Received: 18 July 2022 | Revised: 1 June 2023 | Accepted: 7 June 2023

DOI: 10.1002/hbm.26403

RESEARCH ARTICLE

WILEY

The robustness of persistent homology of brain networks to data acquisition-related non-neural variability in resting state fMRI

Sidharth Kumar[1] | Ahmedur Rahman Shovon[1] | Gopikrishna Deshpande[2,3,4,5,6,7,8,9]

[1]Computer Science Department, University of Alabama at Birmingham, Birmingham, Alabama, USA

[2]Department of Electrical and Computer Engineering, AU MRI Research Center, Auburn University, Alabama, USA

[3]Department of Psychological Sciences, Auburn University, Auburn, Alabama, USA

[4]Alabama Advanced Imaging Consortium, Birmingham, Alabama, USA

[5]Center for Neuroscience, Auburn University, Auburn, Alabama, USA

[6]School of Psychology, Capital Normal University, Beijing, China

[7]Key Laboratory for Learning and Cognition, Capital Normal University, Beijing, China

[8]Department of Psychiatry, National Institute of Mental Health and Neurosciences, Bangalore, India

[9]Centre for Brain Research, Indian Institute of Science, Bangalore, India

Correspondence
Gopikrishna Deshpande, Department of Electrical and Computer Engineering, AU MRI Research Center, Auburn University, AL, USA.
Email: gopi@auburn.edu

Funding information
National Institute of Mental Health, Grant/Award Number: R01MH094639-01; New York State Office of Mental Health; Research Foundation for Mental Hygiene; Child Mind Institute, Grant/Award Number: 1FDN2012-1; NKI Center for Advanced Brain Imaging (CABI); Brain Research Foundation; Stavros Niarchos Foundation

Abstract

There is increasing interest in investigating brain function based on functional connectivity networks (FCN) obtained from resting-state functional magnetic resonance imaging (fMRI). FCNs, typically obtained using measures of time series association such as Pearson's correlation, are sensitive to data acquisition parameters such as sampling period. This introduces non-neural variability in data pooled from different acquisition protocols and MRI scanners, negating the advantages of larger sample sizes in pooled data. To address this, we hypothesize that the topology or shape of brain networks must be preserved irrespective of how densely it is sampled, and metrics which capture this topology may be statistically similar across sampling periods, thereby alleviating this source of non-neural variability. Accordingly, we present an end-to-end pipeline that uses persistent homology (PH), a branch of topological data analysis, to demonstrate *similarity* across FCNs acquired at different temporal sampling periods. PH, as a technique, extracts topological features by capturing the network organization across all continuous threshold values, as opposed to graph theoretic methods, which fix a discrete network topology by thresholding the connectivity matrix. The extracted topological features are encoded in the form of persistent diagrams that can be compared against one another using the earth-moving metric, also popularly known as the Wasserstein distance. We extract topological features from three data cohorts, each acquired at different temporal sampling periods and demonstrate that these features are statistically the same, hence, empirically showing that PH may be robust to changes in data acquisition parameters such as sampling period.

KEYWORDS
fMRI, persistent homology, TDA, topological data analysis

1 | INTRODUCTION

Studying the functional and anatomical connectivity of the human brain has given us a deeper understanding of the characteristics of the brain, from microscale connectivity between single neurons to macroscale connectivity between regions of interest in whole brain images. One of the routinely employed approaches for characterizing macroscale connectivity utilizes measures of statistical association (such as Pearson's correlation) obtained from resting-state functional magnetic resonance imaging (rs-fMRI) times series corresponding to specific ROIs in the human brain. The connectivity matrix thus obtained is an algebraic representation of the weighted brain network, which shows the relationship between all pairs of nodes.

Functional connectivity obtained from rs-fMRI has been shown to be extremely sensitive to mental health disorders (Fornito & Bullmore, 2010) as well as predictive of behavior in healthy individuals (Miller et al., 2016). This has kindled interest in using functional connectivity networks (FCNs) as potential biomarkers of mental disorders. The sensitivity and specificity of these biomarkers and their generalizability to the general population seem to increase with sample size (Schnack & Kahn, 2016). However, acquiring data from larger samples at any given site can be economically and logistically prohibitive. Therefore, there has been a recent impetus toward post-hoc aggregation of data acquired at different sites to form larger datasets. Examples include the Autism Brain Imaging Data Exchange (ABIDE; Di Martino et al., 2014) and ADHD-200 (Bellec et al., 2017). In such large publicly available datasets, it has been found that biomarkers do not generalize well to data acquired at different sites (Abraham et al., 2017; Khalili-Mahani et al., 2017). This has been attributed to the fact that MRI scanners and data acquisition protocols are different across sites, and this induces an element of non-neural variability in the data that tend to make it difficult for us to discover consistent inter-group FCN differences that are at least partly neural in origin.

Topological data analysis (TDA) is an emerging field that has found recent applications in functional neuroimaging. Traditionally, graph-theoretic tools, which can be construed as special cases of the more generic concepts of TDA, have been extensively used to study and quantify FCNs. More recently, advanced tools from TDA, such as persistent homology (PH) (Rubinov & Sporns, 2010) have been used to study complex networks. PH investigates connections between different parts of networks using algorithms designed to encode and measure the significance of relationships across multiple scales (thresholds). In the context of networks, topological features refer to the 0-, 1-, and 2-dimensional homology groups of a metric space that describe its *connected components*, *tunnels*, and *voids*, respectively. Most graph-theoretic techniques for analyzing weighted brain networks can only measure the topological features of the network at a specific threshold and cannot capture how these features change as the threshold value is varied. PH provides a principled approach to quantifying these features for all thresholds; more precisely, it can track when features (such as connected components, loops, and voids) are created and destroyed with varying scales (threshold). The technique quantifies the individual topological events (birth and death of features) in the graph according to their significance (or persistence). This persistence is represented in the form of barcodes, which encode the threshold at which features appear and disappear. The barcodes encoding these sets of features can be seen as a fingerprint for a graph. What makes this fingerprint useful is the presence of metrics such as Wasserstein distance (WD; Edelsbrunner, 2013; Vallender, 1974) that can be used to quantify the statistical difference between two barcodes robustly. The WD is robust to small perturbations in the data and hence can be used to compare and establish similarities between persistent diagrams. We use this ability in our pipeline to establish similarities between barcodes obtained from brain networks derived from data acquired with different acquisition parameters (such as sampling period).

The input to our TDA-based pipeline is subject-specific FCNs from three data cohorts, corresponding to data acquired from the same individuals at three different sampling periods (TR): 645 ms, 1400 ms, and 2500 ms. The cohort consists of rs-fMRI data from 316 subjects (totaling 3×316 scans). In our pipeline, the topological features of each network instance are extracted using PH and encoded with a barcode. The barcodes are compared against each other by using the WD. In particular, we perform two sets of experiments to demonstrate that the FCNs indeed capture the same structure irrespective of the temporal sampling periods: (a) making a direct pairwise comparison of the same subjects across different temporal sampling periods and (b) performing pairwise comparison within a cohort (same temporal sampling period) to extract the overall pattern of subjects and then comparing that pattern across the sampling periods. The statistical analysis is made possible because the barcodes (our metric) can be compared against one another using the WD (Edelsbrunner, 2013; Vallender, 1974). In our first set of experiments, we perform pairwise WD computation of the same subjects but between different sampling periods (across the data cohorts), and this yields three groups of measurements: WD between (645 and 1400 ms), (1400 and 2500 ms), and (2500 and 645 ms). We then demonstrate similarity among these distributions by performing ANOVA and *t*-tests, further establishing the fact that FCNs capture similar information, irrespective of the data acquisition parameters. In our second set of experiments, we calculate pairwise WD between persistence diagrams of all 316 subjects within the same sampling rate. As we have data from three TRs, this step yields three sets of adjacency matrices (of size 316×316). To apply statistical comparison in this high dimensional data, we perform multidimensional scaling (MDS; Carroll & Arabie, 1998; Cox & Cox, 2008) with two components that reduces it to a 2D space and then apply clustering techniques to segregate the subjects into clusters. Finally, we analyze the number of overlapping subjects across the clusters to examine the similarity between the subjects across various parameters using the PH-based pipeline. In particular, our paper makes the following three contributions to the literature:

1. Demonstrate that the barcode obtained from the PH of resting state fMRI data is a compact representation of topological information of an FCN.

2. Present an end-to-end pipeline that uses PH-based techniques to demonstrate that TDA-based metrics derived from FCN networks are statistically similar even when data acquisition parameters such as temporal sampling periods (TR) are varied, potentially removing a source of noise in multi-site case–control studies, thereby improving the effect sizes in group comparisons.
3. Present an open-source and reproducible codebase for all components of our work. Code, including scripts, documentation, and data, can be found here: https://github.com/harp-lab/brainPH

The rest of the paper is organized as follows: In Section 2, we present relevant related work covering both graph-based and topology-based FCN analysis frameworks. In Section 3, we present our end-to-end TDA pipeline comprising four key steps. Finally, in Section 4, we present the result of applying our methods to real data. And we then conclude with a discussion in Section 5 and a conclusion in Section 6.

