologic time series". *Physical Review Letters*.
- Duncan TE (1970). "On the calculation of mutual information". *SIAM Journal on Applied Mathematics* 191, pp. 215–220.
- Ghassemi M, Lehman LH, Snoek J, et al. (2014). "Global optimization approaches for parameter tuning in biomedical signal processing: a focus of multi-scale entropy". *Computers in Cardiology*, 41:993–996.
- Hall J (2010). *Guyton and Hall Textbook of Medical Physiology*. 12th ed. Elsevier.
- Honey GD, Pomarol-Clotet E, Corlett PR, et al. (2005). "Functional dysconnectivity in schizophrenia associated with attentional modulation of motor function". *Brain* 12811, pp. 2597–2611. [PubMed: 16183659]
- Hudson JE (2006). "Signal Processing Using Mutual Information". *IEEE Signal Processing Magazine* 236, pp. 50–54.
- Ivanov PC, Amaral L.a. N, Goldberger AL, et al. (1999). "Multifractality in human heartbeat dynamics". *Nature* 3996735, pp. 461–465. [PubMed: 10365957]
- Kantz H and Schreiber T (2004). *Nonlinear time series analysis*. Cambridge University Press.
- Lee J, Nemati S, Silva I, et al. (2012). "Transfer Entropy Estimation and Directional Coupling Change Detection in Biomedical Time Series". *Biomedical Engineering Online* 111, p. 19. [PubMed: 22500692]
- Lip GYH, Fauchier L, Freedman SB, et al. (2016). "Atrial fibrillation". *Nature Reviews Disease Primers* 2, p. 16016.
- Lo A, Chernoff H, Zheng T, et al. (2015). "Why significant variables aren't automatically good predictors". *Proceedings of the National Academy of Sciences* 11245, pp. 13892–13897.
- Millar A, Espie CA, and Scott J (2004). "The sleep of remitted bipolar outpatients: A controlled naturalistic study using actigraphy". *Journal of Affective Disorders* 802–3, pp. 145–153. [PubMed: 15207927]
- Montaquila JM, Trachik BJ, and Bedwell JS (2015). "Heart rate variability and vagal tone in schizophrenia: A review". *Journal of Psychiatric Research* 699, pp. 57–66. [PubMed: 26343595]
- NICE guideline (CG178) (2014). *Psychosis and schizophrenia in adults: prevention and management*. Tech. rep National institute for Health and Care Excellence.
- Osipov M, Behzadi Y, Kane JM, et al. (2015). "Objective identification and analysis of physiological and behavioral signs of schizophrenia". *Journal of Mental Health* 245, pp. 276–282. [PubMed: 26193048]
- Pencina MJ, D'Agostino RB, Sr, D'Agostino RB, Jr, et al. (2008). "Evaluating the added predictive ability of a newmarker: From area under the ROC curve to reclassification and beyond". *Statistics in medicine* 27, pp. 157–172. [PubMed: 17569110]
- Porta A and Faes L (2016). "Wiener-Granger Causality in Network Physiology with Applications to Cardiovascular Control and Neuroscience". *Proceedings of the IEEE* 1042, pp. 282–309.

- Porta A, Faes L, Bari V, et al. (2014). "Effect of age on complexity and causality of the cardiovascular control: Comparison between model-based and model-free approaches". *PLoS ONE* 9.2.
- Reinertsen E, Osipov M, Liu C, et al. (2017). "Continuous assessment of schizophrenia using heart rate and accelerometer data". *Physiological Measurement* 387, pp. 1456–1471. [PubMed: 28653659]
- Saha S, Chant D, Welham J, et al. (2005). "A systematic review of the prevalence of schizophrenia". *PLoS Medicine* 25, p. 141.
- Saxena S, Thornicroft G, Knapp M, et al. (2007). "Resources for mental health: scarcity, inequity, and inefficiency". *Lancet* 3709590, pp. 878–889. [PubMed: 17804062]
- Schreiber T (2000). "Measuring Information Transfer". *Physical Review Letters* 852, p. 461. [PubMed: 10991308]
- Shahriari B, Swersky K, Wang Z, et al. (2016). "Taking the human out of the loop: a review of Bayesian optimization". *Proceedings of the IEEE*.
- Shashikumar SP, Li Q, Clifford GD, et al. (2017a). "Multiscale network representation of physiological time series for early prediction of sepsis". *Physiological Measurement*.
- Shashikumar SP, Shah AJ, Li Q, et al. (2017b). "A Deep Learning Approach to Monitoring and Detecting Atrial Fibrillation using Wearable Technology". *IEEE EMBS International Conference on Biomedical & Health Informatics*, pp. 141–144.
- Steinhubl SR, Muse ED, and Topol EJ (2015). "The emerging field of mobile health". *Science Translational Medicine* 7283, 283rv3.
- Takens F (1981). "Detecting strange attractors in turbulence". *Lecture notes in mathematics* 8981, pp. 366–381.
- Valencia JF, Porta A, Vallverdu M, et al. (2009). "Refined multiscale entropy: Application to 24-h holter recordings of heart period variability in healthy and aortic stenosis subjects". *IEEE Transactions on Biomedical Engineering* 569, pp. 2202–2213. [PubMed: 19457745]
- Van Someren EJW, Swaab DF, Colenda CC, et al. (1999). "Bright Light Therapy: Improved Sensitivity to Its Effects on Rest-Activity Rhythms in Alzheimer Patients by Application of Nonparametric Methods". *The Journal of Biological and Medical Rhythm Research* 164.
- Xiong W, Faes L, and Ivanov PC (2017). "Entropy measures, entropy estimators, and their performance in quantifying complex dynamics: Effects of artifacts, non-stationarity, and long-range correlations". *Physical Review E* 956, pp. 1–37.

