Dynamic fMRI networks predict success in a behavioral weight loss program among older adults

Fatemeh Mokhtari [a,b], W. Jack Rejeski [a,c,d,e], Yingying Zhu [f], Guorong Wu [f], Sean L. Simpson [a,b,g], Jonathan H. Burdette [a,c], Paul J. Laurienti [a,c,*]

[a] Laboratory for Complex Brain Networks, Department of Radiology, Wake Forest University School of Medicine, Winston Salem, NC, USA
[b] Virginia Tech-Wake Forest University School of Biomedical Engineering and Sciences, Wake Forest University School of Medicine, Winston Salem, NC, USA
[c] Translational Science Center, Wake Forest University, Winston Salem, NC, USA
[d] Department of Health and Exercise Science, Wake Forest University, Winston Salem, NC, USA
[e] Department of Geriatric Medicine, Wake Forest University, Winston Salem, NC, USA
[f] Department of Radiology and BRIC, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
[g] Department of Biostatistical Sciences, Wake Forest School of Medicine, Winston Salem, NC, USA

ARTICLE INFO

ABSTRACT

More than one-third of adults in the United States are obese, with a higher prevalence among older adults. Obesity among older adults is a major cause of physical dysfunction, hypertension, diabetes, and coronary heart diseases. Many people who engage in lifestyle weight loss interventions fail to reach targeted goals for weight loss, and most will regain what was lost within 1–2 years following cessation of treatment. This variability in treatment efficacy suggests that there are important phenotypes predictive of success with intentional weight loss that could lead to tailored treatment regimen, an idea that is consistent with the concept of precision-based medicine. Although the identification of biochemical and metabolic phenotypes are one potential direction of research, neurobiological measures may prove useful as substantial behavioral change is necessary to achieve success in a lifestyle intervention. In the present study, we use dynamic brain networks from functional magnetic resonance imaging (fMRI) data to prospectively identify individuals most likely to succeed in a behavioral weight loss intervention. Brain imaging was performed in overweight or obese older adults (age: 65–79 years) who participated in an 18-month lifestyle weight loss intervention. Machine learning and functional brain networks were combined to produce multivariate prediction models. The prediction accuracy exceeded 95%, suggesting that there exists a consistent pattern of connectivity which correctly predicts success with weight loss at the individual level. Connectivity patterns that contributed to the prediction consisted of complex multivariate network components that substantially overlapped with known brain networks that are associated with behavior emergence, self-regulation, body awareness, and the sensory features of food. Future work on independent datasets and diverse populations is needed to corroborate our findings. Additionally, we believe that efforts can begin to examine whether these models have clinical utility in tailoring treatment.

Introduction

Obesity and its adverse health effects are highly prevalent and uniquely problematic among older adults (65 + years) (Flegal et al., 2012). Obesity is associated with declines in physical functioning, the primary driver of disability and loss of independence with aging (Mathus-Vliegen et al., 2012). Unfortunately, although effective lifestyle weight loss interventions in overweight and obese adults produce mean

weight loss in the range of 5–10% of baseline weight (Rejeski et al., 2011), there is considerable variability in how much people lose (Curioni and Lourenco, 2005; MacLean et al., 2011). In fact, understanding potential phenotypes for success in behavioral weight loss has been identified as a critical area for research by the National Institutes of Health (http://1.usa.gov/1VCI7pA). In the present study, we used baseline (pre-intervention) functional magnetic resonance imaging (fMRI) dynamic networks as a biomarker to identify those older adults that

---

exhibited the greatest success in an 18-month behavioral weight loss intervention.

Although knowledge about nutrition and physical activity is essential to behavioral weight loss, arguably more important are the self-regulatory skills and motivation required to enact needed behavior change (Rejeski et al., 2008; Wing, 2002). Consistent with this perspective are studies showing that high self-efficacy for regulating eating behavior, the ability to manage stress, and motivation to lose weight are positively associated with better initial weight loss and weight maintenance (Elfhag and Rössner, 2005; Williams et al., 1996). More relevant to the current study are data showing that brain activity in neural networks including the hippocampus/amygdala, dorsal striatum, orbitofrontal, prefrontal and anterior cingulate cortices has been shown to be important in overeating and food craving (Berridge et al., 2010; Rothemund et al., 2007; Stice et al., 2013). Specific to behavioral weight loss treatment, we have observed that following an overnight fast, global connectivity of hubs located in the insula, superior temporal pole, limbic regions, and anterior cingulate cortex (ACC) during pre-intervention resting state fMRI (rs-fMRI) was correlated with weight loss 18-months later (Paolini et al., 2015).

Despite the growing emphasis on complex brain networks in the neuroscience literature (Bullmore and Sporns, 2009), most weight loss studies have focused on specific brain regions (Stice et al., 2013; Volkow et al., 2011) rather than integrated networks. However, we would argue that multiple regions of the brain are involved in eating behavior, and that these regions interact in complex patterns with one another. Moreover, there has recently been growing interest in quantifying dynamic brain connections (Hutchison et al., 2013) rather than relying on measures of static connectivity, as brain networks intrinsically change over time due to varying physiological states (Barnes et al., 2009; Deshpande et al., 2006), even in a task-free environment (Chang and Glover, 2010). Thus, we believe it is vital that studies examine whole-brain dynamic networks (Allen et al., 2014; Hutchison et al., 2013), to understand how interacting brain areas contribute to obesity and weight loss.

In contrast to univariate statistical methods commonly used when examining region-based group differences in fMRI data, multivariate machine learning classification algorithms aim to discover a multivariate discriminatory pattern between groups. The machine learning algorithms are first trained with a subset of the data or with a specific training dataset. Once trained, these algorithms are used to predict an outcome or to classify test individuals (Mokhtari et al., 2016; Richiardi et al., 2013; Shahnazian et al., 2012). Multivariate pattern analysis is potentially capable of capturing complex interactions of distant brain sites that is not possible using pairwise connectivity analysis (Richiardi et al., 2013). This methodology allows for single-subject analyses where each participant can be classified rather than group-level summary outcomes such as brain maps of population differences.

The current study aimed to discover multivariate brain connectivity patterns in overweight/obese older adults following an overnight fast that might predict success with a behavioral weight loss intervention. We studied participants in a fasted state because food restraint is a contributing factor to failure with self-regulation of eating behavior (Fedoroff et al., 1997, 2003; Polivy et al., 2005), and variability in body weight (Heatherton et al., 1991). The ability to self-regulate caloric intake is central to intentional weight loss interventions. We believe that deficits in self-regulation are apparent in both automatic and control-based brain networks. We performed a prospective study in which baseline fMRI, baseline weight, and the weight following an 18-month lifestyle intervention were recorded. Here, we brought together dynamic brain networks and classification models for prospectively predicting the weight loss success of older adults who were overweight or obese, and had a documented history of cardiometabolic dysfunction. Given that multivariate prediction models can be used to perform treatment outcome predictions at the individual level, this study could help to provide critical groundwork to personalize weight loss interventions for patients based on individual brain networks.

Material and methods

Participants

Data collection for this study was completed in 2013. A sample of 66 individuals (ages: 60–79 years) were recruited as an ancillary study to the Cooperative Lifestyle Intervention Program-II (CLIP-II) trial (Marsh et al., 2013). Participants were required to be either overweight or obese (BMI ≥ 28 kg/m^2 but < 42 kg/m^2). Additional inclusion criteria involved having a documented history of either cardiovascular disease (CVD) or metabolic syndrome (MetS), low level of physical activity (less than 60 min of moderate physical activity per week), and mobility disability (having difficulty to complete daily physical tasks). All participants signed an informed consent/HIPAA authorization form. The study protocol was approved by the Wake Forest University School of Medicine institutional review board (IRB).

Major exclusion criteria included having a binge eating disorder, high alcoholic intake, a resting blood pressure >160/100 mmHg, a fasting blood glucose ≥ 140 mg/dL, diagnosis of type 1 or type 2 diabetes and insulin therapy, significant visual or hearing impairments, sever systemic diseases (for instance diagnosis of Parkinson's disease, and chronic liver disease), or abnormal health conditions which disable individuals to successfully complete the required study protocol.

Of 66 participants, 2 individuals did not meet eligibility to be randomized in the trial, 2 individuals withdrew from the study because of claustrophobia during magnetic resonance imaging (MRI), and 4 participants withdrew during the first 6-months of intervention for health reasons. Six additional individuals withdrew from the study after the 6-month intensive phase of intervention. Thus, a final sample size of 52 participants (female: 39, male: 13, mean age: 67.62) was used in this analysis. For greater details regarding the data used in this study, refer to (Marsh et al., 2013; Rejeski et al., 2017).

Lifestyle interventions

The CLIP II trial was designed with three different 18-month lifestyle weight loss interventions, including a diet only weight loss condition (WL), WL plus aerobic exercise training (WL+AT), or WL plus resistance exercise training (WL+RT). Participants were required to complete three 6-month phases during the intervention: 1) intensive (months 1–6), 2) transition (months 7–12), and 3) maintenance (months 13–18) phases. Participants were instructed to regulate their food intake, so that the rate of weekly weight loss was approximately 0.3 kg, and eventually after completing the intervention BMI to be decreased 7–10%. The goals of the transition phase varied between subjects based on intensive phase performance. During the maintenance phase, the major focus was on maintaining the weight loss that has been achieved. Participants assigned to the WL+AT group were required to walk on an indoor cushioned track at the YMCA. The intensity of walking was moderate, 4 days/week, with the goal toward 45 min/session. Participants were encouraged to walk at home as well. Participants assigned to the WL+RT intervention were required to perform strength training exercises 4 days/week. Participants completed 3 sets of 10–12 repetitions on 8 machines with initial resistance determined from 1 repetition maximum (RM) testing (Goal of 75% of 1RM). For greater details, refer to (Marsh et al., 2013; Rejeski et al., 2017).

MRI scanning protocol

MRI data was collected using a Siemens MAGNETOM SKYRA 3T MRI Scanner with Tim + Dot Technology. Participants were required to undergo an overnight fast, and were only allowed to consume water after midnight. The scans occurred in the early morning (between 8 a.m. and 11 a.m.) following the overnight fast. Each individual participated in a scanning session which included anatomical imaging, two rs-fMRI series, and a food visualization fMRI series between the two rs-fMRI, with each

series consisting of 157 images. During the rs-fMRI, participants were instructed to relax, not think of anything particular and view a cross sign on the rear projection screen. The food visualization task was designed as alternatively representing four of each participant's favorite food words on an MR-compatible rear projection screen; each word for 30 s. Participants were instructed to imagine the food with all five senses for the entire time that the food word was on the screen. For the current study, only the first resting state and the food visualization fMRI scans were examined. T1 structural images were acquired in the sagittal plane using a single-shot 3D MPRAGE GRAPPA2 sequence (TR = 2.3 s, TE = 2.9 ms, TI = 900 ms, flip angle = 9°, resolution = 1 × 1 × 1 mm). Functional BOLD imaging or T2*-weighted imaging were performed using a single-shot echo-planar imaging sequence (TR = 2.0 s, TE = 25ms, flip angle = 75, resolution = 3.5 × 3.5 × 5.0 mm).