2 | RELATED WORK

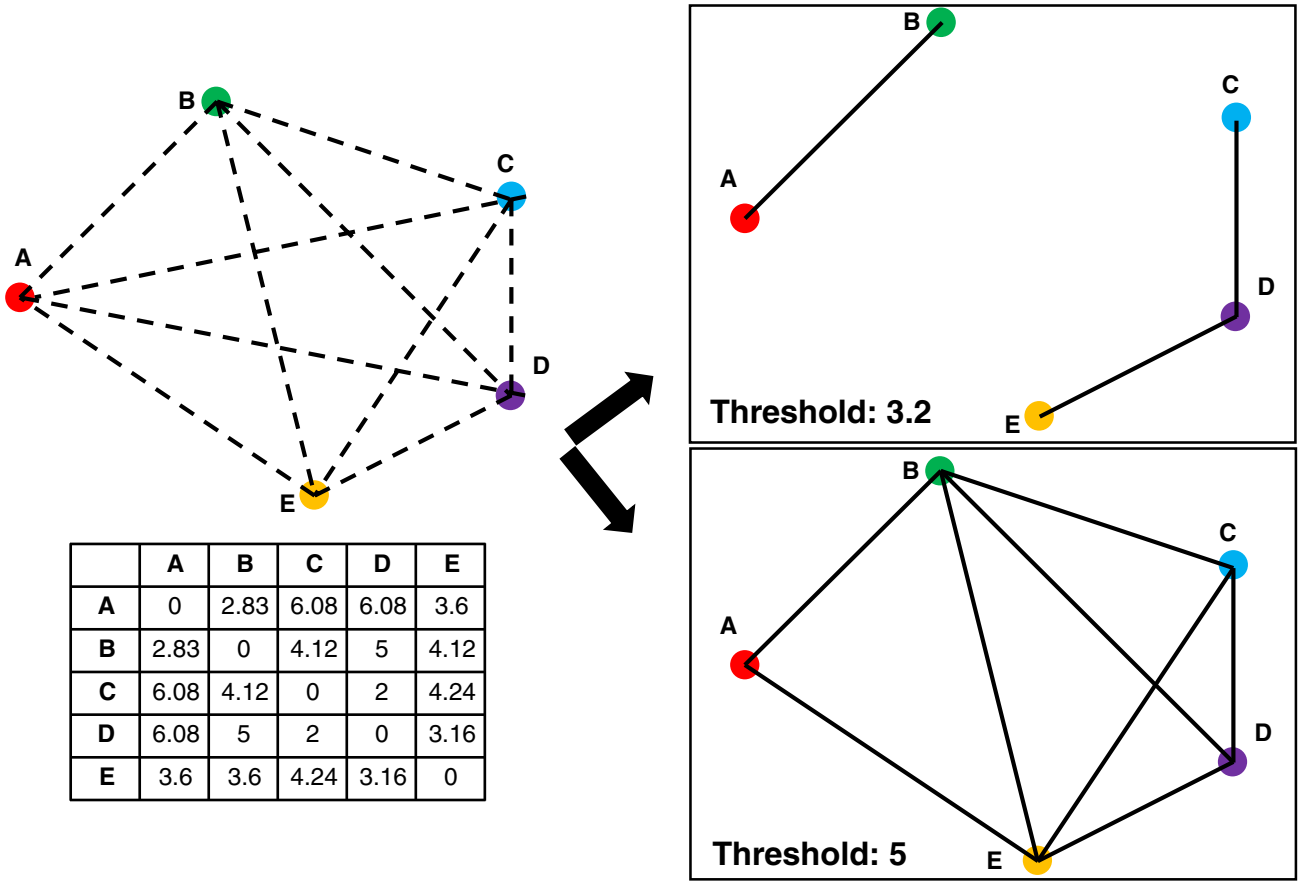
Existing methods in analyzing and characterizing FCNs rely heavily on graph analysis measures (Andellini et al., 2015; Aurich et al., 2015; Ginestet et al., 2014; Termenon et al., 2016) such as clustering coefficient and node degree. These measures are well known to summarize a single weighted network and also to compare FCNs within a collection of networks. For network comparison, typically, there are two cases: comparing two individual networks or comparing two collections of networks where the networks may be paired or unpaired between collections. Networks can be compared either at single nodes and links (Ginestet et al., 2014; Narayan & Allen, 2015), or via some functional or transform summarizing each network (Simpson et al., 2013) using graph analysis approaches. Graph analysis measures, however, have been criticized for being dependent on the choice of network link threshold (Garrison et al., 2015). Graph analysis methods typically use binary measures of link strength, ignoring the weights of links. However, not all FCN links are equal. Some links are expected to be detected more frequently or be more strongly detected than others, and recognizing these differences in link strength allows us to understand the function within networks. Weighted graph analysis methods have been proposed (Rubinov & Sporns, 2010), but are still susceptible to variation with respect to network density (Ginestet et al., 2014). Graph analysis of weighted brain networks at multiple thresholds has also been proposed (Drakesmith et al., 2015).

In recent years, multi-site fMRI studies have become more common as they allow for faster participant recruitment and larger sample sizes, resulting in increased statistical power (Biswal et al., 2010; Noble et al., 2017). This is especially important when studying rare disorders, subtle effects, and diverse populations (Dansereau et al., 2017). However, non-biological variability in these studies can arise from differences in scanner manufacturers, imaging acquisition parameters, and other factors (Shinohara et al., 2017). These sources of variability can decrease statistical power and produce misleading results. Many studies have reported site or scanner effects in fMRI data, but only a few have attempted to standardize protocols and image acquisition parameters to mitigate these effects (Chavez et al., 2018; Shinohara et al., 2017). Despite efforts to standardize protocols and image acquisition parameters, scanner-to-scanner variation caused by the use of scanners from different manufacturers remains present (Noble et al., 2017). An independent component analysis (ICA)-based approach was employed in one study to decrease scanner differences in multi-site resting-state fMRI post-acquisition (Feis et al., 2015). However, this approach was not completely successful in eliminating the structured noise that arises from the use of different scanners (Yu et al., 2018). Thus, the generation of harmonization techniques for multi-site data has become an emerging topic in neuroimaging (Roffet et al., 2022). ComBat is one of the prominent and fast harmonization methods that can reduce the multi-site or multi-scanner effect (Bell et al., 2022; Fortin et al., 2017, 2018; Ingalhalikar et al., 2021; Yu et al., 2018). Existing literature indicates that ComBat may not completely retain inter-subject biological variability following harmonization, particularly in the presence of nonlinear scanner contributions (Cetin-Karayumak et al., 2020, 2023). Even though this was shown for diffusion data, the principles are equally applicable to fMRI data.

TDA of networks goes beyond graph-theoretic analysis by utilizing tools from computational topology to describe the architecture of networks or data structures in more flexible ways (Edelsbrunner & Harer, 2022; Ghrist, 2008). In particular, it encodes higher-order (not just pairwise) interactions in the system and studies the topological features of a network across all possible thresholds. PH, an advanced technique in TDA, is an emerging tool for studying complex networks, including brain networks (Cassidy et al., 2015; Dabaghian et al., 2012; Lee et al., 2011). PH-based methods have shown promising results in modeling transitions between brain states in fMRI data (Saggar et al., 2018). There are many excellent introductions to PH, such as the books (Ghrist, 2014; Oudot, 2015; Zomorodian, 2005) and the papers (Edelsbrunner & Harer, 2008a, 2022; Ghrist, 2008; Patania et al., 2017; Weinberger, 2011).

3 | METHODS

In a typical connectivity analysis framework, the correlation matrix (FCN) is thresholded to obtain a graph of the brain network. The graph thus obtained merely captures a *snapshot* of the FCN whose topology is dictated by the choice of the threshold. The onus, therefore, is to pick the correct threshold, which is a very challenging task, and moreover, usually, one threshold (and one graph) is not enough to capture the complex connectivity information of the FCN (see Figure 1). PH, a branch of TDA, allows one to capture the change of topological structures over the whole range of thresholds. Our approach thus uses PH to extract topological features from rs-fMRI FCNs. The features are used to demonstrate a notion of similarity across FCNs extracted from datasets acquired with different data



	A	B	C	D	E
A	0	2.83	6.08	6.08	3.6
B	2.83	0	4.12	5	4.12
C	6.08	4.12	0	2	4.24
D	6.08	5	2	0	3.16
E	3.6	3.6	4.24	3.16	0

FIGURE 1 An example of graph analytics-based approach for analysis of FCNs. Here, the two thresholds yield graphs of completely different topologies.

acquisition parameters (temporal sampling rate in our case). Our pipeline consists of three key steps: extracting FCNs from fMRI data, followed by PH-based TDA that extracts topological features from FCNs, and finally, statistical analysis to find similarities/differences among the features extracted from data acquired with different TRs. The main goals of this work are to demonstrate the efficacy of PH in extracting relevant features from fMRI networks and show that these techniques are robust to changes in data acquisition parameters. The end-to-end pipeline of our TDA-based framework can be detailed in the following steps:

1. Reduce spatiotemporal fMRI data to FCNs: rs-fMRI data are subjected to a standard pre-processing pipeline followed by estimation of FCNs using Pearson correlation between time series extracted from pre-defined regions across the brain (Section 3.1).
2. Embed FCNs in metric space: This is a pre-processing step required for the execution of TDA. It yields a weighted graph that encodes the correlation between every node pair (Section 3.2).
3. Apply PH to metric space: This step yields barcodes/persistent diagrams. These barcodes/persistent diagrams act as fingerprints of an FCN (Section 3.3).
4. Statistical analysis on persistent diagrams: Two kinds of analysis are performed; we show methods for both comparing the architectures within a collection of networks and between collections of networks; both analyses demonstrate that PH-based techniques are robust to changes in data-acquisition parameters (Section 3.4).

3.1 | Experimental data and network construction

Structural T1-weighted and rs-fMRI data were obtained from the openly available Enhanced Nathan Kline Institute Rockland Sample database (NKI-RS; Nooner et al., 2012). The MRI data were obtained using a 3 T Siemens Magnetom Tim Trio scanner. The data acquisition parameters for the T1-weighted structural data were: 1.0 mm isotropic voxels with 176 slices, repetition time (TR) = 1900 ms, echo time (TE) = 2.52 ms, and field of view (FOV) = 250 × 250. Resting-state fMRI data were acquired using multiband echo-planar imaging (EPI; Feinberg et al., 2010) from each subject using three different acquisition protocols with different parameters as follows: (1) 3.0 mm isotropic voxels with 40 slices, TR = 645 ms, TE = 30 ms, FOV = 222 × 222 mm, number of volumes = 900, and multi-band factor = 4. (2) 2.0 mm isotropic voxels with 64 slices, TR = 1400 ms, TE = 30 ms, FOV = 224 × 224 mm, number of volumes = 404 and

multi-band factor = 4. (3) 2.0 mm isotropic voxels with 38 slices, TR = 2500 ms, TE = 30 ms; FOV = 216 × 216 mm, number of volumes = 120 and multi-band factor = 1. It can be seen that even though we identify the three datasets from each subject with the corresponding TR, the data differ in many other scan parameters, such as the number of volumes, multiband factor, FOV and, voxel size.

