Multiscale Modes of Functional Brain Connectivity

Authors

S Rezvan Farahibozorg,[1]* Samuel J Harrison,[1] Janine D Bijsterbosch,[2] Mark W Woolrich,[1,3] Stephen M Smith,[1]

Affiliations

[1]FMRIB, Oxford Centre for Integrative Neuroimaging, Nuffield Dept. of Clinical Neuroscience, Oxford University, Oxford, UK.

[2]Department of Radiology, Washington University in St Louis, St. Louis, USA.

[3]OHBA, Oxford Centre for Integrative Neuroimaging, Department of Psychiatry, Oxford University, Oxford, UK.

*Correspondence to: rezvan.farahibozorg@ndcn.ox.ac.uk

Abstract

Information processing in the brain spans from localised sensorimotor processes to higher-level cognition that integrates across multiple regions. Interactions between and within these subsystems enable multiscale information processing. Despite this multiscale characteristic, functional brain connectivity is often either estimated based on 10-30 distributed modes or parcellations with 100-1000 localised parcels, both missing *across*-scale functional interactions. We present Multiscale Probabilistic Functional Modes (mPFMs), a new mapping which comprises modes over various scales of granularity, thus enabling direct estimation of functional connectivity within- and across-scales. Crucially, mPFMs were not formulated, but emerged from data-driven multilevel Bayesian modelling of large functional MRI (fMRI) populations and every individual. We demonstrate that mPFMs capture both distributed brain modes and their co-existing subcomponents. In addition to validating mPFMs using simulations and real data, we show that mPFMs can predict ~900 personalised traits from UK Biobank more accurately than current standard techniques. Therefore, mPFMs can offer a new basis for functional connectivity modelling and yield enhanced fMRI biomarkers for traits and diseases.

1 Introduction

Within the complex system of 86 billion neurons (Herculano-Houzel, 2009) in the human brain, large ensembles of neurons work in synchrony, such that they produce distinct modes of extended, correlated activity (Smith et al., 2009). These functional modes exhibit a spectrum of scales and functionalities, from localised brain modes facilitating specialised information processing, such as responses to sensory or motor stimuli, to higher-level cognitive modes integrating across multiple regions through long-range connections. Functional modes underlie spontaneous brain activity during task-free periods (often referred to as resting state networks, RSNs), adapt dynamically during cognitive tasks, and undergo permanent changes in neurodegenerative diseases (Biswal et al., 1995; Buckner and Vincent, 2007; Calhoun et al., 2008; Damoiseaux et al., 2006; Raichle et al., 2001; Smith et al., 2013b). Growing evidence, especially from new datasets with thousands of individuals, has highlighted that spatiotemporal characteristics of functional modes vary systematically across individuals, providing biomarkers for traits and disease, akin to a diagnostic assay for the brain (Elliott et al., 2018; Finn et al., 2015; Jiang et al., 2020; Kong et al., 2019; Smith et al., 2015; Vidaurre et al., 2017). In this paper, we introduce a novel representation, Multiscale Probabilistic Functional Modes (mPFMs), which, unlike the existing single-scale modes, encapsulates an ensemble of modes across multiple scales. mPFMs have the potential to provide a shift in how functional connectivity has been modelled to date, by offering a new basis to gain insights into spatiotemporal connectivity within and across diverse processing scales, and providing more sensitive biomarkers from functional neuroimaging.