Image processing

Statistical parametric mapping 12 (SPM12) software and advanced normalization tools (ANTS) were used to preprocess images. The SPM12 segmentation was used to parcellate the T1 images to gray matter, white matter, cerebrospinal fluid (CSF), bone, soft tissue, and air/background. High dimensional image warping was performed using ANTS symmetric diffeomorphic normalization (Avants et al., 2011), to obtain the deformation transformation map from the native space of each subject to Colin27 average brain template (Holmes et al., 1998).

The first 10 images of functional scans were discarded to assure fMRI signal equilibrium. This resulted in fMRI time series with 147 time points. Functional scans were slice-time corrected, and then realigned to the first volume. Functional scans of each study participant were co-registerd with the same participant's T1 structural data, to assure that they were aligned. The resulting fMRI time series were then band-pass filtered (0.009–0.08 Hz) to filter out physiological noise and low-frequency drifts (Biswal et al., 1995). Finally, confounding signals, including 6 rigid-body transformation parameters generated through the realignment process and 3 mean tissue signals (whole-brain, white matter, and CSF) were regressed out of the functional data. The mask of segmented tissues, created by SPM12 segmentation, was used to extract the mean tissue signals from fMRI data.

The inverse of the deformation field resulting from ANTS registration was applied to transform a functional brain atlas (Shen et al., 2013) to each subject's native brain space. Unlike anatomical brain atlases which may integrate functionally different areas into one region (Tzourio-Mazoyer et al., 2002), the functional atlas used in the present study parcellates brain into finer functionally distinct unites; 268 regions versus 90 regions of automated anatomical labeling (AAL) atlas (Tzourio-Mazoyer et al., 2002). Finally, the resulting atlas in the native space of each study participant was used to extract the mean time series of each region for that individual. It is noteworthy that networks were created and analyzed in native space to limit manipulation and interpolation of the fMRI time series.

Functional network creation

In this study, we used the sliding window technique to account for dynamic alterations associated with functional brain connectivity (Allen et al., 2014). For the sliding window method, a time window of fixed length W was selected and time points within that window were used to compute pairwise Pearson correlation coefficients. The window was then moved across the time and a matrix of correlation coefficients was computed for each shift. For each subject, connectivity data was then represented by a 3rd-order tensor $\mathfrak{C} \in \mathbb{R}^{N \times N \times T}$, where the first and second modes were connectivity and the third mode was time. Here, $N = 268$ was the number of regions of interest (ROIs), and $T = T' - W + 1$, with $T' = 147$, was the length of the preprocessed time series. For instance, the entry c_{ijt}, $i, j \, \varepsilon \, \{1, ..., N\}$, $t \in \{1, 2, ..., T\}$ of the tensor $\mathfrak{C}$ denoted the

functional connectivity between regions O_i and O_j at time window t. No self-connections were allowed.

We varied the window size from 61TRs to 101TRs in increments of 10TRs (where TR = 2 s), to capture dynamic changes of functional connectivity in different frequency ranges, and examine sensitivity of the prediction model to the window size alterations. In order to prevent adding the spurious fluctuations associated with intrinsic changes of fMRI time series to the dynamic connectivity computations, the window size should be greater than $1/f_{min}$ (Leonardi and Van De Ville, 2015), where f_{min} is the lowest frequency of the fMRI time series. As mentioned above, the time series were band-pass filtered at 0.009–0.08 Hz. Thus, in our study the minimum length of window size could be 56 time points (or equivalently 56TRs). To assure passing the lowest value for the window size, the lower limit for window size search range was set at 61 time points. Moreover, the odd length of the window allowed centered sliding window analysis, as it used a symmetric window at each time point. On the other hand, as the window size increases, the dynamic connectivity estimation approaches the static connectivity. Therefore, we limited the search range with the upper bound of 101 time points. To maintain the highest possible temporal resolution, the shift size was set at 1TR throughout the analysis.

To reduce the computational burden, as well as to generate sparse networks with the strong connections retained, the magnitude of correlation tensor was thresholded. It has been shown (Garrison et al., 2015) that proportional thresholdling strategies (e.g. consistent connectivity density) provide more stable network measures compared with absolute thresholding methods (e.g. statistical significance criterion indicated with p-value < 0.05). Thus, in this study, the connectivity tensors were thresholded such that a certain percentage of the high-strength connections was preserved for each subject at each time window. The sign of connections was preserved in the thresholded connectivity tensor. After thresholding, the resulting networks were checked for the fragmentation into multiple subnetworks, and no fragmented network was observed throughout our analysis. To examine the prediction performance consistency versus connectivity density threshold, a density range of 10–30% with increments of 5% was examined. Fig. 1 shows a schematic representation of dynamic connectivity tensor creation.

Connectivity tensor rank reduction using HOSVD

Each connection of the resulting brain network can be considered as a feature or equivalently an axis of the feature space that could contribute to distinguishing study populations. In the current study, the brain networks with 268 nodes have a potential of over 35,000 connections that could serve as features. To avoid including noisy contents which potentially results in low generalization power of the prediction model, the feature space dimension, or rank of the connectivity network data was reduced prior to being used in the multivariate model.

For the cases where datasets are represented with matrices (2nd-order tensors), one popular technique is to perform singular value decomposition (SVD) to obtain a summarized presentation of the data, mitigating the high dimensionality issue (Klema and Laub, 1980; Leonardi et al., 2013). As mentioned above, each feature is represented by an axis/dimension in the feature space; SVD identifies combined features (new axes) as a linear combination of the original features. The new features are called singular vectors or principal axes. The singular vectors are ordered regarding the amount of the data variance that they represent, with the first singular vector as the highest-ordered. Additionally, the variability along each singular vector is computed, namely singular values. The main idea of rank reduction analysis using SVD is to project the data along the R highest-ordered singular vectors such that a certain percent of data variance could be preserved.

In the current study, tensor-based higher-order SVD (HOSVD) (De Lathauwer et al., 2000) was used to reduce the dimensionality of connectivity tensors. HOSVD decomposes a tensor $\mathscr{A} \in \mathbb{R}^{d_1 \times ... \times d_p}$ to a core
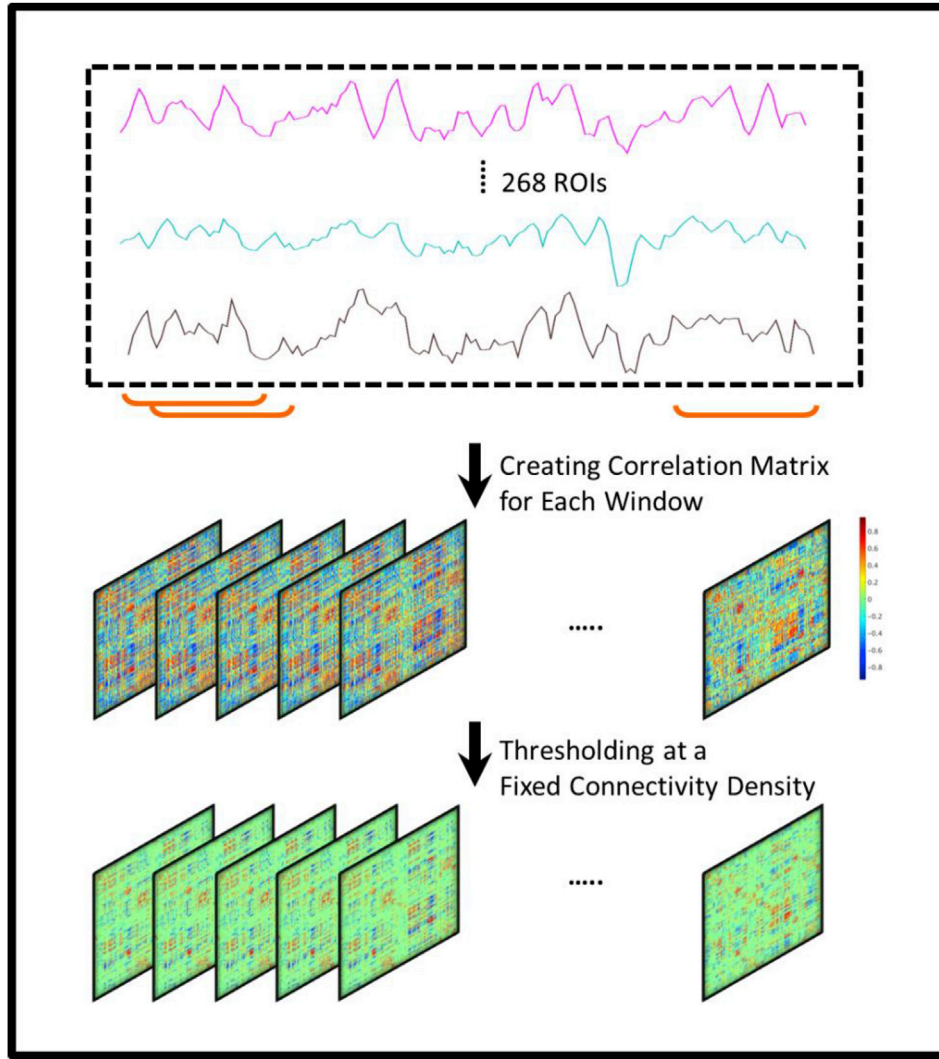
Fig. 1. Dynamic connectivity tensor creation procedure using sliding window technique, for which a window of fixed length is used to split the fMRI time series. For each split, a connectivity matrix is then constructed using pairwise Pearson correlation analysis. By concatenating the resulting matrices, a 3rd-order connectivity tensor is created for each participant. A thresholding value is chosen to remove weak connections, such that for each individual and at each time window a fixed percentage (e.g. 10%) of the total connections is maintained.

tensor $\mathscr{G} \in \mathbb{R}^{d_1 \times \dots \times d_p}$ consisting of singular values, and column spaces $\mathbf{U}^{(n)}$, $n \in \{1, 2, \dots, p\}$, each consisting of singular vectors in a specific mode. HOSVD preserves the information embedded in the structured tensors by computing singular vectors in each mode (connectivity, time and participants in this study) separately, in conjunction with their corresponding singular values. In the present study, the connectivity tensors, $\mathfrak{C} \in \mathbb{R}^{N \times N \times T}$, were concatenated across the M_{Trn} training samples to create a 4th-order connectivity tensor $\mathscr{C}_{Trn} \in \mathbb{R}^{N \times N \times T \times M_{Trn}}$. According to the HOSVD model, $\mathscr{C}_{Trn}$ can be decomposed as

$$\mathscr{C}_{Trn} = \mathscr{G} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times_3 \mathbf{U}^{(3)} \times_4 \mathbf{U}^{(4)} \tag{1}$$

where $\times_n$ denotes the n-mode tensor-matrix product. $\mathbf{U}^{(n)} = [\mathbf{u}_1^{(n)} \mathbf{u}_2^{(n)} \cdots \mathbf{u}_{d_n}^{(n)}] \in \mathbb{R}^{d_n \times d_n}$ is the singular vector matrix (or column space) in the n-th mode, where $d_1 = d_2 = N$, $d_3 = T$ and $d_4 = M_{Trn}$. The vector $\mathbf{u}_i^{(n)} \in \mathbb{R}^{d_n \times 1}$, $i \in \{1, 2, \dots, d_n\}$, is the i-th singular vector in the n-th mode. As the brain connectivity matrix for each subject at each time is symmetric, $\mathbf{U}^{(1)}$ and $\mathbf{U}^{(2)}$ are theoretically equal. $\mathscr{G} \in \mathbb{R}^{N \times N \times T \times M_{Trn}}$ is the core tensor with orthogonal sub-tensors, $\mathscr{G}_k^{(n)}$, and corresponding ordered Frobenius norms, $\sigma_k^{(n)} = \mathscr{G}_k^{(n)}$, where sub-tensors $\mathscr{G}_k^{(n)}$ are obtained by fixing the n-th index of $\mathscr{G}$ at $k \in \{1, 2, \dots d_n\}$, while the other indices are free. For instance, for a 3rd-order tensor, orthogonality means that different horizontal matrices of the tensor (where the first index is kept

fixed, while the two other indices are free) are mutually orthogonal. In other words, the inner product of every pair of the horizontal matrices is zero. Furthermore, the same rule applies to different frontal and vertical matrices as well. The same notion is generalized to nth-order tensors, $n > 3$, where orthogonal matrices are substituted with orthogonal sub-tensors. Details on the tensor algebra relevant to this project are provided in Appendix A.