The MRI data were subjected to a standard pre-processing pipeline, including the first five volume removal, slice time correction, and motion correction. T1-weighted anatomical images were coregistered to the mean functional images, using which the fMRI images were spatially registered to a standard MNI152 template. Nuisance variables such as low-frequency drifts, and motion parameters were regressed out. Unwanted physiological fluctuations (white-matter and cerebrospinal fluid signals) were removed using aCOMPCor (anatomical component-based noise correction). After removing subjects that failed quality control, 316 subjects were identified to have usable data from all three acquisition protocols. Subsequently, mean time series from 113 brain regions (obtained using the Yeo parcellation template; Thomas Yeo et al., 2011) were obtained for each subject and acquisition protocol. Using Pearson's correlation, FCN matrices were estimated from these mean time series. At the end of this phase, we end up with one weighted network represented per fMRI scan. The network is stored in the form of a symmetric adjacency matrix W of dimension 113×113, where w_{ij} corresponds to the correlation coefficient between brain parcels (nodes) i and j. For the three data cohorts and 316 subjects per cohort, we get 948 (3×336) adjacency matrices. The FCN for subject 31 for the three sampling periods (645, 1400, and 2500 ms) is shown in Figure 2 as a representative example.

3.2 | Mapping network to metric space

Well-understood techniques in TDA typically focus on the study of point cloud data under the metric space setting. In order to study FCN data, our approach is to embed the weighted graph in a metric space where topological techniques can be applied. Once a network is embedded in the metric space, topological features in the form of *barcodes* can be effectively extracted using PH.

In this paper, the association between any two nodes u and v in the brain network is measured by their Pearson's correlation coefficients $corr(u,v)$. The idea is to map this association to a distance measure such that higher correlations between nodes map to smaller distances. We use the following mapping to all 3×316 (316 subjects across three temporal frequencies) FCNs:

$$d(u,v) = 1 - corr(u,v)$$

Although other measures of association between time series have been explored for constructing FCNs, Pearson's correlation is by far the most popular method, and the mentioned mapping is used in existing literature (Cassidy et al., 2015; Edelsbrunner & Harer, 2008b; Lee et al., 2011, 2012; Smith et al., 2011). That said, the pipeline proposed here does not depend on how the FCN is constructed and could be applied to scenarios where FCN is constructed using measures other than Pearson's correlation. In such cases, the mapping of the network to metric space would have to be adapted accordingly. The applied correlation value will be different for correlated and anti-correlated nodes. As the correlation between two nodes increases, the distance between them will decrease. We also experiment with positive and negative correlation separately and observe that the positive correlation produces a similar result to the original dataset (discussed in the supplement). We hypothesize that this is due to the fact that only 10% of all connections (across all subjects) have a negative correlation; therefore, it does not have a major impact on the analysis. Hence, we have included both positive and negative correlations in our analyses.

3.3 | Extracting topological features

We apply PH on the metric space representation of brain fMRI graphs to extract topological features. To describe our process, we first briefly review PH, including barcodes and the WD metric. For more details and background on persistence homology, see (Edelsbrunner & Harer, 2008a, 2022).

3.3.1 | Topological features

Homology deals with the topological features of a space. More precisely, the homology of a topological space $\mathbb{X}$ is represented by its homology groups $H_0(\mathbb{X})$, $H_1(\mathbb{X})$, and $H_2(\mathbb{X})$ ..., where the k^{th} homology group $H_k(\mathbb{X})$ describes, informally, the number of k-dimensional holes in $\mathbb{X}$. The 0-, 1-, and 2-dimensional homology groups, denoted as $H_0(\mathbb{X})$, $H_1(\mathbb{X})$, and $H_2(\mathbb{X})$, respectively, correspond to (connected) components, tunnels, and voids of $\mathbb{X}$. In our experimental setup, we focus on the 0-dimensional topological features of an FCN (W), which correspond to $H_0(\mathbb{X})$ of its metric space representation. The 0-dimensional topological features, roughly speaking, capture connected components by vertices in the instances.

3.3.2 | Persistent homology

In practice, these features, such as connected components, loops, and voids in an FCN, can also be captured by the traditional graph analytic approach using thresholding. An example of this pipeline is shown in Figure 1; here, we have an FCN with five nodes and the pairwise distance between every node. We demonstrate how different thresholds can yield different instances of graphs with completely different topologies and connectivity. For instance, a threshold of 3.2 yields a sparse graph with two connected components, and a threshold of 5 yields a relatively denser graph with only one connected component. In practice, there does not exist a unique scale (threshold) that captures the topological structure and connectivity of a weighted

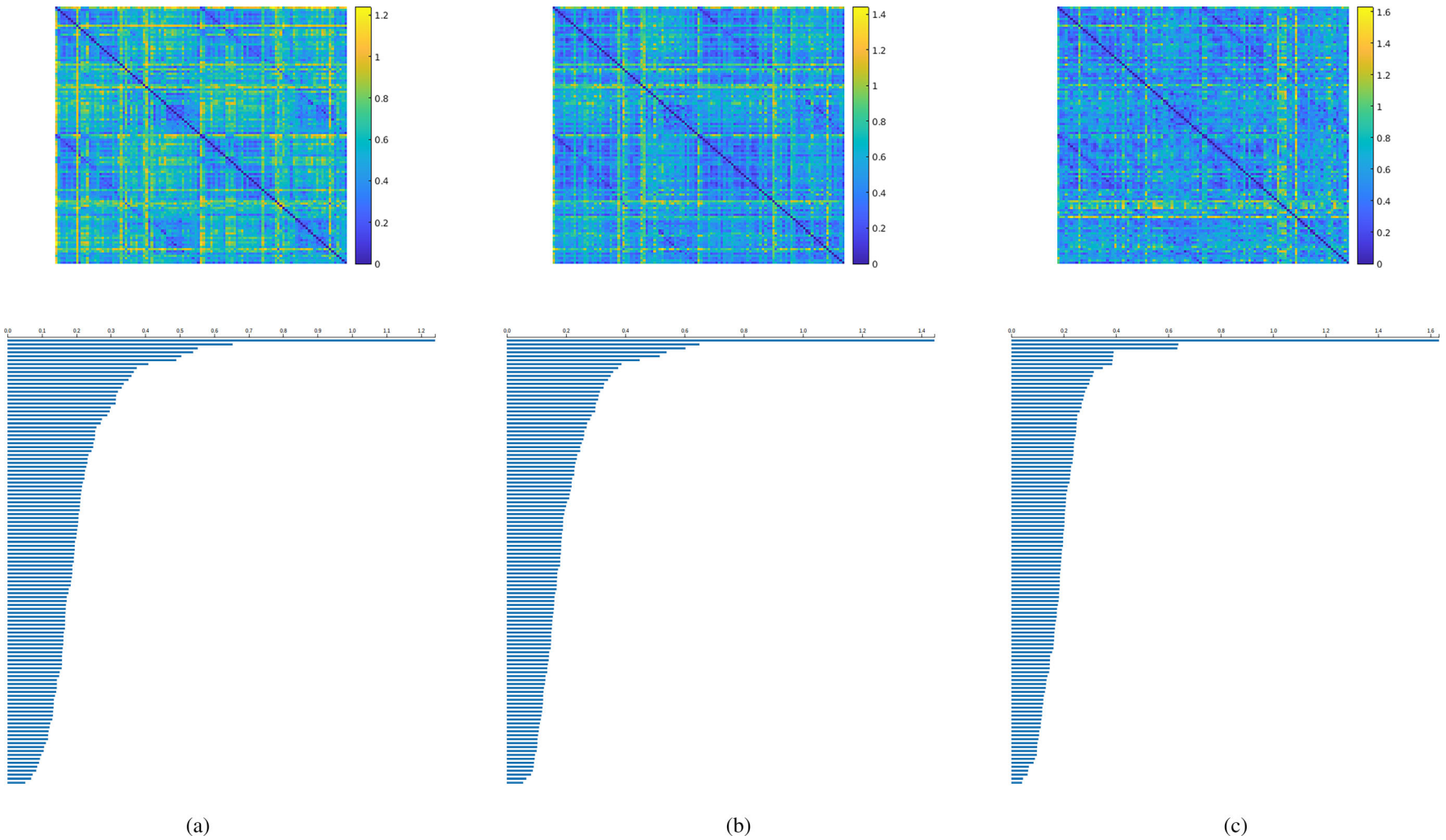


FIGURE 2 FCN (top-row) and extracted topological feature (bottom-row) for Subject 31 for temporal sampling period 645 ms (left), 1400 ms (center), and 2500 ms (right). The adjacency matrix is of size 113 × 113. (a) Temporal sampling period 645 ms. (b) Temporal sampling period 1400 ms. (c) Temporal sampling period 2500 ms.

FCN. Instead of trying to determine one proper threshold that may not really capture the structure of the FCN, we decided to look at the overall change of topological structure over the whole range of thresholds using PH.

PH is a multi-scale notion of homology that captures the change of topological structures over the whole range of thresholds. Similar to the example of Figure 1, we start with a set F of point cloud data consisting of p points, with a known distance between every pair of points. Two points i and j will be connected by an edge if the distance $d(i,j)$ is smaller than ϵ, forming a 1-simplex. Similarly, a 2-simplex (a triangular face) will be formed among three points if there is an edge between every pair of points. The generated graph is a Rips complex and denoted by $Rips(F,\epsilon)$.