Functional MRI (fMRI), with a spatial resolution of ~2mm and a temporal resolution of ~1s, provides a suitable spatiotemporal resolution to characterise functional modes non-invasively. This has largely been done using two distinct approaches: a) high-dimensional (highD) decomposition of fMRI data into functional parcellations with 100-1000 modes (i.e., parcels), where each mode is typically *localised* to a single brain region (Bellec et al., 2010; Craddock et al., 2012; Schaefer et al., 2018); b) low dimensional (lowD) decomposition of fMRI data into ~20-30 modes (Beckmann and Smith, 2004; Calhoun et al., 2001) - the conventional RSNs - each *distributed* across multiple distant brain regions, and often with some spatial overlap between the modes. The latter is often carried out using spatial Independent Component Analysis (ICA) (McKeown et al., 1998), or lowD parcellations (Yeo et al., 2011). Each of these two approaches offer distinct benefits. On the one hand, fine-grained, highD parcellations capture local details in the brain's organisation, especially details that have proved useful for cross-individual variability modelling, and prediction of traits and disease (Dadi et al., 2020; Pervaiz et al., 2020). On the other hand, lowD large-scale modes capture the more global functional configuration of the brain. Distant brain regions with intrinsic long-range connections are grouped into biologically

meaningful neural entities such as language and attention networks, allowing for easier interpretation and putative links to cognitive functions (Power et al., 2011; Smith et al., 2009; Yeo et al., 2011). Additionally, lower data dimensionality makes subsequent statistical analyses more manageable. Therefore, distributed vs localised modes each offer distinct benefits and limitations for capturing global vs local functional architecture of the brain. Crucially, however, neither of these approaches directly model inherent interactions across multiple scales of information processing in the brain.

In this context, we report the finding of mPFMs, which can bridge this gap by providing a unified representation that encompasses modes across multiple scales. mPFMs were identified through high dimensional decomposition of resting state fMRI using the sPROFUMO framework (Farahibozorg et al., 2021). sPROFUMO defines a Bayesian hierarchy with two levels of group and individuals across big data populations, where group priors are iteratively used for top-down regularisation of individuals, and evidence across individuals is accumulated and fed back to the group. Two key distinguishing aspects of the model are that: 1) sPROFUMO is subject-specific, i.e., inferring the modes for populations and individuals *simultaneously*; 2) unlike ICA or parcellation-based techniques, sPROFUMO does not impose strong constraints on spatial or temporal independence between the modes, thus yielding modes that are more flexibly interacting, spatially and temporally (Farahibozorg et al., 2021; Harrison et al., 2020).

We report that this added flexibility in modelling population variations, and in spatiotemporal connectivity between the modes, resulted in the discovery of mPFMs (Figure 1), which preserve conventional large-scale modes, and add new modes with different scales of granularity to the ensemble. Figure 1 shows direct comparisons between lowD vs highD Probabilistic Functional Modes (PFM) decomposition of rfMRI data in HCP, where a force-directed network layout is used for visualisation (Bastian et al., 2009). LowD decomposition (shown in yellow) yielded the conventional distributed-only resting state networks, hereafter referred to as large-scale RSNs. The highD decomposition yielded a mixture of modes across various scales. By spatial pairing of the modes from highD and lowD decompositions, we found a one-to-one correspondence between a set of highD modes (shown in orange) and large-scale RSNs, indicating that as we increase the dimensionality from 25 to 150, conventional RSNs are preserved, and new modes are added (shown in blue). This is a novel type of decomposition, and since it includes functional modes across multiple scales, we refer to it as Multiscale Probabilistic Functional Modes, mPFMs.

mPFMs can be considered an alternative to conventional ("single scale") resting state networks (RSNs). We report that: a) mPFMs emerge to explain the existence of multiple distinct time-courses within each large-scale mode, thus capturing voxel-to-voxel variability of temporal dynamics that cannot be captured by these distributed modes alone; b) we validate mPFMs using several methods, including synthetic data, reproducibility between Human Connectome Project

(HCP) and UK Biobank (UKB) data, and link to HCP's Multimodal Parcellation (Glasser et al., 2016); c) we further show the added benefit of cross-scale interactions of mPFMs, compared with within-scale interactions, for capturing cross-individual variability and predicting personalised traits; d) finally, we show that, compared with two commonly used techniques, spatial ICA modes and parcellations from the Schaefer Atlas (Schaefer et al., 2018), mPFMs (of matched high dimensionality) yield more accurate prediction of a wide range of imaging- and non-imaging-derived phenotypes (IDPs/nIDPs) related to cognition, cardiovascular health, brain anatomy, and task fMRI in UK Biobank.