As the norm of the sub-tensors $\sigma_k^{(n)} = \mathscr{G}_k^{(n)}$, $k \in \{1, 2, \dots, d_n\}$, are in descending order, one can take advantage of the truncated version of the core tensor to reduce the rank of the connectivity tensors (De Lathauwer et al., 2000; Rabusseau and Kadri, 2016). The optimal solution of the truncated decomposition of the tensor $\mathscr{C}_{Trn} \in \mathbb{R}^{N \times N \times T \times M_{Trn}}$ is obtained when the core tensor is in the hypercube form, $\mathscr{G}_r \in \mathbb{R}^{R \times R \times R \times R}$ where $R \leq \min\{N, N, T, M_{Trn}\}$ (Chen and Saad, 2009). It is shown that in such a decomposition the diagonal entries of the core tensor are maximized (Chen and Saad, 2009). In this study, R was first set at $M_{Trn} = 42$. However, since 85% of the total variance was captured by the first 21 singular vectors, ultimately rank $R = 21$ was chosen for the final analyses throughout the study. Please refer to Fig. B.1 (Appendix B), to see the supporting data regarding rank reduction.

According to the definitions provided in Appendix A, each entry c_{ijkl}, of the connectivity tensor, $\mathscr{C}_{Trn}$, is obtained as below

$$c_{ijkl} = \sum_{pqrs} u_{ip}^{(1)} u_{jq}^{(2)} u_{kr}^{(3)} u_{ls}^{(4)} g_{pqrs} = \sum_{pqrs} \left(u_{ip}^{(1)} u_{jq}^{(2)} \right) g_{pqrs} \left(u_{kr}^{(3)} u_{ls}^{(4)} \right) \tag{2}$$

where $u_{ij}^{(n)}$ indicates the entry (i, j) of $\mathbf{U}^{(n)}$, and g_{pqrs} is the entry (p, q, r, s) of the core tensor $\mathscr{G}$. Eq. (2) can be rewritten in the matrix multiplication form as below

$$\mathbf{C}_{Trn} = \left(\mathbf{U}^{(1)} \otimes \mathbf{U}^{(2)}\right)\mathbf{G}\left(\mathbf{U}^{(3)} \otimes \mathbf{U}^{(4)}\right)^T \in \mathbb{R}^{N^2 \times TM_{trn}} \tag{3}$$

where $\mathbf{C}_{Trn} \in \mathbb{R}^{N^2 \times TM_{Trn}}$ and $\mathbf{G} \in \mathbb{R}^{N^2 \times TM_{Trn}}$ are the reshaped format of the tensor $\mathscr{C}_{Trn} \in \mathbb{R}^{N \times N \times T \times M_{Trn}}$, and $\mathscr{G} \in \mathbb{R}^{N \times N \times T \times M_{Trn}}$, respectively. Using the truncated version of HOSVD, while only the R highest ranked singular vectors are kept in each mode, we can approximate the connectivity matrix, $\mathbf{C}_{Trn}$, as below

$$\tilde{\mathbf{C}}_{Trn} = \left(\mathbf{U}_r^{(1)} \otimes \mathbf{U}_r^{(2)}\right)\mathbf{G}_r\left(\mathbf{U}_r^{(3)} \otimes \mathbf{U}_r^{(4)}\right)^T \tag{4}$$

where $\tilde{\mathbf{C}}_{Trn} \in \mathbb{R}^{N^2 \times TM_{Trn}}$ is the reconstructed connectivity matrix using the R highest-ordered singular vectors, subscript r denotes reduced-rank, $\mathbf{G}_r \in \mathbb{R}^{R^2 \times R^2}$, $\mathbf{U}_r^{(1)} \in \mathbb{R}^{N \times R}$, $\mathbf{U}_r^{(2)} \in \mathbb{R}^{T \times R}$, $\mathbf{U}_r^{(3)} \in \mathbb{R}^{T \times R}$ and $\mathbf{U}_r^{(4)} \in \mathbb{R}^{M_{Trn} \times R}$. It should be noted that only the training samples were used in finding the singular vectors. Therefore, the projection of the training samples onto the reduced rank connectivity space, $\mathbf{U}_r^{(1)} \otimes \mathbf{U}_r^{(2)}$, can be represented by

$$\mathbf{C}_{rTrn} = \mathbf{G}_r\left(\mathbf{U}_r^{(3)} \otimes \mathbf{U}_r^{(4)}\right)^T \tag{5}$$

where $\mathbf{C}_{rTrn} \in \mathbb{R}^{R^2 \times TM_{Trn}}$ represents the reduced-rank connectivity networks of the training samples, which are reshaped to the tensor format as $\mathscr{C}_{rTrn} \in \mathbb{R}^{R \times R \times T \times M_{Trn}}$, prior to being used in the prediction model. Consequently, with the reduced-rank connectivity space represented by $\mathbf{U}_r^{(1)} \otimes \mathbf{U}_r^{(2)}$, and the reshaped version of the test samples connectivity tensors $\mathbf{C}_{Tst} \in \mathbb{R}^{N^2 \times TM_{Tst}}$, the dimensionality of the test samples can be reduced through

$$\mathbf{C}_{rTst} = \left(\mathbf{U}_r^{(1)} \otimes \mathbf{U}_r^{(2)}\right)^{-1}\mathbf{C}_{Tst} \tag{6}$$

where $\mathbf{C}_{rTst} \in \mathbb{R}^{R^2 \times TM_{Tst}}$ are the reduced-rank connectivity networks of these samples that are reshaped to the tensor format $\mathscr{C}_{rTst} \in \mathbb{R}^{R \times R \times T \times M_{Tst}}$ before testing the predictive model.

The resulting reduced-rank connectivity tensors of the training samples $\mathscr{C}_{rTrn} \in \mathbb{R}^{R \times R \times T \times M_{Trn}}$ and the test samples $\mathscr{C}_{rTst} \in \mathbb{R}^{R \times R \times T \times M_{Tst}}$ were finally averaged across time to obtain an $R \times R$ summarized representation of the connectivity tensor for each participant. Eventually, the dimension of the new feature vector representation was R^2.

It is also worth mentioning that by having principal vectors in the reduced-rank connectome space $\mathbf{U}_r^{(1)} = [\mathbf{u}_1^{(1)}, \mathbf{u}_2^{(1)}, ..., \mathbf{u}_R^{(1)}] \in \mathbb{R}^{N \times R}$, connectivity network components $FC_{ij} \in \mathbb{R}^{N \times N}$ associated with the singular vectors i and j, are obtained through the outer product of the corresponding vectors, i.e. $FC_{ij} = \mathbf{u}_i^{(1)}(\mathbf{u}_j^{(1)})^T$. Nevertheless, as previously mentioned, in the HOSVD estimation, the diagonal entries of the core tensor are maximized. Hence, the connectivity components, FC_{ij}, where $i = j$, should carry the highest level of information. In this study, only these components are presented.

This analysis was conducted independently for the resting state and food visualization task. The resulting low-rank connectivity networks were used in a linear SVM model (Cortes and Vapnik, 1995). LIBSVM toolbox was used to implement support vector classification (Chang and Lin, 2011).

Assessing prediction performance

The sample set was split to the low- and high-weight loss groups using the median of the weight change percentage [(weight after 18 months – baseline weight)/baseline weight $\times$ 100]. The performance of the models was evaluated using random sub-sampling cross validation approach, for which 52 samples were randomly permuted, the first 21 samples from each group (low- and high-weight loss groups) were assigned to the training set and the remaining to the test set (5 samples from each group). Thus, in each permutation, training and test sets consisted of 42 and 10 samples, respectively. It should be noted that before running the prediction model, 100 random permutations of the participants were recorded, and the same permutations of training and test participants were used for all the analyses of this study. Prediction performance measures were evaluated as the average of the resulting scores across the 100 permutations (or validation folds).

The grouping label of sample i was set as $[a_{i1}, a_{i2}]$, with $a_{i1} = 1$ and $a_{i2} = 0$ for the low weight loss group, and vice versa for the high weight loss group. The label predicted for the test sample i, $[a_{i1}', a_{i2}']$, was not a binary array. In other words, a_{i1}' and a_{i2}' indicate the probability of belongingness of the test sample i to the group 1 and 2, respectively. The label array $[a_{i1}', a_{i2}']$ was then binarized, to determine the grouping label of each sample. Equivalently, subject i was assigned to the low weight loss group, $[a_{i1}, a_{i2}] = [1, 0]$, if $a_{i1}' > a_{i2}'$, and to the high weight loss group, $[a_{i1}, a_{i2}] = [0, 1]$, if $a_{i1}' < a_{i2}'$. The probability of belongingness of each sample to either of the groups was estimated as suggested in (Chang and Lin, 2011).

We estimated cross entropy (CE) of the prediction model that quantified how close the predicted probabilities (prior to binarizing) were to the real binary labels. In other words, a high accuracy might not always imply a highly confident decision regarding grouping assignments, characterized by low CE. For instance, a sample with grouping probability [0.49, 0.51] would be assigned to the high weight loss group, but there would not be a high confidence margin in this prediction. CE was then defined as below

$$\text{CE} = -\frac{1}{M_{tst}} \sum_i^{M_{tst}} \sum_d^{D} \left(a_{id} \ln(a_{id}') + (1 - a_{id})\ln(1 - a_{id}')\right) \tag{7}$$

where D denotes the number of groups (2 in this study), M_{tst} indicates the number of test samples, and a_{id} and a_{id}' denote real and predicted probability of belongingness of the test sample i to the group d, respectively. For a perfect probability prediction CE is 0, and for a chance level probability, i.e. $a_{id}' = 0.5$, CE is $D \ln 2$.