Figure 3c–f shows a toy example of constructed Rips complex with different ϵ. It can be seen that $Rips(F,t) \subset Rips(F,s)$ whenever $t \le s$. This topological transition with increasing filtration value ϵ is called as a Rips filtration and is the key idea behind PH.

More formally, for a real number $t \ge 0$, a Rips complex denoted as $Rips(F,t)$, is formed by considering a set of balls of radius $t/2$ centered at points in F. A 1-simplex (an edge) is formed between two points in F if and only if their balls intersect (see Figure 3c). Given a finite point set F, continuously increasing the radius forms a one-parameter family of nested unions of balls; and, correspondingly, a one-parameter family of nested Rips complexes referred to as a Rips filtration. Let $0 = r_0 \le r_1 \le r_2 \le r_3 \ldots \le r_m$ denote a finite sequence of increasing radius. The Rips filtration is a sequence of Rips complexes connected by inclusions, $R(r_0) \subset R(r_1) \subset R(r_2) \subset R(r_3) \subset \ldots \subset R(r_m)$.

As t increases, we focus on the important events when the topology of the space changes. This change occurs when components merge with one another to form larger components. We track the *birth* and *death* times of each topological feature (a *component* or a *tunnel*), and visualize them in the form of *barcodes*. The lifetime of a feature ($t_{death} - t_{birth}$) is called its persistence. In Figure 3, each colored point (red, green, blue, purple, and yellow) is born (appears) at $t = 0$ as an independent (connected) component. At $t = 2$, the blue component merges with the purple component and dies (disappears). Therefore, the blue component has a persistence of 2. At $t = 2.2$, the yellow component merges into the purple component and dies. Hence, it has a persistence of 2.2. Similarly, the green component dies at $t = 2.8$, and the red component dies at $t = 3.15$. The purple component born at time 0 never dies, and thus it has a persistence of ∞. In Figure 3g, we visualize the appearance (birth), the disappearance (death), and the persistence of these topological features via the *barcode*, where each feature is summarized by a horizontal bar that begins at its birth time and ends at its death time in the X-axis. The Y-axis depicts that the bars are from 0-dimensional homology groups ($H_0(\mathbb{X})$; Ghrist, 2008). In our setting, for every FCN, all 113 vertices form a finite set of points F, and d_i encodes the pairwise distance among points in F. We apply PH on all our 3×316 FCNs (3 scans for every 316 subject) and extract its barcode. The right part of Figure 2 shows the barcodes generated from experimental fMRI data from an illustrative subject. We used the Gudhi library to compute the 0-dimensional persistence diagram from the simplex tree (The GUDHI Project, 2015). As we set *min_persistence* to zero, any value greater than this in the simplex tree will be considered for computing the persistent diagram. For range, the method finds all the barcodes until there is only one connected component, as we set *max_dimension* to 1 in simplex tree creation.

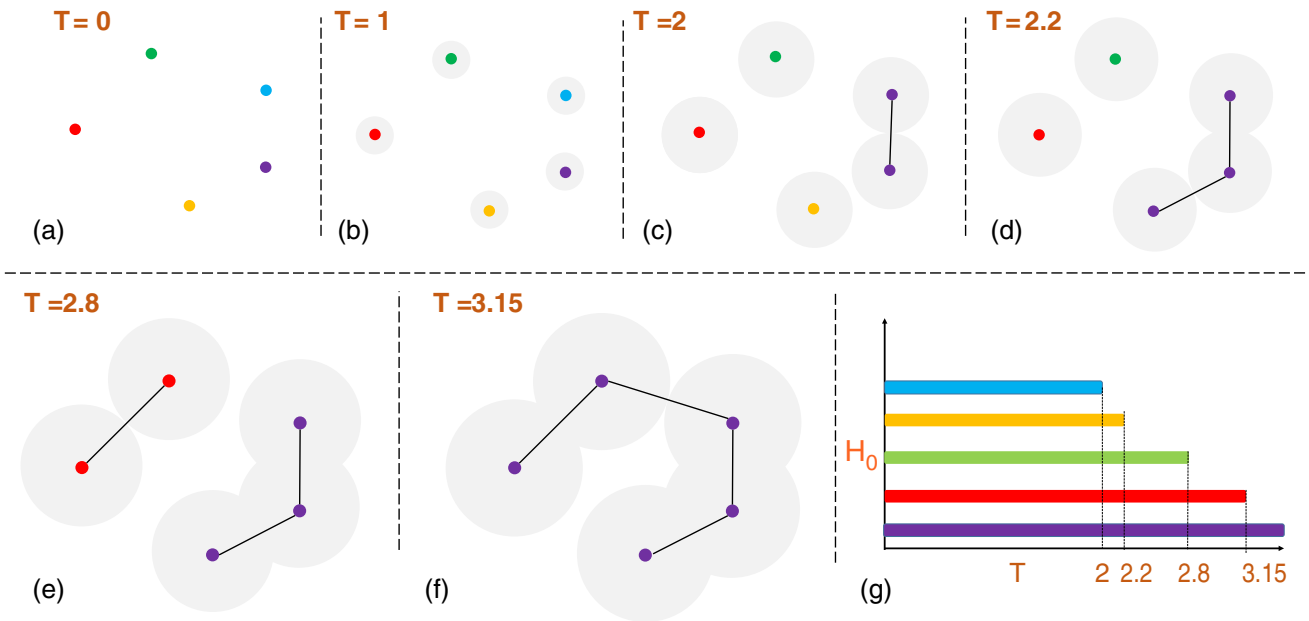


FIGURE 3 Computing the persistent homology of a point cloud. At $T = 0$, there are five topological features, corresponding to the five connected components, as we increase T, we see these connected components merge. (c–f), each shows a T where connected components merge, which corresponds to the death of topological features. The birth and death of these topological features are captured in the form of *barcodes* shown in (g).

3.4 | Statistical inferencing

A barcode can be interchanged with a persistent diagram without the loss of any information. A barcode can be drawn in the form of persistent diagrams that records the birth threshold (X-axis) and death threshold (Y-axis) of all topological features (bars). Every bar of the barcode is reduced to a point in the persistent diagram. The *persistent diagrams* are then used in a statistical inference framework to establish that PH, in general, is robust to data acquisition parameters such as sampling rate. The core idea behind the statistical inference framework is our ability to quantify the difference between two persistent diagrams using the earth moving distance, such as the WD (Edelsbrunner, 2013; Vallender, 1974).

To quantify the structural difference between two FCN instances W_i and W_j, we compute the WD between their persistent diagrams. WD is defined as the minimum value achieved by a perfect matching between the points (features) of the two persistent diagrams. It is achieved by capturing the perturbation across every pair of points (feature) in the diagrams. In particular, the WD measures the similarity between two persistence diagrams using the sum of all edge lengths between every feature pair. The ability to compare two persistent diagrams using the Wasserstein metrics provides the foundation to develop our statistical inference framework. The goal is to establish a notion of similarity between FCNs that are derived at different sampling rates. We perform three kinds of statistical analysis: (a) comparison across cohorts, (b) comparison within a cohort, and (c) comparison with traditional FCN analysis. We used the Gudhi framework (Kerber et al., 2017) to compute the WD between two persistent diagrams.

3.4.1 | Comparison across cohorts

In this first set of experiments, we perform a comparison of persistent diagrams across cohorts. In particular, we compute the WD between the persistent diagram of the same subject but with data acquired at different sampling rates and other acquisition parameters. Since we have data at three temporal sampling rates (TR): 645, 1400, and 2500 ms, for each subject, we compute three WD between persistent diagrams of data acquired at (1) TR = 645 and 1400 ms, (2) TR = 1400 and 2500 ms, and (3) TR = 2500 and 645 ms. For the three pairs, we get three distributions, each with 316 data points (= number of subjects). We perform the Analysis of Variance (ANOVA) test to demonstrate statistical similarity among the three distributions. ANOVA is a collection of statistical models and their associated estimation procedures (such as the *variation* among and between groups) used to analyze the differences among means. ANOVA provides a statistical test of whether two or more population means are equal and therefore generalizes the t-test beyond two means. The ANOVA test yields a p-value, and typically in statistics, a high p-value (>0.05) is a strong indication that distributions have identical means, that is, the null hypothesis is true. In our case, a high p-value would indicate that three WD between the pairs is statistically the same data and that the three data cohorts 645, 1400, and 2500 ms are encoding the same underlying data, and thus would infer that the PH-based technique is indeed robust to data acquisition parameters. Along with the ANOVA test, we also perform pairwise t-value tests between the three distributions, establishing statistical similarity across the distributions in a more conservative sense. An illustration of this pipeline is given in Figure 4.