Figure 1.

Schematic of data processing and classification algorithm. DV partitions are computed from time-lagged and coarse-grained HR and locomotor activity, which are transformed to a network representation. Topological attributes of the networks are used as input features to a machine learning classifier. DV partitions are also used to compute transfer entropy for between HR and locomotor activity (and vice-versa) for varying lags and time scales. Finally, mutual information and sample entropy are calculated for varying time scales. These features are used to train a classifier to estimate the probability of a subject belonging to the unhealthy class.

Figure 2.

MMI between HR and activity for A) patients with schizophrenia and controls, and B) AFib patients and controls. Data is shown via notched box plots; the horizontal red line denotes the median, the notches denote 95th percent confidence intervals of the median, borders of the blue box denote the 25th and 75th percentiles, and the lower and upper whiskers denote the minimum and maximum values, respectively. The y-axis is mutual information in bits, and the x-axis denotes different time scales and sick versus healthy subjects. Asterisks indicates $P < 0.05$ via the Wilcoxon rank-sum test.

Figure 3.

MTE from A) HR to activity ($\mathcal{T}_{HR \rightarrow act}$) for patients with schizophrenia and controls, B) activity to HR ($\mathcal{T}_{act \rightarrow HR}$) for patients with schizophrenia and controls, C) A) HR to activity ($\mathcal{T}_{HR \rightarrow act}$) for AFib patients and controls, and D) activity to HR ($\mathcal{T}_{act \rightarrow HR}$) for AFib patients and controls. Data is shown via notched box plots; the horizontal red line denotes the median, the notches denote 95th percent confidence intervals of the median, borders of the blue box denote the 25th and 75th percentiles, and the lower and upper whiskers denote the minimum and maximum values, respectively. The y-axis is transfer entropy in bits, and the x-axis denotes different time scales and sick versus healthy subjects. Asterisks indicates $P < 0.05$ via the Wilcoxon rank-sum test.

Figure 4.

MSNR of HR and activity data; each colored circle represents a six-dimensional state defined by a value of HR (e.g. 74 BPM), locomotor activity (e.g. RMS of accelerometry value of 1.7), two time-lagged values of HR, and two time-lagged values of activity. Thus, each state represents a temporal trajectory through physiological and behavioral states. Lines between nodes denote transitions in time from one node to the next. A) Network representations of data from a single subject with schizophrenia (denoted in red) demonstrate a higher number of physiological and behavioral states at τ_2 , and a lower number of states at τ_3 , compared to states from a healthy control subject (denoted in blue). τ_i indicates the i_{th} time scale. B) Network representations of data from a single subject with AFib (denoted in red) demonstrate a higher number of physiological and behavioral states and more state transitions compared to a healthy control subject (denoted in blue). The properties of these networks were quantified using graph theoretical approaches, and these properties were used as features to train a support vector machine to classify patients from healthy controls.

Figure 5.

ROC curves of models for classifying patients with A) schizophrenia or B) AFib from healthy controls using combinations of different features. Stat moments is statistical moments, MSE is multiscale entropy, MMI is multiscale mutual information, and MSNR is multiscale network representations. The Y-axis is the sensitivity and the X-axis is 1 – specificity.

Table 1.

AUCs indicating classifier performance for nine feature groups, or models. The model is described in column 1, results from the schizophrenia group are reported in columns 2–3, and results from the AFib group are reported in columns 4–5.

Model	Schizophrenia		Atrial fibrillation	
	Training AUC	Test AUC	Training AUC	Test AUC
Statistical moments	0.98	0.84	0.91	0.89
MSE	0.98	0.95	0.94	0.93
MMI	0.90	0.81	0.81	0.77
MTE	0.98	0.90	0.94	0.90
MSNR	1.00	1.00	0.92	0.90
MSE + MTE	1.00	0.98	0.97	0.93
MSE + MSNR	1.00	1.00	0.96	0.90
MTE + MSNR	1.00	1.00	0.97	0.92
MTE + MSE + MSNR	1.00	1.00	0.99	0.93

Table 2.

Comparison of model performance on test set data via the IDI. A positive IDI with $P < 0.05$ indicates the new model achieves a statistically significant improvement in classification performance versus the old model. Models are listed in column 1, results from the schizophrenia group are reported in columns 2–3, and results from the AFib group are reported in columns 4–5.

Models	Schizophrenia		Atrial fibrillation	
	IDI	P-value	IDI	P-value
MSE vs. MMI	−0.316	0.073	−0.297	0.005
MMI vs. MTE	0.182	0.200	0.102	0.349
MMI vs. MSNR	0.449	0.009	0.243	0.026
MSE vs. MTE	−0.133	0.420	−0.195	0.035
MSE vs. MSNR	0.133	0.317	−0.054	0.452
MTE vs. MSNR	0.267	0.024	0.141	0.171