For each validation fold, the accuracy, sensitivity and specificity of the model were determined using the test samples as below

$$\text{Accuracy} = \frac{\text{Number of correctly classified samples}}{\text{Total number of samples}} \times 100$$

$$\text{Sensitivity} = \frac{\text{Number of correctly classified samples from the low weight loss group}}{\text{Total number of low weight loss group samples}} \times 100 \tag{8}$$

$$\text{Specificity} = \frac{\text{Number of correctly classified samples from the high weight loss group}}{\text{Total number of high weight loss group samples}} \times 100$$

Results

Weight loss and hunger ratings following the overnight fast in the study sample

Within the entire sample, the mean (SD) baseline weight was 93.28 (12.80) kg, and the percent weight loss was 8.41 (7.14)%. The mean percent weight loss for those below the median was 2.87% (95% CI = 1.41, 4.33), while it was 13.96% (95%CI = 11.86, 16.05) for those above the median. These group differences reflect a poor and successful outcome for weight loss treatment, a conclusion that is in agreement with the clinically significant threshold of 5% established by The Obesity Society (TOS), since the loss of 5% or more has been associated with positive health consequences (Jensen et al., 2014). At the time of entering the scanner, the group was in the state of moderate hunger as evident by hunger ratings acquired using a 100 point VAS scale: Mean (95%CI) for hunger scale = 56.56 (48.30, 64.63). For greater details of the descriptive and clinical data, refer to Table B.1 (Appendix B) and the existing publications using this dataset (Mokhtari et al., 2016; Rejeski et al., 2017).

Prediction performance of the dynamic brain networks

The networks of the resting state and food cue task were used in separate prediction analyses. There was no significant difference in the prediction accuracy between resting-state and food-cue task; hence, the remainder of the manuscript focuses on resting-state data with detailed results for the food-cue task being presented in Fig. B.2.

Direct comparisons were made between the dynamic networks, static networks and dynamic networks with random grouping assignment. The dynamic connectivity networks were created using a sliding window with length of 61 time points and connectivity density of 10%. A single static connectivity matrix was also generated for each participant using the entire fMRI time series and thresholded at a density of 10%. For the networks with random grouping assignments, the labels of the 52 samples were randomly permuted and the resulting randomly assigned samples were used in the same rank reduction and prediction algorithms with the same 100 random subsampling cross validation folds. The entire random grouping assignment analysis was repeated 10 times and the average performance across the 10 runs was used. These networks were also created with a window size of 61 time points and a density of 10%. Fig. 2 contrasts the prediction performance measures from the different connectivity networks. The dynamic model exhibited the best performance with an average accuracy (SD) of 96.80% (5.67). The static networks resulted in a prediction accuracy of 48.90% (14.90). The prediction accuracy of the random grouping model was 49.40% (16.00) over the 10 runs; its correct prediction rate failed to exceed 70% in any of the validation folds.

Fig. 3 contrasts the prediction performance measures for the dynamic networks across different window sizes and different connectivity density thresholds. As shown on the right vertical axis, to explore how different density thresholds can affect prediction performance, a search range of 10–30% with 5% increments was examined. Similarly, prediction performance was evaluated with different window sizes ranging from 61 to 101 time points with 10 time points increments. To utilize information across a wider range of connectivity fluctuation frequencies, for each participant, a "Total" score was generated by concatenating the low-rank networks estimated for the same participant across all window sizes.

Results from a two-factor (window size and density) repeated measures analysis of variance (ANOVA) test with significance level of 0.05 showed that window size was a significant factor in the prediction performance measures including accuracy, specificity and cross entropy (p-value < 0.05). In addition, connectivity density was significantly associated with accuracy, specificity and cross entropy (p-value < 0.05). However, the interaction of window size and density was not
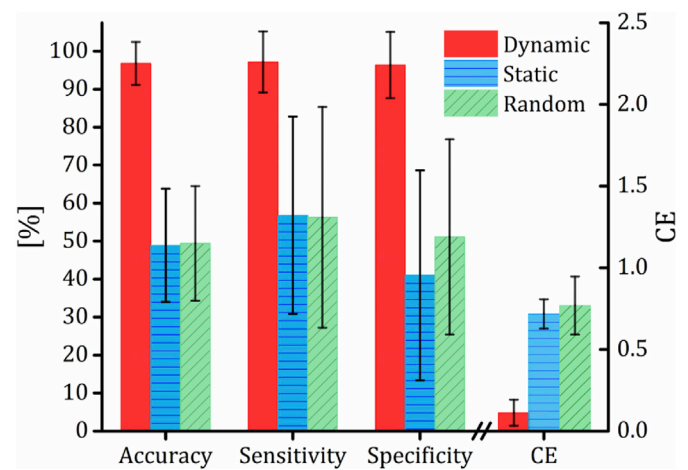

Fig. 2. Prediction performance, quanitified using accuracy (rate of correct prediction), sensitivity (rate of correct prediction for low-weight loss group), specificity (rate of correct prediction for high-weight loss group) and CE (a measure of prediction error defined as the difference between the prediction grouping probabilities and the real grouping labels, see section 2.7 for detailed definitions) for the dynamic and static connectivity and random grouping analysis. Dynamic connectivity tensors were constructed using a sliding window length of 61 time points and thresholded at connectivity density of 10%. No sliding window was used to make the static netowork; the entire fMRI time series was used to compute a single pairwise correlation matrix for the static networks. For the random grouping analysis, the particpants' labels were randomly permuted. HOSVD was used to reduce the networks' rank, HOSVD linearly decomposes the dynamic connectivity tensors to a set of linearly independent (equivalently orthogonal) dynamic connectivity components. To identify the components, HOSVD maximizes the amount of variance captured by the components.

significantly associated with the performance measures (p-value > 0.05). Further, multiple comparison tests were performed to identify which pairs of window sizes and density values contributed to the significant difference in the prediction performance (see Fig. B.3). As indicated in Fig. B.3 (a), the model integrating networks from different window sizes, namely "Total", significantly outperformed the models trained and tested using a single window size for the data. Additionally, a connectivity density of 10% resulted in significantly lower cross entropy (Fig. B.3 (b)).

Connectivity principal components of the dynamic connectivity networks

The 9 highest-ranked connectivity singular vectors captured ~81% of the variance of the reduced-rank data. Each singular vector identified by HOSVD represents a connectivity pattern, namely a connectivity principal component. Fig. 4 shows the principle components described by the 9 highest-ordered singular vectors, averaged over 100 validation folds. The amount of the variance carried by each component (equivalent to the corresponding singular value squared), normalized to the first singular vector variance, and averaged over 100 cross-validation folds, is shown on the top of that component. There were no qualitative differences between the network components generated using different window sizes. Due to space limitation and for simplicity, we only present the network components created with a window size of 61 time points and a density of 10%. For interested readers, the network components were provided in Appendix B as an Excel file, each sheet shows the connections (edges) of a specific component as a set of node pairs and corresponding weights. The number of each node is determined based on its label in the functional atlas (Shen et al., 2013). Only the connections located in the 99th percentile of the connection strength distribution are shown for each component.

Given that these components represent the complex patterns that emerge from temporally high level interactions between single
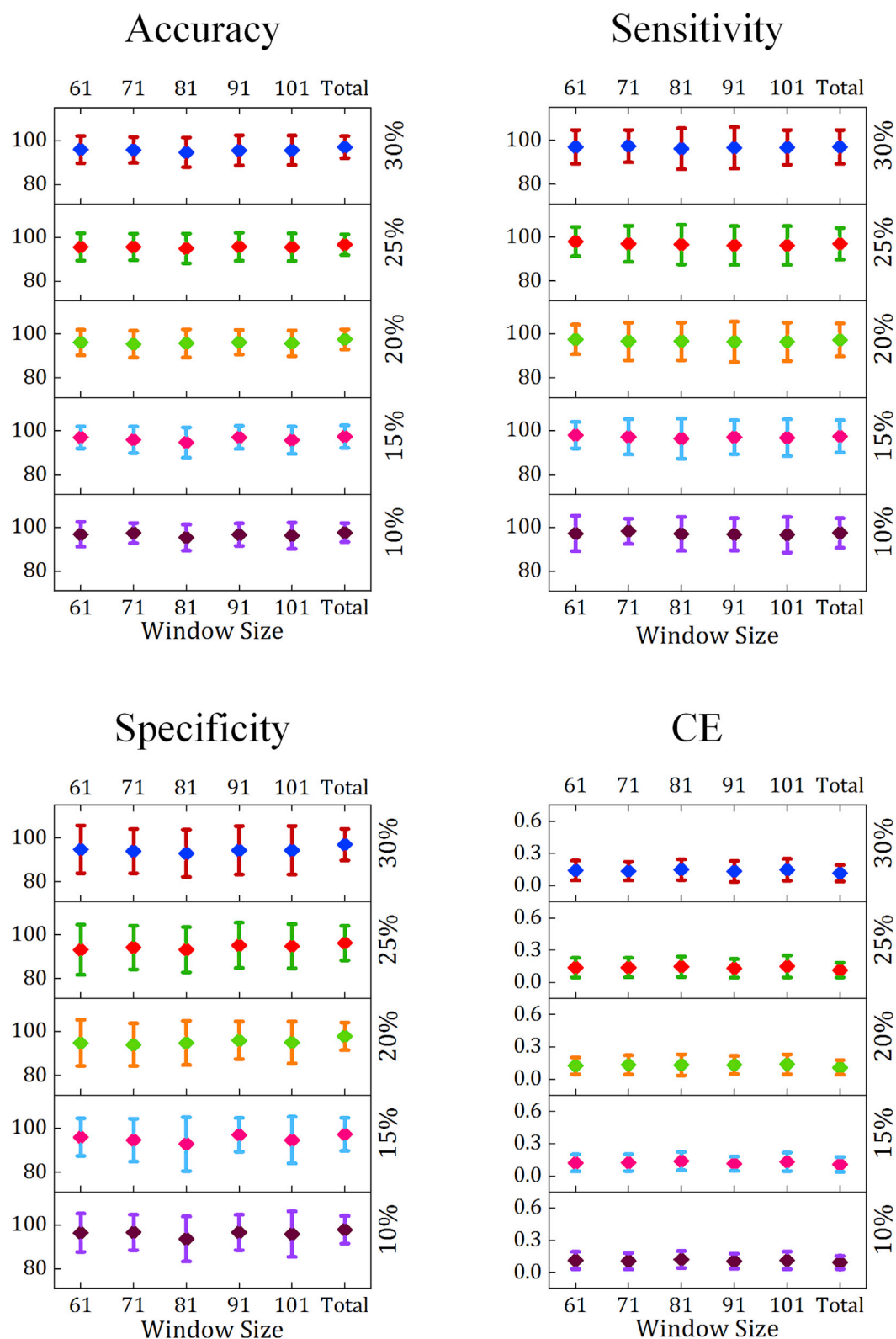
Fig. 3. Prediction performance measures [mean with error bars reflecting ±SD across 100 validation folds] of the SVM model estimated for different window sizes and different connectivity density thresholds. The numbers on the left vertical axis of each figure illustrate the prediction performance values, and the numbers on the right side indicate connectivity density thresholding values.