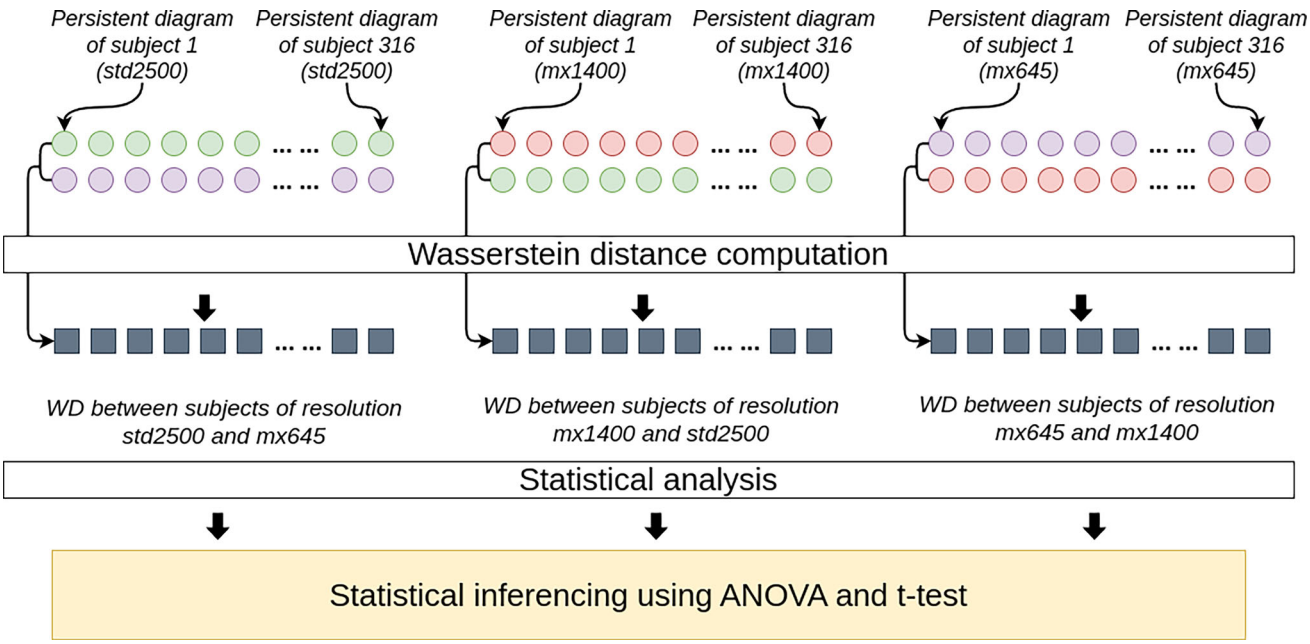


FIGURE 4 Comparison of WD across cohorts. Here, WD is obtained by comparing persistent diagrams obtained from the same subjects but with data acquired using different temporal samplings.

3.4.2 | Comparison within a cohort

In this second set of experiments, we compute the pairwise WD between persistence diagrams of subjects within each of the data cohorts. Since we have data from three sampling periods, we get three adjacency matrices (of size 316×316) encoding the pairwise WD for all 316 subjects. An adjacency matrix of 316×316 is representative of a high dimensional space (of dimension 316), and it is often a challenging task to perform an interpretable form of statistical comparison for high dimensional data. We, therefore, use the commonly used dimensionality reduction technique, MDS. For an adjacency matrix, storing the distances between each pair of objects in a set and a chosen number of dimensions, N, an MDS algorithm places each object into N-dimensional space (N is lower-dimensional) such that the between-object distances are preserved as well as possible. We use $N = 2$ in our experiments, which results in scatter plots. Applying clustering to the scatter plot generated by MDS is common, yielding segregated groups/clusters. Therefore, we apply k-means clustering using the silhouette coefficient to automatically divide the subjects into clusters. As the ground truth of the clustering is not known, we specify the number of clusters using the silhouette coefficient (Pedregosa et al., 2011; Rousseeuw, 1987). We perform this pipeline on all three data cohorts corresponding to the three temporal sampling frequencies. Only the number of clusters is not a good parameter to analyze the similarity of subjects across different TRs—as this method is coarse-grained and loses the identity information of the subjects. We, therefore, performed further fine-grained analysis on the clustering result, taking into account the identities of all subjects. In particular, we compute the set intersection (overlaps) of subjects across the clusters to keep the identity information of the subjects. Ideally, the same set of subjects should be clustered together in the three TR plots. The higher number of overlapping subjects across the clusters will indicate the higher similarity of subjects across different temporal sampling rates. We also compare the result obtained from original data with a null distribution obtained from random FC matrices to verify that the set intersection (overlaps) of subjects across the clusters is statistically significant as compared to random chance. An illustration of this pipeline is given in Figure 5.

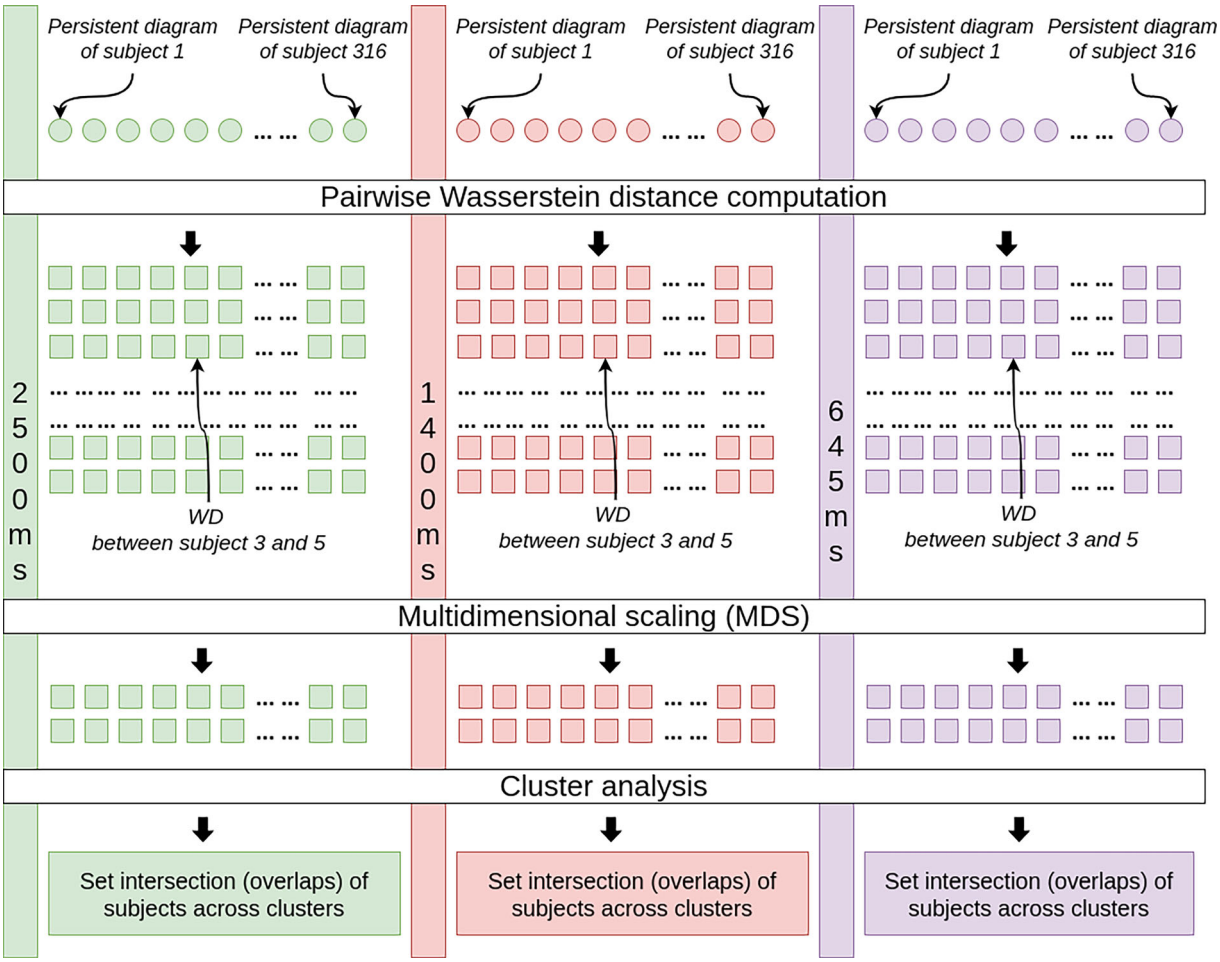


FIGURE 5 Comparison of WD within a given cohort. Here, WD is obtained by comparing persistent diagrams obtained from different subjects but within the same data cohort, acquired using the same temporal sampling.

3.4.3 | Comparison with traditional FCN analysis

The true efficacy of the TDA-based pipeline can be demonstrated further by comparing it against the traditional FC analysis pipeline. Traditionally, raw FCNs are used instead of using a topological feature, for example, a common pipeline is to perform ANOVA across all subjects for every node of the raw FCN. This gives one p-value per entry of the FCN. With our FCN of size 113×113, we get 113×113 p-values. The p-values can then be thresholded using a typical significance level of $\alpha = 0.05$. A value of 1 is assigned if the p-value is less than 0.05, or else 0 is assigned. The thresholded output can then be analyzed to see if a majority of values are less than 1 or not. If a majority of values are 1 (p-value less than 0.05), then this is a strong indication that FC is finding differences in the brain network of the same individual constructed from data acquired with different parameters. This is not ideal since we want a metric that is robust to non-neural variability in the data.

4 | RESULTS

The first step (Section 3.1) yields one FCN per fMRI scan, stored in the form of an adjacency matrix. We obtain a total of 3×316 matrices, corresponding to three scans of different temporal samplings for 316 subjects. As an illustrative example, we have plotted the FCN for subject 31 for the three data acquisition protocols in Figure 2. The next step (Section 3.2) embeds the FCN in a metric space, limiting the entries between 0 and 1. Finally, PH is applied to the metric space FCN to extract dimension-0 features, which are stored in the form of persistent barcodes. Figure 2 also shows the barcodes for subject 31 for the three temporal sampling periods. With the barcodes extracted for all 3×316 fMRI scans, we move on to statistical inference. We perform two sets of TDA-based experiments, as presented in Section 3.4, where we compare FCNs across cohorts and within the cohort.

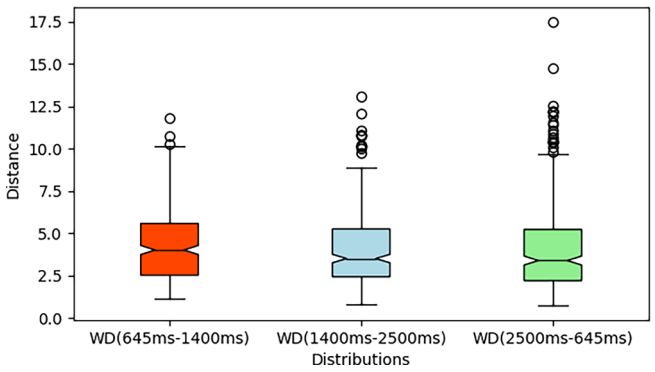


FIGURE 6 Wasserstein distance computation for all 316 subjects between the three cohorts: WD (TR = 645 and 1400 ms), WD (TR = 1400 and 2500 ms), WD (TR = 2500 and 645 ms), yielding three distributions. The three distributions are plotted using box plots.