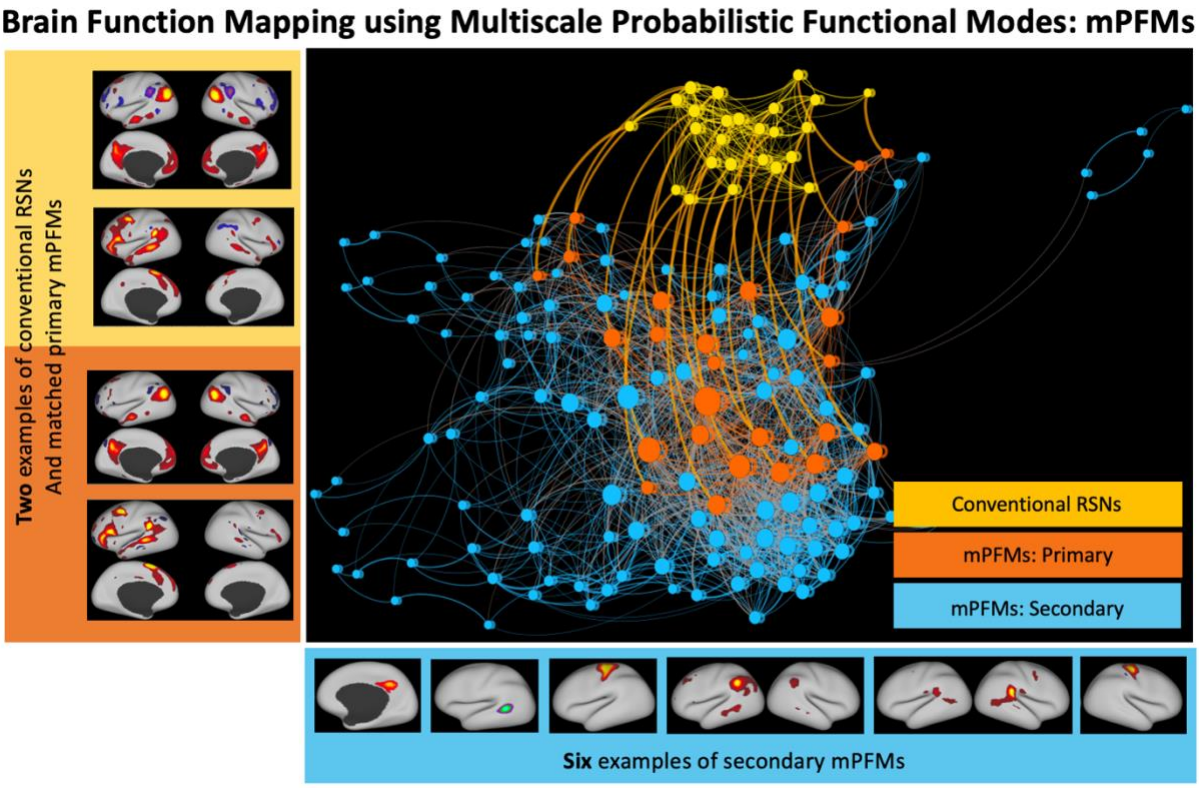

Figure 1 Brain function mapping using Multiscale Probabilistic Functional Modes (mPFMs): High-dimensional (150) Probabilistic Functional Modes estimated from resting state fMRI (1003 HCP subjects) yield a new representation of the brain's functional organisation, which consists of an ensemble of conventional distributed-only (i.e., large-scale) resting state networks (RSNs) and newly-added modes across more fine-grained scales. We refer to this new representation as Multiscale Probabilistic Functional Modes, mPFMs. When paired with large-scale RSNs from 25-mode decomposition (yellow), 25 of the mPFMs showed a clear matching, labelled as Primary mPFMs (orange). Secondary mPFMs (blue) are the remaining 125 mPFMs, that start appearing with increased dimensionality. A force-directed layout is used to visualise the spatial similarities (i.e. correlations) between the modes. In this layout, each circle denotes a functional mode, which is a node in the connectivity matrix, and each line is the connection between a pair of modes. The size of each mode shows the average strength of its connections, line widths denote the strength of spatial correlations (i.e. overlaps) between two modes, and the distance between the modes denotes how close two modes are in the connectivity matrix, both within and between low-dimensional vs high-dimensional decompositions. All connections within-large-scale and within-multiscale, as well as paired connections between large-scale and multiscale were maintained in visualisation. Matrix-based visualisation of these results is shown in Figure S 2.

2 Materials and Methods

Figure S 1 shows a flowchart of different subsections of materials and methods, and their link to the corresponding results subsections.

2.1 Data

2.1.1 Human Connectome Project (HCP)

We used minimally pre-processed fMRI data from Human Connectome Project (HCP) from S1200 data release (https://www.humanconnectome.org/study/hcp-young-adult), with acquisition and processing details outlined in (Glasser et al., 2013; Smith et al., 2013a). Resting state data from 1003 subjects aged 22-35 years were used here and included 4x15 minute runs per subject (i.e., 1 hour recording per subject). With repetition time (TR) of 0.72s, this dataset includes 4800 time points per individual. Study protocol of the HCP study was approved by the WU-Minn HCP Consortium's institutional review boards, and written informed consent was obtained from all participants prior to data collection. We used two data formats, volumetric and HCP CIFTI. CIFTI data includes grey matter voxels only (i.e., greyordinates), where cortical grey matter is surface-registered using MSMAll (Robinson et al., 2014) and subcortical grey matter is in volumetric space. This results in inherent data