connections, they should be viewed as whole connectivity components, rather than sets of independent connections between regions. Nevertheless, it is noteworthy that the network components contain brain connections that are hypothesized to be involved in many aspects of

eating behavior and behavioral change skills necessary for successful weight loss (Rothemund et al., 2007; Stice et al., 2013; Volkow et al., 2008). The components were made up of the following interconnected regions:
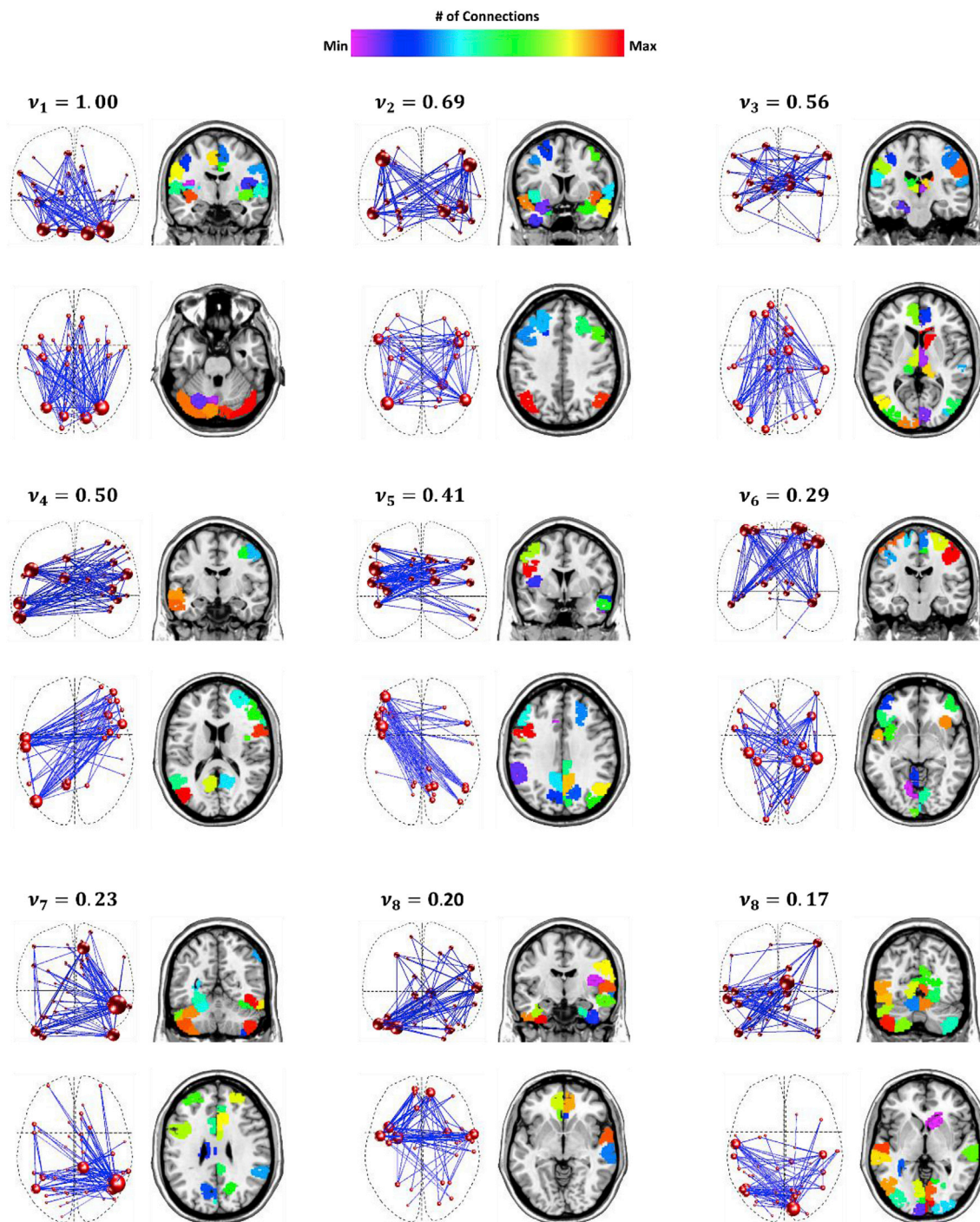
Fig. 4. The connectivity principal components revealed by HOSVD. For each component, the left column displays a network-based representation and the right column shows the individual nodes in their anatomical locations. Each network component is a collection of nodes and edges that captures the greatest amount of variance across time and across individuals. The size of each node is directly related to its number of connections. The variance explained by each component, ν_i, where i indicates the component index, is noted on the top of the same component's graph. For each component, the variance was computed as the square of the corresponding singular value normalized by the first singular value, that was finally averaged across 100 validation folds.

1. Dorsal insula, basal ganglia, anterior cingulate, and cerebellum
2. Attention network and limbic regions such as amygdala, hippocampus and inferior insula
3. Right caudate, visual cortex, motor and sensory regions, thalamus, medial prefrontal and dorsolateral prefrontal cortices
4. Left posterior cingulate/precuneus, left temporal pole and right ventrolateral prefrontal cortex
5. Right posterior cingulate/precuneus, right temporal pole and left ventrolateral prefrontal cortex
6. Bilateral anterior insula, bilateral inferior frontal, and early visual and primary motor cortices
7. Cingulate gyrus, lateral visual cortex and frontal poles
8. Orbitofrontal cortex, temporal poles and parietal regions
9. Visual cortex and cerebellum

As mentioned above, to provide a clearer visualization of the network components, only connections located in the 99th percentile of the connection strength distribution are shown for each component. Thus,
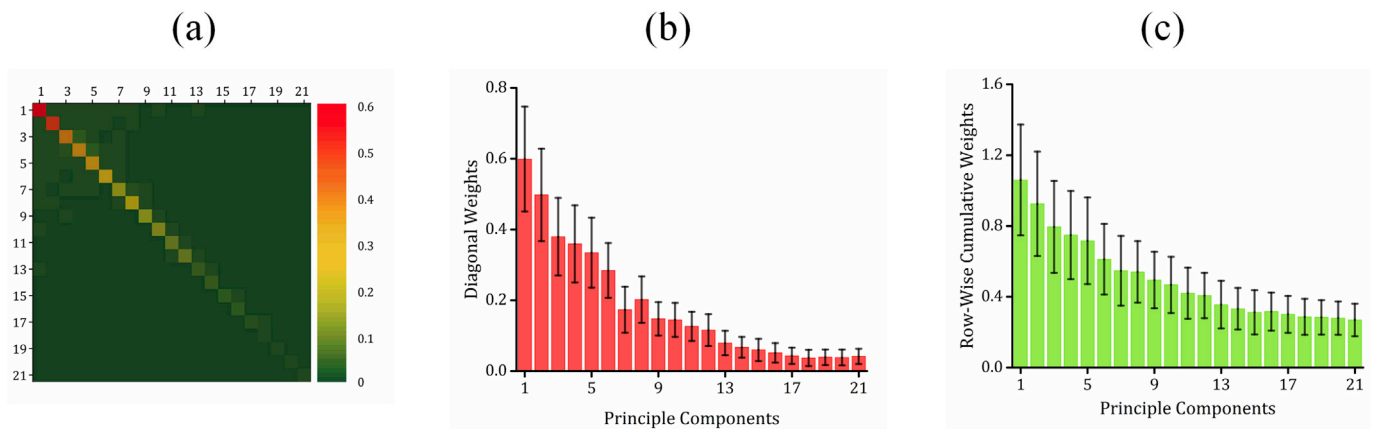
(a)   (b)   (c)

Fig. 5. (a) Strength of SVM weight matrix, $|W| \in \mathbb{R}^{21 \times 21}$, averaged over 100 validation folds, (b) average weights on the diagonal entries of the SVM weight matrix $|W| \in \mathbb{R}^{21 \times 21}$, (c) average of row-wise sum of the matrix $|W| \in \mathbb{R}^{21 \times 21}$.

the images in Fig. 4 were selected to highlight the key regions that might contribute to the classification of successful weight loss. Note that the connections shown in this figure do not represent simple temporal correlation between the regions. Rather, connections between two regions indicate a complex coupling between the connections of those two regions across time and across individual participants.

The SVM model exploits an optimum parametrized linear combination of the features to build the maximum-margin hyperplane between the two groups (Cortes and Vapnik, 1995). A parameter is assigned to each feature that represents the weight of the same feature in the linear maximum-margin separation. Thus, for the reduced-rank connectivity data $\in \mathbb{R}^{21 \times 21}$, the parameters, also known as the weights, can be represented with the matrix $W \in \mathbb{R}^{21 \times 21}$. The strength of parameter matrix $|W| \in \mathbb{R}^{21 \times 21}$, averaged over 100 permutations, is illustrated in Fig. 5(a). The diagonal entry, w_{ii} $i \in \{1, 2, \ldots, 21\}$, represents the weight of the principal component i. The off-diagonal entry, w_{ij} $i, j \in \{1, 2, \ldots, 21\}$, $i \neq j$, represents the weight assigned to the interaction of principal components i and j. As apparent in the diagonal entries, the weights, as a measure of principal components discriminatory power, decrease as the order of the components decreases. Moreover, it shows that overall the diagonal elements are substantially stronger than the off-diagonal entries in the corresponding row and column. In other words, the individual principal networks, play a significant role in discriminating the groups whereas the interaction between principal components contributes much less. To clarify this point, Fig. 5(b) and (c) depict the average and standard deviation of the diagonal entries and the cumulative weight associated with each principal component (equivalently row-wise sum of the matrix $|W|$).

Discussion

This study involved an examination of whether functional brain networks of older, obese adults assessed in a fasted state prior to randomization to intentional weight loss could be used to predict success with weight loss 18 months after initiation of treatment. Evaluating functional brain networks following an overnight fast was a critical feature of the study design in that periods of food restraint result in nutritional deficits that trigger hunger and hence food consumption (Fedoroff et al., 1997, 2003; Polivy et al., 2005). In line with Kavanagh and colleagues' (Kavanagh et al., 2005) Elaborated Intrusion Theory of Desire, the motivation to eat is driven by both unconscious and conscious processes. Moreover, physiological deficits, created by manipulations such as food restraint, trigger cravings for food and challenge self-regulatory capacities (Fedoroff et al., 1997, 2003; Polivy et al., 2005). Indeed, we have observed this pattern of response in our prior work with food restraint (Rejeski et al., 2010). Consistent with our study

hypothesis, we found that HOSVD in combination with machine learning applied to pre-treatment functional brain networks generated from resting-state scans resulted in network components that were highly accurate (>95%) in predicting 18-month weight loss.