In our first set of experiments, we compare FCNs across cohorts. For all 316 subjects, we compute the WD for the following three pairs: (1) TR = 645 and 1400 ms, (2) TR = 1400 and 2500 ms, and (3) TR = 2500 and 645 ms. The results of the three distributions (with 316 samples each) are plotted using box plots in Figure 6 (Williamson et al., 1989). As can be seen, the box plots are a strong qualitative indication of the similarity among the three distributions, with the median, interquartile region, minimum and maximum (excluding the outliers) all aligning across the three distributions. To further test their statistical equivalence, an ANOVA test was performed on these three distributions. The p-value yielded by the ANOVA test determines whether the differences between group means are statistically significant. To determine whether any of the differences between the means are statistically significant, one typically compares the p-value to a chosen *significance level* to assess the null hypothesis. The null hypothesis states that the population means are all equal. Usually, a significance level (denoted as α) of 0.05 works well. A significance level of 0.05 indicates a 5% risk of concluding that a difference exists when there is no actual difference. The ANOVA test on our distribution yielded a p-value of 0.29. Since the p-value is greater than the significance level ($\alpha = 0.05$), implying that the null hypothesis cannot be rejected, and hence the WD between the three data cohorts indeed have statistically identical means. In addition to the ANOVA test, we also performed t-value tests on the three pairs—the results of which are shown in Table 1. Note that this is a more stringent test as compared to an ANOVA. As can be seen in Table 1, all the p-values are greater than 0.05, also implying that the means of WD distributions for each cohort comparison are statistically similar.

We also investigate the three distributions in isolation. For all the 316 subjects, the WD between FCNs with temporal sampling period 645 and 1400 ms, we find that 13% of subjects have a WD less than 2, 65% of subjects have a WD less than 5, and only 1% of subjects have a WD greater than 10. The mean WD for the three distributions across all 316 subjects is 4.30, 4.01, and 4.14, respectively.

		t-value	p-value
WD (TR = 645, TR = 1400)	WD (TR = 1400, TR = 2500)	.059	.088
WD (TR = 1400, TR = 2500)	WD (TR = 2500, TR = 645)	.460	.519
WD (TR = 2500, TR = 645)	WD (TR = 645, TR = 1400)	.286	.387

TABLE 1 t-values and p-values obtained by pairwise t-tests comparing the WDs between data cohorts.

Note: Since all p-values are greater than .05, the means of WD distributions for each cohort comparison are statistically similar.

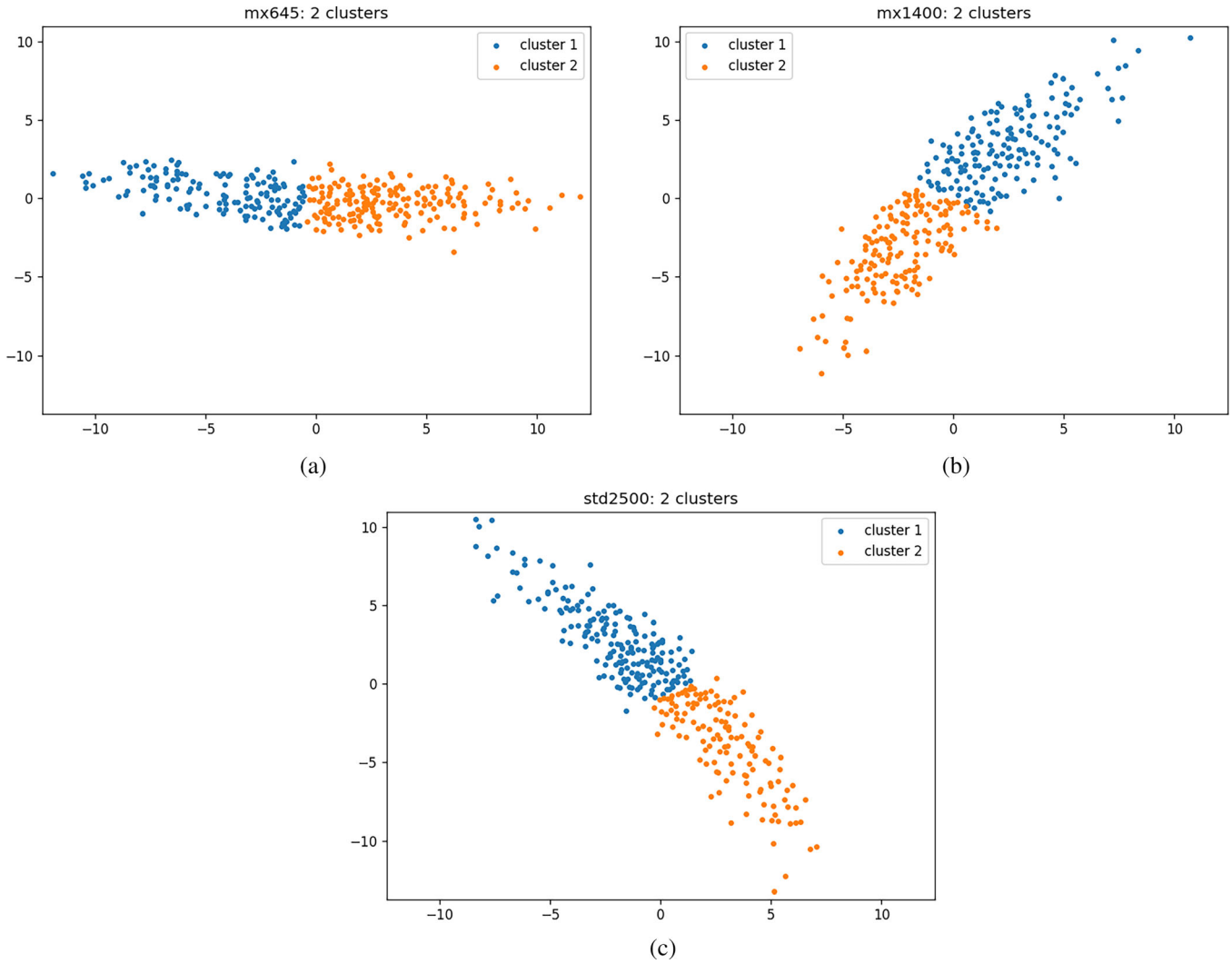


FIGURE 7 MDS plots of the three data cohorts (TR = 645, 1400, and 2500 ms) have similar structures, and all 316 subjects within the cohorts are segregated into two clusters. (a) Clustering result on MDS data for TR = 645 ms. (b) Clustering result on MDS data for TR = 1400 ms. (c) Clustering result on MDS data for TR = 2500 ms.

Next, we perform a pairwise WD computation among all 316 subjects for the three data cohorts in the second set of experiments. This yields three adjacency (A) matrices for the three temporal cohorts where a_{ij} correspond to the WD between persistent diagram of subjects i and j. It is difficult to compare adjacency matrices (both statistically and visually), and therefore we perform dimensionality reduction using MDS to reduce the matrix (316-dimensional data) to a 2D plane (2-dimensional data). We plot the MDS plots for the three adjacency matrices in Figure 7. We also applied clustering on the MDS plots to segregate the subjects, and we can see that all three cohorts result in two distinct clusters (as determined by the Silhouette criterion). When we compute the set intersection (overlaps) of subjects across the clusters of different TRs, we observe 198 out of 316 subjects reside in the same cluster group for all three TRs. This number rises to 251 subjects for pairwise cluster intersection. This intersection was statistically very significant (a p-value of 0) when compared with the null distribution obtained by adopting the same procedure using random FC matrices. This result is encouraging as it qualitatively confirms the quantitative result we obtained in the first set of experiments, further bolstering our hypothesis that PH-based techniques are robust to data acquisition parameters such as sampling period.

Finally, we perform the third set of experiments detailed in Section 3. These experiments are not based on TDA and follow the traditional analysis pipeline. We plot the result of the thresholded image in Figure 8. It can be seen from the figure that the majority of pixels are black, which corresponds to a p-value of less than 0.05. In particular, there are 7732 connections that showed significant differences in connectivity across the same subjects for data acquired from different sampling periods versus 5037 connections that did not. This result strongly indicates that the traditional analysis method fails to identify connections from the same subjects but acquired with different sampling periods as the same connections in a majority of cases.

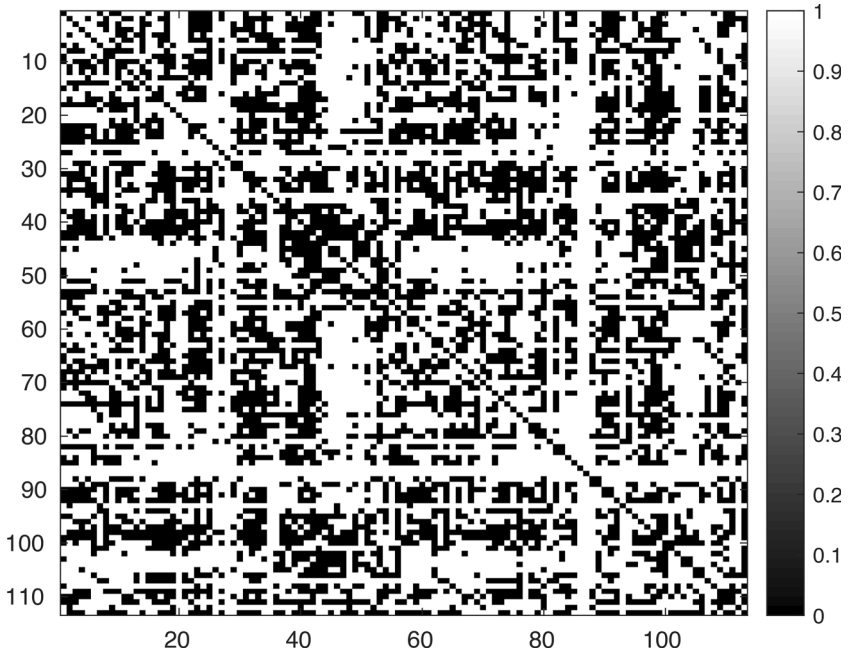


FIGURE 8 Result after ANOVA test performed for all 316 subjects for every node of the FCN. White pixels correspond to those connections which were significantly different across the same subjects for data acquired from different sampling periods.