differences such as signal to noise ratio between cortical and subcortical regions, affecting matrix factorisation techniques in, e.g., finding an unbalanced number of cortical vs subcortical modes. To prevent such unbalances and make the results more easily comparable across CIFTI and Volumetric fMRI, we only used cortical greyordinates of CIFTI in this paper. ICA-FIX was used to remove artefacts from volumetric data before resampling to 32k_fs_LR space and projecting onto the surface. Standard HCP preprocessing includes spatial smoothing with a 2mm FWHM smoothing kernel. We applied additional smoothing to obtain 5mm FWHM (for both volumetric and CIFTI data), which improves the Signal to Noise Ratio (SNR) and was found to be useful for obtaining reliable subject-specific modes (Farahibozorg et al., 2021).

HCP data were predominantly used for analyses that aimed at examining timecourse and functional connectivity modelling using mPFMs, their reproducibility across datasets (HCP and UK Biobank), and validation of mPFMs against existing functional parcellations. Given the longer recordings per subject (1 hour vs 6 minutes), we deemed HCP more suitable for these analyses.

2.1.2 UK Biobank (UKB)

We randomly selected 4999 subjects aged > 45 years from the May 2019 release of UK Biobank data (application number 8107). The UKB has approval from the North West Multi-centre Research Ethics Committee (MREC) to obtain and disseminate data and samples from the participants (http://www.ukbiobank.ac.uk/ethics/), and these ethical regulations cover the work in our study. Written informed consent was obtained from all participants prior to data collection. Resting state fMRI data in UKB (at the time that this research was conducted) was in standard volumetric space, and included 1 recording of ~6 minute per subject. With a TR of 0.735s, data consisted of 490 time points per session. The standard UKB pipeline, which includes quality control, brain extraction, motion correction, artefact rejection using FSL-FIX, high-pass temporal filtering (sigma = 50.0s, Gaussian-weighted least-squares straight line fitting) and registration to standard MNI-2mm space was used for preprocessing (Alfaro-Almagro et al., 2018). Standard UKB preprocessing includes spatial smoothing with a 2mm FWHM smoothing kernel. We applied additional smoothing to obtain 5mm FWHM, which improves SNR and is useful for obtaining reliable subject-specific high dimensional PFMs. UKB was used for evaluating reproducibility of mPFMs (i.e. high-dimensional PFMs) across datasets, and their prediction accuracy for personalised traits. Given the wide range of imaging derived phenotypes (IDPs) and non-imaging derived phenotypes (nIDPs) in UK Biobank, this dataset was deemed more suitable for prediction of traits (see section 2.6 for details).