There is growing consensus that self-regulatory skills and the motivation required to follow-through with behavior change (Rejeski et al., 2008; Wing, 2002) are vital for successful weight loss or adopting any new health behavior. Whereas weight loss research in neuroscience has focused on specific regions of interest (Stice et al., 2013), the current methodology is highly innovative in that we used HOSVD in conjunction with machine learning to identify patterns of multivariate connectivity across the whole brain that effectively discriminated between individuals who either exceeded or failed to exceed the 5% weight loss criterion that defines success with intentional weight loss. Furthermore, the network components used in the predictive model have at least three important implications. First, they could assist in the development of a comprehensive neural-based theory of intentional weight loss, since studies involving specific regions of interest are restrictive and ignore the inherent complexity of the brain. Indeed, recent comprehensive reviews of research into obesity suggest that a whole brain approach is warranted (Stice et al., 2013). Second, a broader view of dynamic brain networks could lead to the development of new treatment options. These options could complement or improve upon current behavioral methods for weight management which largely rely on conscious, decision based theories (Wing, 2002). And third, we believe that a more comprehensive understanding of neural networks involved in successful weight management could lead to the tailoring of treatment consistent with precision-based models of medicine. It is unlikely that "one size fits all" when it comes to implementing strategies for promoting health behavior change. Because the analytic approach in the current investigation relied on multivariate relationships across the whole brain, prior to discussing key brain networks illuminated by these analyses, we first want to emphasize what the phenotypic characterizations can and cannot tell us.

As noted above, the predictive analyses combined HOSVD with machine learning to identify patterns of multivariate connectivity that were effective in discriminating between individuals who either exceeded or failed to exceed the 5% criterion for successful weight loss. The multivariate patterns indicate that, as opposed to specific regions of interest, complex connectivity patterns between widespread areas of the brain are key to understanding intentional weight loss. This finding is vital when attempting to gain a mechanistic understanding of the network components that contribute to successful weight loss. The network components cannot be dissected and interpreted as individual brain regions or individual connections between brain regions. It is the entire connectivity pattern of the network components that was essential to effective classification.

The results clearly demonstrated that dynamic network analyses were successful at predicting weight loss whereas static network analyses were not. This means that a series of networks that evolved over time was necessary to discriminate between the weight loss groups. Each component (see Fig. 4) represents a brain subnetwork that accounts for variability in connectivity across the groups. The expression of each component varies over time (see Fig. B.4) in each individual participant. The machine learning algorithm was able to identify spatio-temporal patterns that distinguished the high from low weight loss groups. This does not simply translate to a given component being more or less likely to be connected in the high or low weight loss groups. Rather, the temporal profile of the network components and the interaction between the temporal dynamics of the components distinguished the study populations. Due to the fact that the multivariate pattern spans many interacting spatial and temporal dimensions, it is not possible to simply identify a single discriminant feature for the two groups. The complex pattern is what discriminated between the two groups and it required a machine learning algorithm to identify that pattern in high dimensional space.

With these caveats in mind, 9 principal components captured in excess of 80% of dynamic connectivity variance. Consistent with Kavanagh and colleagues' (Kavanagh et al., 2005) Elaborated Intrusion Theory of Desire, these components underscore that effective self-regulation involves both unconscious processes—sensory, motor, cognitive and affective—that are likely responsible for the intrusive quality of food craving as well as conscious processes—both cognitive and emotional—that serve to elaborate on these intrusive cues to promote consumption. Although we want to reemphasize the importance of interpreting each network as a whole component, and not as distinct regions of interests, the implication of regional interactions in each component is briefly discussed below.

Component 1 represents a network that is dominated by interactions between the cerebellum, lateral sensorimotor areas (including face, mouth, and throat), posterior insula, mid-anterior cingulate cortex, as well as the early visual cortex. Given this connectivity pattern, the component 1 may play an active role in higher-order cognitive processing, e.g. self-regulation and goal-directed behaviors for optimizing context-appropriate performance (Imamizu and Kawato, 2009; Schmahmann, 1991, 1996; Stoodley, 2012; Stoodley et al., 2012), feeding behavior (Zhu and Wang, 2008; Zhu et al., 2006), and voluntary movement commands (Daskalakis et al., 2004; Guye et al., 2003).

A bilateral interacting pattern between the executive attention network and hedonic/goal directed network including the amygdala, hippocampus and inferior insula was found to serve as the core for component 2. We postulate that this network component represents the process of top-down control that the attention network projects onto the limbic regions known to be important in goal-oriented behavior (Dosenbach et al., 2008; Hopfinger et al., 2000). Altered connectivity within this network component could adversely affect inhibitory control processing and consequently lead to altered behavioral regulation (Volkow et al., 2008). Interestingly connections between the control and reward/motivation networks are known to play a central role in regulating food consumption (Brooks et al., 2012; Volkow et al., 2011).

Component 3 represents a sensory network that involves extensive connectivity with the visual cortex, thalamus, caudate, ACC, medial prefrontal cortex (MPFC) and dorsal lateral prefrontal cortex (DLPFC). While the ACC, MPFC and DLPFC are known to be associated with salience and the rewarding qualities of sensory input (Ridderinkhof et al., 2004; Rogers et al., 2004), the DLPFC also plays a key role in working memory.

The importance and complexity of elaboration and inhibition to intentional weight loss is evident in network components 4–8. The presence of an interacting pattern consisting of the left temporal pole, right ventrolateral prefrontal (RVLPF) and posterior cingulate cortices defines component 4 as an important network in cognitive elaboration/inhibition (Levy and Wagner, 2011; Noesselt et al., 2007). Interestingly,

the structure of network component 4 was also present in network component 5, albeit with opposite laterality. Network component 5 includes emotional elaboration/delayed discounting network comprising the interactions between right temporal pole and left ventrolateral prefrontal cortex (LVLPFC) (McClure et al., 2004).

Network component 6 represents communication between the bilateral anterior insula, bilateral inferior frontal areas, early visual areas, and superior and medial motor cortex. This network component can be considered influential in self-regulatory processes involving emotional awareness and inhibition (Rueda et al., 2004), whereas network component 7 captures interactions between inferior temporal areas and the posterior cingulate cortex, key elements of the resting state brain default mode network (DMN) (Bluhm et al., 2007; Uddin et al., 2009).

The 8th network component includes high-level interactions between temporal poles, anterior cingulate, and orbital frontal cortices. This component resembles the hot state brain network of appetite (HBN-A) -including insula, anterior cingulate cortex (ACC), superior temporal pole (STP), amygdala and the parahippocampal gyrus- (Paolini et al., 2014, 2015) that was found to be an important control-based network contributing to the prediction of weight loss in prior work (Paolini et al., 2015).

Finally, network component 9 is dominated by connections between the visual cortex and the cerebellum. In general, we posit that the appearance of the interactions between visual cortex and cerebellum in this dedicated network component, underscores the powerful role that visual imagery and embodied long term memory play in the desire for food and hence the strength of self-regulatory capacities. In fact, the role of visual imagery in food craving is a topic that has attracted considerable attention (Bullins et al., 2013; Harvey et al., 2005; Kemps and Tiggemann, 2010; Kemps et al., 2008; Tiggemann and Kemps, 2005). Not surprisingly, visual imagery and other body related/sensory inputs are core features of Kavanagh and colleagues' Elaborated Intrusion Theory of Desire (Kavanagh et al., 2005).

Despite the strengths of the current method in classifying intervention success and the role that this model played in elucidating the complex, whole brain connectivity involved in intentional weight loss, the study is not without limitations. First, the sample size was small and involved a relatively homogeneous group of overweight/obese older adults with metabolic syndrome and/or cardiovascular disease (CVD). We are aware that the extraordinary prediction accuracy in this study could be artificially inflated by these factors. There is a need to replicate these findings with larger independent samples. Given the complexity of the data and technique, we expect that the prediction performance will fall off some, when using independent samples. Nevertheless, in real-life applications, even a prediction performance of ~60% could significantly impact medical decisions. Second, although sex has been identified as a determinant of brain function and self-regulation (Cosgrove et al., 2007; Kimura, 1992; Rolls et al., 1991), we did not find that sex was an important consideration in the success of classifying who was likely to fail or succeed at weight loss. It is possible that in a larger more heterogeneous sample of men and women, classification accuracy could be compromised if sex was ignored. The same could be said for age, race and other potentially important moderating variables that do compromise the external validity of these results. Third, previous studies have shown that overeating and consequently obesity can occur in the absence of hunger (Faith et al., 2006; Fisher and Birch, 2002; Hill et al., 1991). In the current study, baseline fMRI scans were acquired in a physiological state of hunger following an overnight fast, and the successful weight loss performance was prospectively predicted using the resulting dynamic connectivity networks. However, we did not compare fasted and sated treatment conditions. This topic is certainly worthy of study in subsequent research.

Finally, there are at least two other important areas of study that deserve attention. First, if one prescribes an intervention based on the multivariate connectivity patterns, will this lead to an improved outcome? In other words, although this study shows that there are well-

defined network components that are useful in predicting weight loss success, the current study does not tell us whether these components can be modified in a manner that changes the outcome for those who failed. In addition to this limitation in interpreting the component networks, one should also keep in mind that the interaction of individual characteristics and intervention type could make the interpretation of network components even more challenging. A second and related issue is that, although this is the first step toward specifying weight loss treatment at individual level, separate models should be generated for each specific intervention of conceptual interest to demonstrate its efficacy, eventually developing a clinical paradigm in which individuals could be matched to highly specific behavioral and/or pharmacological interventions. In other words, if we had sufficient sample sizes (e.g. 200 participants) for different interventions, we could construct a separate model for each specific intervention to identify phenotypes that are predictive of success. The probability of successful weight loss for each future patient would then be estimated with all the resulting models based on their brain phenotypes. The intervention associated with the model resulting in the highest probability would then be assigned to that patient. In this way, we can extend the proposed idea in this paper to a personalized medicine paradigm.

Conclusions

In summary, the current study offers an exciting new frontier for behavioral neuroscience in that it demonstrates the utility of studying connectivity across the whole brain as opposed to focusing on limited regions of interest. Although the study sample was relatively homogeneous, obesity is a major factor in the high rates of metabolic syndrome and/or CVD as people age and creates an enormous burden on public health. The combined use of HOSVD with machine learning was highly effective at using baseline dynamic fMRI networks to identity participants who were likely to either succeed or fail at intentional weight loss 18-months after the initiation of treatment. This is a novel and highly innovative approach to identifying phenotypes that are best suited for current behavioral weight loss interventions and offers guidance for tailoring treatment, a research priority at the National Institutes of Health. For example, it may well be that individuals who most closely resemble the phenotype of those who were successful with our intervention, could be effectively treated with less intensive behavioral therapy. This has important implications for public health. Moreover, we believe the 9 components that accounted for over 80% of the variance of the data will serve as an impetus for theory development in future exploratory studies that aim to investigate baseline brain functional characteristics as predictive phenotypes of obesity treatment success.

Conflicts of interest

None.