5 | DISCUSSION

MRI scanners across the world have different configurations and field strengths.[1] Data acquired from different scanners with different parameters have some degree of noise in them due to non-neural variability introduced by different scanner configurations and data acquisition parameters, which makes it difficult for datasets acquired from different machines to be pooled into one large dataset for analysis within a single framework. As a result, conducting research on brain networks obtained from fMRI is primarily concentrated at localized sites. Such studies are limited by the fact that the number of subjects that can be scanned at a single scanner is limited, which tends to reduce the sample size and hence the generalizability of the results. This can be overcome by conducting multi-site studies with sites that are geographically closer to the population of interest. However, as mentioned above, data acquired from different scanners and parameters introduces noise, which reduces the effect of interest that is neural in origin, and hence reduces the utility of such a multi-site effort. Our paper aims to address precisely this aspect. Accordingly, we have proposed PH-based techniques (which are based on TDA) as a means to overcome noise introduced by non-neural factors such as different data acquisition parameters and extract the inherent structural topology underlying brain networks characterized by fMRI. Computational topology is known to extract underlying shapes from complex data structures. We have shown that the topology-based metric for the brain network is invariant to data acquisition parameters such as sampling period. The metric captures the underlying shape of the brain network in topological space, thus creating a common ground to facilitate multi-site data-driven analysis of fMRI datasets (acquired from different scanners).

We were able to establish the efficacy of the TDA-based pipeline by comparing it against the traditional data analysis pipeline presented in Section 3. In a traditional pipeline, raw FCNs are used directly instead of using a topological feature. The results of this pipeline are a strong indication that these traditional techniques are not able to establish similarity across FCNs of the same subjects acquired with different TRs and acquisition parameters. While the opposite was true for the TDA-based metric, wherein we showed both qualitatively and quantitatively that the metric remains statistically invariant across the same subjects irrespective of the sampling period with which resting-state fMRI data was acquired. This demonstrates the utility of TDA-based analysis because, in principle, data acquired using different parameters from the same subject should still capture the same brain network. Certain limitations of our work must be kept in mind while interpreting the results. Even though the sampling period was different across the three different acquisitions, some other parameters, such as number of volumes, multiband factor, FOV, and voxel size, also varied across the three cohorts. Ideally, we would want to investigate the limits of parameter variability that would show invariance in the TDA metric by varying only one parameter at a time. This will be part of our future work in this area. Furthermore, future work can investigate whether the TDA metric is invariant to differences in other variables, such as vendors (such as Siemens, GE, and Philips) and scanner models within a given vendor.

6 | CONCLUSION

In this paper, we have demonstrated the efficacy of TDA-based techniques in establishing a notion of similarity between FCN. FCNs of subjects acquired from different scanners or at different sampling

[1]For example, the MRI machines available for research in the state of Alabama in the United States—Siemens 7 T Magnetom and 3 T Verio at Auburn University, Siemens 3 T Prisma at the University of Alabama Birmingham, Philips 3 T Achieva at the University of South Alabama, and Siemens 3 T Prisma at the University of Alabama Tuscaloosa—all have different configurations.

rates have some degree of noise, reducing the efficacy of traditional network analysis pipelines. We have used PH to extract topological features in the form of barcodes from FCNs; these barcodes can then be used to compare and contrast different FCNs. We demonstrated that barcodes extracted from scans of the same subject but acquired with different temporal sampling periods capture the same structure. We presented two analysis pipelines for TDA-based analysis. In the first pipeline, we compared scans of the same subject acquired at different sampling rates, and in the second pipeline, we compared scans of all subjects with the same sampling period. Both our analysis pipelines demonstrated the efficacy of PH in performing analysis of FCNs that are acquired with differing parameters (such as sampling periods). Finally, we have open-sourced all our code, script, data, and documentation at https://github.com/harp-lab/brainPH, making this work reproducible.

ACKNOWLEDGMENTS

We thank Xi-Nian Zuo from the Beijing Normal University for sharing some of the data with us. We are also grateful to the entire Enhanced Nathan Kline Institute—Rockland Sample (NKI-RS) team for providing free access to the raw MRI data. Principal support for the enhanced NKI-RS project was provided by the NIMH BRAINS R01MH094639-01 (PI Milham). Funding for key personnel was also provided in part by the New York State Office of Mental Health and Research Foundation for Mental Hygiene. Funding for the decompression and augmentation of administrative and phenotypic protocols was provided by a grant from the Child Mind Institute (1FDN2012-1). Additional personnel support was provided by the Center for the Developing Brain at the Child Mind Institute, as well as NIMH R01MH081218, R01MH083246, and R21MH084126. Project support was also provided by the NKI Center for Advanced Brain Imaging (CABI), the Brain Research Foundation, and the Stavros Niarchos Foundation.

DATA AVAILABILITY STATEMENT

The code, script, data, and documentation has been publicly shared on GitHub and can be accessed using the following link: https://github.com/harp-lab/brainPH.

ORCID

Gopikrishna Deshpande https://orcid.org/0000-0001-7471-5357

REFERENCES

Abraham, A., Milham, M. P., Di Martino, A., Craddock, R. C., Samaras, D., Thirion, B., & Varoquaux, G. (2017). Deriving reproducible biomarkers from multi-site resting-state data: An autism-based example. *NeuroImage*, *147*, 736–745.

Andellini, M., Cannatà, V., Gazzellini, S., Bernardi, B., & Napolitano, A. (2015). Test-retest reliability of graph metrics of resting state MRI functional brain networks: A review. *Journal of Neuroscience Methods*, *253*, 183–192.

Aurich, N. K., Alves Filho, J. O., Marques da Silva, A. M., & Franco, A. R. (2015). Evaluating the reliability of different preprocessing steps to estimate graph theoretical measures in resting state fmri data. *Frontiers in Neuroscience*, *9*, 48.

Bell, T. K., Godfrey, K. J., Ware, A. L., Yeates, K. O., & Harris, A. D. (2022). Harmonization of multi-site mrs data with combat. *NeuroImage*, *257*, 119330.

Bellec, P., Chu, C., Chouinard-Decorte, F., Benhajali, Y., Margulies, D. S., & Craddock, R. C. (2017). The neuro bureau adhd-200 preprocessed repository. *NeuroImage*, *144*, 275–286.

Biswal, B. B., Mennes, M., Zuo, X.-N., Gohel, S., Kelly, C., Smith, S. M., Beckmann, C. F., Adelstein, J. S., Buckner, R. L., Colcombe, S., Dogonowski, A. M., Ernst, M., Fair, D., Hampson, M., Hoptman, M. J., Hyde, J. S., Kiviniemi, V. J., Kötter, R., Li, S. J., … Milham, M. P. (2010). Toward discovery science of human brain function. *Proceedings of the National Academy of Sciences*, *107*(10), 4734–4739.

Carroll, J. D., & Arabie, P. (1998). Multidimensional scaling. In *Measurement, Judgment and Decision Making* (pp. 179–250). Elsevier.

Cassidy, B., Rae, C., & Solo, V. Brain activity: Conditional dissimilarity and persistent homology. In *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*. 2015, 1356–1359.

Cetin-Karayumak, S., Stegmayer, K., Walther, S., Szeszko, P. R., Crow, T., James, A., Keshavan, M., Kubicki, M., & Rathi, Y. (2020). Exploring the limits of combat method for multi-site diffusion MRI harmonization. *bioRxiv*. https://www.biorxiv.org/content/early/2020/11/21/2020.11.20.390120

Cetin-Karayumak, S., Zhang, F., Billah, T., Zekelman, L., Makris, N., Pieper, S., O'Donnell, L. J., & Rathi, Y. (2023). Harmonized diffusion mri data and white matter measures from the baseline adolescent brain cognitive development (ABCD) study. *bioRxiv*.

Chavez, S., Viviano, J., Zamyadi, M., Kingsley, P. B., Kochunov, P., Strother, S., & Voineskos, A. (2018). A novel DTI-QA tool: Automated metric extraction exploiting the sphericity of an agar filled phantom. *Magnetic Resonance Imaging*, *46*, 28–39.

Cox, M. A., & Cox, T. F. (2008). Multidimensional scaling. In *Handbook of data visualization* (pp. 315–347). Springer.

Dabaghian, Y., Mémoli, F., Frank, L., & Carlsson, G. (2012). A topological paradigm for hippocampal spatial map formation using persistent homology. *PLoS Computational Biology*, *8*(8), 1–14. doi:10.1371/journal.pcbi.1002581

Dansereau, C., Benhajali, Y., Risterucci, C., Pich, E. M., Orban, P., Arnold, D., & Bellec, P. (2017). Statistical power and prediction accuracy in multisite resting-state fmri connectivity. *NeuroImage*, *149*, 220–232.

Di Martino, A., Yan, C.-G., Li, Q., Denio, E., Castellanos, F. X., Alaerts, K., Anderson, J. S., Assaf, M., Bookheimer, S. Y., Dapretto, M., Deen, B., Delmonte, S., Dinstein, I., Ertl-Wagner, B., Fair, D. A., Gallagher, L., Kennedy, D. P., Keown, C. L., Keysers, C., … Milham, M. P. (2014). The autism brain imaging data exchange: Towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular Psychiatry*, *19*(6), 659–667.

Drakesmith, M., Caeyenberghs, K., Dutt, A., Lewis, G., David, A., & Jones, D. (2015). Overcoming the effects of false positives and threshold bias in graph theoretical analyses of neuroimaging data. *NeuroImage*, *118*, 313–333.