2.2 Estimating Resting State Networks

2.2.1 Probabilistic Functional Modes (PFMs)

Details of PROFUMO/sPROFUMO models and inference are elaborated in our previous work (Farahibozorg et al., 2021; Harrison et al., 2020). Here we provide a brief summary of the model and its application to the data. sPROFUMO is a matrix factorisation model for big fMRI data, which uses hierarchical Bayes with two levels of modelling and inference: population and individual.

At the subject level, fMRI timeseries ($\mathbf{D}^{sr}$) are decomposed into a set of spatial maps ($\mathbf{P}^{s}$), time courses ($\mathbf{A}^{sr}$) and time course amplitudes ($\mathbf{H}^{sr}$), with residuals $\boldsymbol{\varepsilon}^{sr}$:

$$\boldsymbol{D}^{sr} = \boldsymbol{P}^{s}\boldsymbol{H}^{sr}\boldsymbol{A}^{sr} + \boldsymbol{\varepsilon}^{sr} \qquad \boldsymbol{D}^{sr} \in \mathbb{R}^{N_v \times N_t}, \boldsymbol{P}^{s} \in \mathbb{R}^{N_v \times N_m}, \boldsymbol{H}^{sr} \in \mathbb{R}^{N_m \times N_m}, \boldsymbol{A}^{sr} \in \mathbb{R}^{N_m \times N_t} \qquad (1)$$

where s denotes subject, r denotes a recording session and $\mathbf{H}^{sr}$ denotes mode amplitudes. N_v, N_m and N_t denote the number of voxels, modes and time points, respectively.

$\mathbf{P}^s$ denotes spatial mode layout across brain voxels, hereafter referred to as *spatial maps*, or just *maps*. These are modelled using a Double-Gaussian Mixture Model (DGMM), with one Gaussian component used to model signal and a second Gaussian distribution to model the background spatial noise in each voxel. To model the spatial maps hierarchically, in addition to $\mathbf{P}^s$ for each subject, a consensus set of group-level parameters is also estimated to capture both the group mean and variance. $\mathbf{A}^{sr}$ represents mode time courses per subject and scan. The time courses also model signal and noise elements separately, where the signal element is constrained by the Haemodynamic Response Function (HRF) and the noise element follows a Gaussian distribution. Connectivity between the timecourses of functional modes can have a consensus structure across individuals. To model this simultaneously for population and individuals, a second hierarchy is defined over the precision matrix $\boldsymbol{\alpha}^{sr}$ using Wishart distributions. To further allow for modelling of haemodynamic responses that govern the Blood Oxygenation Level Dependent (BOLD) signal, HRF-constrained autocorrelations are incorporated in the modelling of this functional connectivity between the modes. $\mathbf{H}^{sr}$ denotes mode amplitudes. These are modelled using a multivariate normal distribution, with a Bayesian hierarchy between the group and the subject levels. $\mathbf{H}^{sr}$ is diagonal (one value per PFM) and captures timecourse variance for each subject and recording.

To find a solution for the probabilistic model described above, sPROFUMO uses stochastic Variational Bayesian (VB) inference. This includes dividing subjects from large data into small batches, and optimising the parameters of an approximating distribution q, with the aim that it is as close as possible to the true posterior. The group model is maintained across batches and continuously updated over time, and within each iteration of each batch, it is used for top-down regularisation of subject-specific matrix factorisations, from which the posterior distributions for subject-specific spatial maps, time course correlations and amplitudes are inferred. Subsequently, posterior distributions are accumulated across individuals and fed back to the group to update the group model. The model iterates between these two levels of estimation until convergence.