Appendix A. Tensor Algebra Details

A tensor $\mathscr{A} \in \mathbb{R}^{d_1 \times \ldots \times d_p}$ can be represented by a multidimensional array $(\mathscr{A}_{i_1, \ldots, i_p} : i_n \in \{1, 2, \ldots, d_n\}, n \in \{1, \ldots, p\})$. The n-mode fibers of the $\mathscr{A}$ are the vectors obtained by letting the n-th index to change while the remaining indices are fixed, e. g. $\mathscr{A}_{:, i_2, \ldots, i_p} \in \mathbb{R}^{d_1 \times 1}$ represents the 1-mode fiber of $\mathscr{A}$. The n-mode matrix of $\mathscr{A}$, $A_{(n)} \in \mathbb{R}^{d_n \times d_1 d_2 \ldots d_{n-1} d_{n+1} \ldots d_p}$ is constructed by stacking n-mode fibers of $\mathscr{A}$ as columns. The vectorization form of the tensor $\mathscr{A} \in \mathbb{R}^{d_1 \times \ldots \times d_p}$ is defined as $\text{vec}(\mathscr{A}) = \text{vec}(A_{(1)})$. The inner product between two same-size tensors $\mathscr{A}$ and $\mathscr{B}$ is defined by $\langle \mathscr{A}, \mathscr{B} \rangle = \langle \text{vec}(\mathscr{A}), \text{vec}(\mathscr{B}) \rangle$, and the Ferobinus norm of $\mathscr{A}$ is defined as $\mathscr{A}_F^2 = \langle \mathscr{A}, \mathscr{A} \rangle$.

The n-mode matrix product of the tensor $\mathscr{A}$ with a matrix $X \in \mathbb{R}^{m \times d_n}$, shown with the symbol $\times_n$, is defined by the relationship below

$$\mathscr{Y} = \mathscr{A} \times_n X \Leftrightarrow Y_{(n)} = X A_{(n)} \tag{A.1}$$

where the tensor $\mathscr{Y} \in \mathbb{R}^{d_1 \times d_2 \times \ldots \times d_{n-1} \times m \times d_{n+1} \times \ldots \times d_p}$. The n-mode rank of $\mathscr{A}$ is defined as the rank of n-mode matricization of the tensor denoted by $A_{(n)}$, $\text{rank}_n(\mathscr{A}) = \text{rank}(A_{(n)})$, that represents the dimension of the space spanned by the n-mode fibers of the tensor. The multilinear rank of the tensor $\mathscr{A}$, denoted by $\text{rank}(\mathscr{A})$, is the tuple of the n-mode ranks of $\mathscr{A}$, denoted by $\text{rank}(\mathscr{A}) = (R_1, \ldots, R_p)$ where $R_n = \text{rank}_n(\mathscr{A})$, $n \in \{1, 2, \ldots p\}$. It is written that $\text{rank}(\mathscr{A}) \le (S_1, \ldots, S_p)$, whenever $\text{rank}_1(\mathscr{A}) \le S_1$, $\text{rank}_1(\mathscr{A}) \le S_2$, ..., $\text{rank}_p(\mathscr{A}) \le S_p$.

Having these basic definitions established, according to the higher order singular value decomposition (HOSVD), the tensor $\mathscr{A}$ can be decomposed as below

$$\mathscr{A} = \mathscr{G} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times_3 \mathbf{U}^{(3)} \times \ldots \times_p \mathbf{U}^{(p)} \tag{A.2}$$

where $\mathscr{G} \in \mathbb{R}^{d_1 \times \ldots \times d_p}$ is the core tensor with orthogonal subtensors, $\mathbf{G}_k^{(n)}$, and corresponding ordered Frobenius norms $\sigma_k^{(n)} = \mathbf{G}_k^{(n)}$. The subtensor $\mathbf{G}_k^{(n)}$ is obtained by fixing the n-th index of $\mathscr{G}$ at $k \in \{1, 2, \ldots d_n\}$, $n \in \{1, 2, \ldots, p\}$, while the other indices are free. $\mathbf{U}^{(n)} = [\mathbf{u}_1^{(n)} \mathbf{u}_2^{(n)} \cdots \mathbf{u}_{d_n}^{(n)}] \in \mathbb{R}^{d_n \times R_n}$ is a unitary matrix, $\mathbf{U}_i^T \mathbf{U}_i = \mathbf{I}$, and represents the singular vectors in the n-th mode.

According to the tensor rank definition provided above, the number of parameters needed to fully represent the tensor $\mathscr{A}$ using HOSVD should be equal or smaller than $d_1 d_2 \ldots d_p$. Finding the optimum decomposition of the tensor $\mathscr{A}$, with the multilinear rank $(R_1, R_2, \ldots R_p) \le (d_1, d_2, \ldots, d_p)$ is a very hard and time consuming problem. In practice, truncated HOSVD algorithm provides a good approximation of the multilinear tensor decomposition (Rabusseau and Kadri, 2016). For the truncated HOSVD, only the l_n singular vectors with the largest singular values (λ_k, $k \in \{0, 1, \ldots d_n\}$) within n-th mode are chosen to reconstruct the data, that results in a retained variance percent of $\sum_{k=1}^{l_n} \lambda_k^2 / \sum_{k=1}^{d_n} \lambda_k^2 \times 100$ in that mode. Furthermore, it has also been proved (Chen and Saad, 2009) that there exists an optimal solution for the tensor-based low rank decomposition problem when $R_1 = R_2 = \ldots = R_p \le min(d_1, d_2, \ldots, d_p)$.

Appendix B. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.neuroimage.2018.02.025.

References

Allen, E.A., Damaraju, E., Plis, S.M., Erhardt, E.B., Eichele, T., Calhoun, V.D., 2014. Tracking whole-brain connectivity dynamics in the resting state. Cerebr. Cortex 24, 663–676.

Avants, B.B., Tustison, N.J., Song, G., Cook, P.A., Klein, A., Gee, J.C., 2011. A reproducible evaluation of ANTs similarity metric performance in brain image registration. Neuroimage 54, 2033–2044.

Barnes, A., Bullmore, E.T., Suckling, J., 2009. Endogenous human brain dynamics recover slowly following cognitive effort. PLoS One 4, e6626.

Berridge, K.C., Ho, C.-Y., Richard, J.M., DiFeliceantonio, A.G., 2010. The tempted brain eats: pleasure and desire circuits in obesity and eating disorders. Brain Res. 1350, 43–64.

Biswal, B., Zerrin Yetkin, F., Haughton, V.M., Hyde, J.S., 1995. Functional connectivity in the motor cortex of resting human brain using echo-planar mri. Magn. Reson. Med. 34, 537–541.

Bluhm, R.L., Miller, J., Lanius, R.A., Osuch, E.A., Boksman, K., Neufeld, R., Théberge, J., Schaefer, B., Williamson, P., 2007. Spontaneous low-frequency fluctuations in the BOLD signal in schizophrenic patients: anomalies in the default network. Schizophr. Bull. 33, 1004–1012.

Brooks, S.J., Rask-Andersen, M., Benedict, C., Schiöth, H.B., 2012. A debate on current eating disorder diagnoses in light of neurobiological findings: is it time for a spectrum model? BMC Psychiatr. 12, 76.

Bullins, J., Laurienti, P.J., Morgan, A.R., Norris, J., Paolini, B.M., Rejeski, W.J., 2013. Drive for consumption, craving, and connectivity in the visual cortex during the imagery of desired food. Frontiers in Aging Neuroscience 5. Front. Aging Neurosci. 5.

Bullmore, E., Sporns, O., 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. Nat. Rev. Neurosci. 10, 186–198.

Chang, C.-C., Lin, C.-J., 2011. LIBSVM: a library for support vector machines. ACM Trans. Intell Syst Technol (TIST) 2, 27.

Chang, C., Glover, G.H., 2010. Time–frequency dynamics of resting-state brain connectivity measured with fMRI. Neuroimage 50, 81–98.

Chen, J., Saad, Y., 2009. On the tensor SVD and the optimal low rank orthogonal approximation of tensors. SIAM J. Matrix Anal. Appl. 30, 1709–1734.

Cortes, C., Vapnik, V., 1995. Support-vector networks. Mach. Learn. 20, 273–297.

Cosgrove, K.P., Mazure, C.M., Staley, J.K., 2007. Evolving knowledge of sex differences in brain structure, function, and chemistry. Biol. Psychiatr. 62, 847–855.

Curioni, C.C., Lourenco, P.M., 2005. 2005. Long-term weight loss after diet and exercise: a systematic review. Int. J. Obes. 29, 1168–1174.

Daskalakis, Z.J., Paradiso, G.O., Christensen, B.K., Fitzgerald, P.B., Gunraj, C., Chen, R., 2004. Exploring the connectivity between the cerebellum and motor cortex in humans. J. Physiol. 557, 689–700.

De Lathauwer, L., De Moor, B., Vandewalle, J., 2000. A multilinear singular value decomposition. SIAM J. Matrix Anal. Appl. 21, 1253–1278.

Deshpande, G., LaConte, S., Peltier, S., Hu, X., 2006. Directed transfer function analysis of fMRI data to investigate network dynamics. Engineering in Medicine and Biology Society, 2006. In: EMBS'06. 28th Annual International Conference of the IEEE. IEEE, pp. 671–674.

Dosenbach, N.U., Fair, D.A., Cohen, A.L., Schlaggar, B.L., Petersen, S.E., 2008. A dual-networks architecture of top-down control. Trends Cognit. Sci. 12, 99–105.

Elfhag, K., Rössner, S., 2005. Who succeeds in maintaining weight loss? A conceptual review of factors associated with weight loss maintenance and weight regain. Obes. Rev. 6, 67–85.

Faith, M.S., Berkowitz, R.I., Stallings, V.A., Kerns, J., Storey, M., Stunkard, A.J., 2006. Eating in the absence of hunger: a genetic marker for childhood obesity in prepubertal boys? Obesity 14, 131–138.

Fedoroff, I., Polivy, J., Herman, C.P., 2003. The specificity of restrained versus unrestrained eaters' responses to food cues: general desire to eat, or craving for the cued food? Appetite 41, 7–13.

Fedoroff, I.D., Polivy, J., Herman, C.P., 1997. The effect of pre-exposure to food cues on the eating behavior of restrained and unrestrained eaters. Appetite 28, 33–47.

Fisher, J.O., Birch, L.L., 2002. Eating in the absence of hunger and overweight in girls from 5 to 7 y of age. Am. J. Clin. Nutr. 76, 226–231.

Flegal, K.M., Carroll, M.D., Kit, B.K., Ogden, C.L., 2012. Prevalence of obesity and trends in the distribution of body mass index among us adults, 1999-2010. JAMA, J. Am. Med. Assoc. 307, 491–497.

Garrison, K.A., Scheinost, D., Finn, E.S., Shen, X., Constable, R.T., 2015. The (in) stability of functional brain network measures across thresholds. Neuroimage 118, 651–661.

Guye, M., Parker, G.J., Symms, M., Boulby, P., Wheeler-Kingshott, C.A., Salek-Haddadi, A., Barker, G.J., Duncan, J.S., 2003. Combined functional MRI and tractography to demonstrate the connectivity of the human primary motor cortex in vivo. Neuroimage 19, 1349–1360.

Harvey, K., Kemps, E., Tiggemann, M., 2005. The nature of imagery processes underlying food cravings. Br. J. Health Psychol. 10, 49–56.