Edelsbrunner, H. (2013). Persistent homology: Theory and practice. Lawrence Berkeley National Laboratory. https://escholarship.org/uc/item/2h33d90r

Edelsbrunner, H., & Harer, J. (2008a). Surveys on discrete and computational geometry. *Contemporary Mathematics (American Mathematical Society, Providence, RI)*, *453*, 257–282.

Edelsbrunner, H., & Harer, J. (2008b). Persistent homology—a survey. *Contemporary Mathematics*, *453*(26), 257–282.

Edelsbrunner, H., & Harer, J. (2022). *Computational topology: An introduction*. American Mathematical Society.

Feinberg, D. A., Moeller, S., Smith, S. M., Auerbach, E., Ramanna, S., Glasser, M. F., Miller, K. L., Ugurbil, K., & Yacoub, E. (2010). Multiplexed echo planar imaging for sub-second whole brain fmri and fast diffusion imaging. *PLoS One*, *5*(12), e15710.

Feis, R. A., Smith, S. M., Filippini, N., Douaud, G., Dopper, E. G., Heise, V., Trachtenberg, A. J., van Swieten, J. C., van Buchem, M. A., Rombouts, S. A., & Mackay, C. E. (2015). ICA-based artifact removal diminishes scan site differences in multi-center resting-state fMRI. *Frontiers in Neuroscience*, *9*, 395.

Fornito, A., & Bullmore, E. T. (2010). What can spontaneous fluctuations of the blood oxygenation-level-dependent signal tell us about psychiatric disorders? *Current Opinion in Psychiatry*, *23*(3), 239–249.

Fortin, J.-P., Cullen, N., Sheline, Y. I., Taylor, W. D., Aselcioglu, I., Cook, P. A., Adams, P., Cooper, C., Fava, M., McGrath, P. J., McInnis, M., Phillips, M. L., Trivedi, M. H., Weissman, M. M., & Shinohara, R. T. (2018). Harmonization of cortical thickness measurements across scanners and sites. *NeuroImage*, *167*, 104–120.

Fortin, J.-P., Parker, D., Tunç, B., Watanabe, T., Elliott, M. A., Ruparel, K., Roalf, D. R., Satterthwaite, T. D., Gur, R. C., Gur, R. E., Schultz, R. T., Verma, R., & Shinohara, R. T. (2017). Harmonization of multi-site diffusion tensor imaging data. *NeuroImage*, *161*, 149–170.

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.

Ghrist, R. (2008). Barcodes: The persistent topology of data. *Bulletin of the American Mathematical Society*, *45*(1), 61–75.

Ghrist, R. W. (2014). *Elementary applied topology* (Vol. 1). Createspace Seattle.

Ginestet, C. E., Fournel, A. P., & Simmons, A. (2014). Statistical network analysis for functional MRI: Summary networks and group comparisons. *Frontiers in Computational Neuroscience*, *8*, 51.

Ingalhalikar, M., Shinde, S., Karmarkar, A., Rajan, A., Rangaprakash, D., & Deshpande, G. (2021). Functional connectivity-based prediction of autism on site harmonized ABIDE dataset. *IEEE Transactions on Biomedical Engineering*, *68*(12), 3628–3637.

Kerber, M., Morozov, D., & Nigmetov, A. (2017). Geometry helps to compare persistence diagrams. *Journal of Experimental Algorithmics (JEA)*, *22*, 1–20.

Khalili-Mahani, N., Rombouts, S. A., van Osch, M. J., Duff, E. P., Carbonell, F., Nickerson, L. D., Becerra, L., Dahan, A., Evans, A. C., Soucy, J.-P., Wise, R., Zijdenbos, A. P., & van Gerven, J. M. (2017). Biomarkers, designs, and interpretations of resting-state fMRI in translational pharmacological research: A review of state-of-the-art, challenges, and opportunities for studying brain chemistry. *Human Brain Mapping*, *38*(4), 2276–2325.

Lee, H., Chung, M. K., Kang, H., Kim, B., & Lee, D. S. (2011). Discriminative persistent homology of brain networks. In IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 841–844.

Lee, H., Kang, H., Chung, M. K., Kim, B., & Lee, D. S. (2012). Persistent brain network homology from the perspective of dendrogram. *IEEE Transactions on Medical Imaging*, *31*(12), 2267–2277.

Miller, K. L., Alfaro-Almagro, F., Bangerter, N. K., Thomas, D. L., Yacoub, E., Xu, J., Bartsch, A. J., Jbabdi, S., Sotiropoulos, S. N., Andersson, J. L., Griffanti, L., Douaud, G., Okell, T. W., Weale, P., Dragonu, I., Garratt, S., Hudson, S., Collins, R., Jenkinson, M., … Smith, S. M. (2016). Multimodal population brain imaging in the UK biobank prospective epidemiological study. *Nature Neuroscience*, *19*(11), 1523–1536.

Narayan, M., & Allen, G. I. (2015). Population inference for node level differences in multi-subject functional connectivity. International Workshop on Pattern Recognition in NeuroImaging, IEEE, 53–56.

Noble, S., Scheinost, D., Finn, E. S., Shen, X., Papademetris, X., McEwen, S. C., Bearden, C. E., Addington, J., Goodyear, B., Cadenhead, K. S., Mirzakhanian, H., Cornblatt, B. A., Olvet, D. M., Mathalon, D. H., McGlashan, T. H., Perkins, D. O., Belger, A., Seidman, L. J., Thermenos, H., … Constable, R. T. (2017). Multisite reliability of mr-based functional connectivity. *NeuroImage*, *146*, 959–970.

Nooner, K. B., Colcombe, S., Tobe, R., Mennes, M., Benedict, M., Moreno, A., Panek, L., Brown, S., Zavitz, S., Li, Q., Sikka, S., Gutman, D., Bangaru, S., Schlachter, R. T., Kamiel, S. M., Anwar, A. R., Hinz, C. M., Kaplan, M. S., Rachlin, A. B., … Milham, M. P. (2012). The NKI-Rockland sample: A model for accelerating the pace of discovery science in psychiatry. *Frontiers in Neuroscience*, *6*, 152.

Oudot, S. Y. (2015). *Persistence theory: From quiver representations to data analysis* (Vol. 209). American Mathematical Society Providence.

Patania, A., Vaccarino, F., & Petri, G. (2017). Topological analysis of data. *EPJ Data Science*, *6*, 1–6.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Roffet, F., Delrieux, C., & Patow, G. (2022). Assessing multi-site rs-fMRI-based connectomic harmonization using information theory. *Brain Sciences*, *12*(9), 1219.

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, *20*, 53–65.

Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage*, *52*(3), 1059–1069.

Saggar, M., Sporns, O., Gonzalez-Castillo, J., Bandettini, P. A., Carlsson, G., Glover, G., & Reiss, A. L. (2018). Towards a new approach to reveal dynamical organization of the brain using topological data analysis. *Nature Communications*, *9*(1), 1399. https://doi.org/10.1038/s41467-018-03664-4

Schnack, H. G., & Kahn, R. S. (2016). Detecting neuroimaging biomarkers for psychiatric disorders: Sample size matters. *Frontiers in Psychiatry*, *7*, 50.

Shinohara, R. T., Oh, J., Nair, G., Calabresi, P. A., Davatzikos, C., Doshi, J., Henry, R. G., Kim, G., Linn, K. A., Papinutto, N., Pelletier, D., Pham, D. L., Reich, D. S., Rooney, W., Roy, S., Stern, W., Tummala, S., Yousuf, F., Zhu, A., … the NAIMS Cooperative. (2017). Volumetric analysis from a harmonized multisite brain MRI study of a single subject with multiple sclerosis. *American Journal of Neuroradiology*, *38*(8), 1501–1509.

Simpson, S. L., Bowman, F. D., & Laurienti, P. J. (2013). Analyzing complex functional brain networks: Fusing statistics and network science to understand the brain. *Statistics Surveys*, *7*, 1–36.

Smith, S. M., Miller, K. L., Salimi-Khorshidi, G., Webster, M., Beckmann, C. F., Nichols, T. E., Ramsey, J. D., & Woolrich, M. W. (2011). Network modelling methods for fMRI. *NeuroImage*, *54*(2), 875–891.

Termenon, M., Jaillard, A., Delon-Martin, C., & Achard, S. (2016). Reliability of graph analysis of resting state fMRI using test-retest dataset from the human connectome project. *NeuroImage*, *142*, 172–187.

The GUDHI Project. (2015). *GUDHI user and reference manual*. GUDHI Editorial Board. http://gudhi.gforge.inria.fr/doc/latest/

Thomas Yeo, B., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zöllei, L., Polimeni, J. R., Fischl, B., Liu, H., & Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, *106*(3), 1125–1165.

Vallender, S. (1974). Calculation of the Wasserstein distance between probability distributions on the line. *Theory of Probability & its Applications*, *18*(4), 784–786.

Weinberger, S. (2011). What is … Persistent homology. *Notices of the AMS*, *58*(1), 36–39.

Williamson, D. F., Parker, R. A., & Kendrick, J. S. (1989). The box plot: A simple visual method to interpret data. *Annals of Internal Medicine*, *110*(11), 916–921.

Yu, M., Linn, K. A., Cook, P. A., Phillips, M. L., McInnis, M., Fava, M., Trivedi, M. H., Weissman, M. M., Shinohara, R. T., & Sheline, Y. I. (2018). Statistical harmonization corrects site effects in functional connectivity measurements from multi-site fMRI data. *Human Brain Mapping*, *39*(11), 4213–4227.

Zomorodian, A. J. (2005). *Topology for computing* (Vol. 16). Cambridge University Press.

How to cite this article: Kumar, S., Shovon, A. R., & Deshpande, G. (2023). The robustness of persistent homology of brain networks to data acquisition-related non-neural variability in resting state fMRI. *Human Brain Mapping*, 1–15. https://doi.org/10.1002/hbm.26403