We used the model's default parameters for application to both HCP and UKB datasets. The forget rate (β) was set to 0.6; this controls the degree to which global parameter updates rely on the current compared with previous batches. Delay parameter (τ) was set to 5; this controls the degree to which initial batches influence the overall inference. Batch size was set to 50 subjects, and the number of batches was tuned such that each subject will be visited 2.5 times on average (i.e., normally chosen in 2 to 3 batches). Given the different number of subjects in the two datasets, to achieve the abovementioned number of updates per subjects, the overall group model was set to be updated at least 5000 times for UKB and 4000 times for HCP. For the UK Biobank (taking 150-mode decomposition as an example), the model was run on 10 processing cores on a compute node, taking ~210 hours to complete, and the memory usage peaked at

210 ∼100GB. For the HCP (taking 150-mode decomposition as an example), the model was run on 10
211 processing cores on a compute node, taking ∼103 hours to complete, and the memory usage
212 peaked at ∼90GB.

213 When applying to each dataset, and as a part of sPROFUMO's internal initialisation, voxel-wise
214 de-meaning and variance normalisation was applied to rfMRI recordings. Additionally, to initialise
215 the model using a realistic set of initial group maps that can place the parameters within a
216 realistic ballpark, sPROFUMO internally applies an online PCA algorithm called MELODIC's
217 Incremental Group-PCA (MIGP) (Smith et al., 2014) followed by variational ICA. These initial
218 maps, together with hyperpriors, are used to initialise the group-level parameters before
219 conducting the full Bayesian inference at subject- and group-level. It is worth noting that, to make
220 the results fully comparable across HCP and UKB, the same set of initial maps were used to
221 initialise both datasets. Two features in sPROFUMO help the model escape local minima: firstly,
222 we use 10 random re-initialisations to generate the initial group-level maps; secondly, the mini-
223 batching technique in stochastic variational Bayesian inference reduce the risk of getting stuck in
224 local minima.

225 We estimated 25, 50, 100, and 150 PFMs, each used in a subset of the analyses, as elaborated in
226 Results section. Initial maps were estimated for each dimensionality independently, since the
227 number of the initialised modes will have to match the final number of the estimated modes.
228 However, it is worth noting that the probabilistic framework has the capacity to switch off some
229 of these modes in the final decomposition, if there is not enough posterior evidence in the data
230 to support specific modes in higher dimensionalities (e.g., in noisier datasets or noisier modes for
231 specific subjects). This feature, which reduces the risk of false positive modes, has been
232 investigated in detail in our previous work (Farahibozorg et al., 2021).

233 After estimating PFMs with various dimensionalities, we validated whether each mode existed at
234 both group-level and subject-level. A few artefactual modes (typically ~2-3 due to, e.g., white
235 matter artefacts) were manually removed from subsequent analyses. At subject-level, spatial and
236 temporal aspects of the modes were used for evaluation. In particular, subject-to-group similarity
237 of spatial maps is a useful metric. Very low similarities (~<0.1) would indicate highly noisy
238 estimation of a mode at subject-level. Whereas, artefactually high similarities (>0.99) would
239 indicate that posterior parameters were not estimable at subject level, and thus, the spatial map
240 for that specific subject-PFM was reverted back to group priors. An example of subject-group
241 spatial map similarity for d=150 in UK Biobank is shown in **Figure S 3**. In addition to this, we used
242 mode amplitudes (H^{sr}) and standard deviation of spatial maps across voxels to identify PFMs with
243 weak signal and subject level. PFMs included in the subsequent analyses were reliably estimated
244 at both subject- and group-level.