Heatherton, T.F., Polivy, J., Herman, C.P., 1991. Restraint, weight loss, and variability of body weight. J. Abnorm. Psychol. 100, 78.

Hill, A.J., Weaver, C.F., Blundell, J.E., 1991. Food craving, dietary restraint and mood. Appetite 17, 187–197.

Holmes, C.J., Hoge, R., Collins, L., Woods, R., Toga, A.W., Evans, A.C., 1998. Enhancement of MR images using registration for signal averaging. J. Comput. Assist. Tomogr. 22, 324–333.

Hopfinger, J.B., Buonocore, M.H., Mangun, G.R., 2000. The neural mechanisms of top-down attentional control. Nat. Neurosci. 3, 284–291.

Hutchison, R.M., Womelsdorf, T., Allen, E.A., Bandettini, P.A., Calhoun, V.D., Corbetta, M., Della Penna, S., Duyn, J.H., Glover, G.H., Gonzalez-Castillo, J., 2013. Dynamic functional connectivity: promise, issues, and interpretations. Neuroimage 80, 360–378.

Imamizu, H., Kawato, M., 2009. Brain mechanisms for predictive control by switching internal models: implications for higher-order cognitive functions. Psychol. Res. PRPF 73, 527–544.

Jensen, M.D., Ryan, D.H., Apovian, C.M., Ard, J.D., Comuzzie, A.G., Donato, K.A., Hu, F.B., Hubbard, V.S., Jakicic, J.M., Kushner, R.F., 2014. 2013 AHA/ACC/TOS guideline for the management of overweight and obesity in adults: a report of the american college of cardiology/american heart association task force on practice guidelines and the obesity society. J. Am. Coll. Cardiol. 63.

Kavanagh, D.J., Andrade, J., May, J., 2005. Imaginary relish and exquisite torture: the elaborated intrusion theory of desire. Psychol. Rev. 112, 446–467.

Kemps, E., Tiggemann, M., 2010. A cognitive experimental approach to understanding and reducing food cravings. Curr. Dir. Psychol. Sci. 19, 86–90.

Kemps, E., Tiggemann, M., Christianson, R., 2008. Concurrent visuo-spatial processing reduces food cravings in prescribed weight-loss dieters. J. Behav. Ther. Exp. Psychiatr. 39, 177–186.

Kimura, D., 1992. Sex differences in the brain. Sci. Am. 267, 118–125.

Klema, V., Laub, A., 1980. The singular value decomposition: its computation and some applications. IEEE Trans. Automat. Contr. 25, 164–176.

Leonardi, N., Richiardi, J., Gschwind, M., Simioni, S., Annoni, J.-M., Schluep, M., Vuilleumier, P., Van De Ville, D., 2013. Principal components of functional connectivity: a new approach to study dynamic brain connectivity during rest. Neuroimage 83, 937–950.

Leonardi, N., Van De Ville, D., 2015. On spurious and real fluctuations of dynamic functional connectivity during rest. Neuroimage 104, 430–436.

Levy, B.J., Wagner, A.D., 2011. Cognitive control and right ventrolateral prefrontal cortex: reflexive reorienting, motor inhibition, and action updating. Ann. N. Y. Acad. Sci. 1224, 40–62.

MacLean, P.S., Bergouignan, A., Cornier, M.-A., Jackman, M.R., 2011. Biology's response to dieting: the impetus for weight regain. Am. J. Physiol. Regul. Integr. Comp. Physiol. 301, R581–R600.

Marsh, A.P., Janssen, J.A., Ambrosius, W.T., Burdette, J.H., Gaukstern, J.E., Morgan, A.R., Nesbit, B.A., Paolini, J.B., Sheedy, J.L., Rejeski, W.J., 2013. The cooperative lifestyle intervention program-ii (CLIP-II): design and methods. Contemp. Clin. Trials 36, 382–393.

Mathus-Vliegen, E.M., Basdevant, A., Finer, N., Hainer, V., Hauner, H., Micic, D., Maislos, M., Roman, G., Schutz, Y., Tsigos, C., 2012. Prevalence, pathophysiology, health consequences and treatment options of obesity in the elderly: a guideline. Obesity facts 5, 460–483.

McClure, S.M., Laibson, D.I., Loewenstein, G., Cohen, J.D., 2004. Separate neural systems value immediate and delayed monetary rewards. Science 306, 503–507.

Mokhtari, F., Paolini, B.M., Burdette, J.H., Marsh, A.P., Rejeski, W.J., Laurienti, P.J., 2016. Baseline gray-and white-matter volume predict successful weight loss in the elderly. Obesity 24, 2475–2480.

Noesselt, T., Rieger, J.W., Schoenfeld, M.A., Kanowski, M., Hinrichs, H., Heinze, H.-J., Driver, J., 2007. Audiovisual temporal correspondence modulates human multisensory superior temporal sulcus plus primary sensory cortices. J. Neurosci. 27, 11431–11441.

Paolini, B.M., Laurienti, P.J., Norris, J., Rejeski, W.J., 2014. Meal Replacement: Calming the Hot-state Brain Network of Appetite.

Paolini, B.M., Laurienti, P.J., Simpson, S.L., Burdette, J.H., Lyday, R.G., Rejeski, W.J., 2015. Global integration of the hot-state brain network of appetite predicts short term weight loss in older adult. Front. Aging Neurosci. 7.

Polivy, J., Coleman, J., Herman, C.P., 2005. The effect of deprivation on food cravings and eating behavior in restrained and unrestrained eaters. Int. J. Eat. Disord. 38, 301–309.

Rabusseau, G., Kadri, H., 2016. Low-rank regression with tensor responses. Adv. Neural Inf. Process. Syst. 1867–1875.

Rejeski, W.J., Ambrosius, W.T., Burdette, J.H., Walkup, M.P., Marsh, A.P., 2017. Community weight loss to combat obesity and disability in at-risk older adults. J. Gerontol. Biol. Med. Sci. 72 (11), 1547–1553.

Rejeski, W.J., Blumenthal, T.D., Miller, G.D., Lobe, M., Davis, C., Brown, L., 2010. State craving, food availability, and reactivity to preferred snack foods. Appetite 54, 77–83.

Rejeski, W.J., Brawley, L.R., Jung, M.E., 2008. Self-management in geriatric medicine. In: Halter, J. (Ed.), Hazzard's Principles of Geriatric Medicine and Gerontology. McGraw-Hill, New York.

Rejeski, W.J., Bribaker, P.H., Goff, D.C., Bearon, L.B., McClelland, J.W., Perri, M.G., Ambrosius, W.T., 2011. Translating weight loss and physical activity programs into the community to preserve mobility in older, obese adults in poor cardiovascular health. Achieve Intern. Med. 171, 880–886.

Richiardi, J., Achard, S., Bunke, H., Van De Ville, D., 2013. Machine learning with brain graphs: predictive modeling approaches for functional imaging in systems neuroscience. IEEE Signal Process. Mag. 30, 58–70.

Ridderinkhof, K.R., Van Den Wildenberg, W.P., Segalowitz, S.J., Carter, C.S., 2004. Neurocognitive mechanisms of cognitive control: the role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. Brain Cognit. 56, 129–140.

Rogers, R.D., Ramnani, N., Mackay, C., Wilson, J.L., Jezzard, P., Carter, C.S., Smith, S.M., 2004. Distinct portions of anterior cingulate cortex and medial prefrontal cortex are activated by reward processing in separable phases of decision-making cognition. Biol. Psychiatr. 55, 594–602.

Rolls, B.J., Fedoroff, I.C., Guthrie, J.F., 1991. Gender differences in eating behavior and body weight regulation. Health Psychol. 10, 133.

Rothemund, Y., Preuschhof, C., Bohner, G., Bauknecht, H.-C., Klingebiel, R., Flor, H., Klapp, B.F., 2007. Differential activation of the dorsal striatum by high-calorie visual food stimuli in obese individuals. Neuroimage 37, 410–421.

Rueda, M.R., Posner, M.I., Rothbart, M.K., 2004. Attentional control and self-regulation. In: Handbook of self-regulation: Research, theory, and applications, 2, pp. 284–299.

Schmahmann, J.D., 1991. An emerging concept: the cerebellar contribution to higher function. Arch. Neurol. 48, 1178–1187.

Schmahmann, J.D., 1996. Dysmetria of thought: correlations and conundrums in the relationship between the cerebellum, learning, and cognitive processing. Behav. Brain Sci. 19, 472–473.

Shahnazian, D., Mokhtari, F., Hossein-Zadeh, G.-A., 2012. A method based on the granger causality and graph kernels for discriminating resting state from attentional task. Biomedical Engineering (ICoBE). In: 2012 International Conference on. IEEE, pp. 83–88.

Shen, X., Tokoglu, F., Papademetris, X., Constable, R.T., 2013. Groupwise whole-brain parcellation from resting-state fMRI data for network node identification. Neuroimage 82, 403–415.

Stice, E., Figlewicz, D.P., Gosnell, B.A., Levine, A.S., Pratt, W.E., 2013. The contribution of brain reward circuits to the obesity epidemic. Neurosci. Biobehav. Rev. 37, 2047–2058.

Stoodley, C.J., 2012. The cerebellum and cognition: evidence from functional imaging studies. Cerebellum 11, 352–365.

Stoodley, C.J., Valera, E.M., Schmahmann, J.D., 2012. Functional topography of the cerebellum for motor and cognitive tasks: an fMRI study. Neuroimage 59, 1560–1570.

Tiggemann, M., Kemps, E., 2005. The phenomenology of food cravings: the role of mental imagery. Appetite 45, 305–313.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15, 273–289.

Uddin, L.Q., Clare Kelly, A., Biswal, B.B., Xavier Castellanos, F., Milham, M.P., 2009. Functional connectivity of default mode network components: correlation, anticorrelation, and causality. Hum. Brain Mapp. 30, 625–637.

Volkow, N.D., Wang, G.-J., Baler, R.D., 2011. Reward, dopamine and the control of food intake: implications for obesity. Trends Cognit. Sci. 15, 37–46.

Volkow, N.D., Wang, G.-J., Fowler, J.S., Telang, F., 2008. Overlapping neuronal circuits in addiction and obesity: evidence of systems pathology. Philos. Trans. R. Soc. Lond. B Biol. Sci. 363, 3191–3200.

Williams, G.C., Grow, V.M., Freedman, Z.R., Ryan, R.M., Deci, E.L., 1996. Motivational predictors of weight loss and weight-loss maintenance. J. Pers. Soc. Psychol. 70, 115.

Wing, R.R., 2002. Behavioral Weight control., Handbook of Obesity Treatment. Guilford Press, New York, pp. 301–316.

Zhu, J.-N., Wang, J.-J., 2008. The cerebellum in feeding control: possible function and mechanism. Cell. Mol. Neurobiol. 28, 469–478.

Zhu, J.-N., Yung, W.-H., Chow, B.K.-C., Chan, Y.-S., Wang, J.-J., 2006. The cerebellar-hypothalamic circuits: potential pathways underlying cerebellar involvement in somatic-visceral integration. Brain Res. Rev. 52, 93–106.