245

2.2.1.1 Volumetric vs CIFTI PFMs:

sPROFUMO was applied to volumetric and CIFTI data from HCP separately, and group-level spatial maps were used as the basis for computing the reproducibility scores reported in Results section 3.2.2. In order to compare volumetric and CIFTI results, volumetric PFMs were mapped onto CIFTI using linear regression at the subject level. More specifically, for each subject and run, we used volumetric PFM timeseries ($\boldsymbol{A}_{VOL}^{sr}$) and amplitudes ($\boldsymbol{H}_{VOL}^{sr}$), as well as raw CIFTI fMRI timeseries $\boldsymbol{D}_{CIFTI}^{sr}$. We computed spatial maps in CIFTI space ($\boldsymbol{P}_{CIFTI}^{s}$) through back projection of $\boldsymbol{A}_{VOL}^{sr}\boldsymbol{H}_{VOL}^{sr}$ onto $\boldsymbol{D}_{CIFTI}^{sr}$ using linear regression with L2 regularisation. Regression betas were then converted to t-stat values. We next averaged $\boldsymbol{P}_{CIFTI}^{s}$ across subjects and obtained volume-to-surface mapped group spatial maps to compare against surface based spatial maps.

2.2.2 Spatial ICA and Dual Regression

We applied the widely-used group level spatial ICA followed by dual regression (DR) to characterise subject-specific RSNs and compare against sPROFUMO results. In standard pipelines of large datasets such as HCP and UKB, ICA/ICA-DR have been used for RSN characterisation, thus providing a suitable baseline to compare sPROFUMO results against. We applied ICA to UKB data and identified 100 group-level spatial components using FSL *MELODIC,* and subsequently mapped these group-level results onto single subject data using FSL *dual_regression*. Subject-specific time courses for each mode and subject-specific spatial maps we obtained from stages 1 and 2 of Dual Regression, respectively. Mode amplitudes were computed as standard deviation of the time courses and functional connectivity was computed using Tikhonov/L2-regularised partial correlations. In order to make the final sPROFUMO and ICA/ICA-DR results fully comparable, the same PCA initialisation (using the MIGP approach) was used for both models. ICA-DR results were used to predict non-imaging phenotypes and compare PFM's performance against.

2.3 Identifying temporally distinct subcomponents within large-scale RSNs

Low-dimensional representation of functional modes yields large-scale RSNs that are distributed across multiple brain areas. Matrix factorisation techniques (e.g. sPROFUMO or ICA) describe each of these modes with a single timecourse. We tested if there are distinct temporal subcomponents within each of these modes, and whether or not these subcomponents will be better represented by lowD large-scale RSNs vs highD mPFMs. For this purpose, we focused on 25-mode decomposition from HCP CIFTI rfMRI data of 1003 subjects, and conducted the following tests:

2.3.1 Characterising distinct subcomponents within each conventional RSN

This step aimed to test whether a single prominent subcomponent can explain a majority of *group-level* temporal variance in the voxel-wise activity within each large-scale mode. For this purpose, we used PCA as follows:

a) $N_{voxel} \times N_{time}$ subject rfMRI timeseries were weighted by each subject PFM (d=25) spatial map, obtaining $N_{voxel} \times N_{time}$ weighted rfMRI timeseries (the data dimensions are unchanged, but voxels within a given map are up-weighted);
b) PCA was applied to each of these weighted timeseries, and 10 components were estimated per subject per mode, reducing the weighted subject fMRI data of each mode to $N_{voxel} \times 10$. The resulting PCA maps (for each mode separately) were concatenated across subjects, resulting in a feature space of size $N_{voxel} \times (10*N_{subject})$. A second group-level PCA was next applied, and 10 group-level components were extracted per mode. The explained variance per PC was estimated (by considering the 10 singular values); the results of this analysis are reported in 3.1.1.

2.3.2 Are temporal subcomponents best represented by mPFMs?

This step aimed to test whether multiple temporally distinct subcomponents exist within each large-scale RSN (i.e., distributed only), and whether subcomponents originating from large-scale RSNs can be better represented by large-scale (lowD) vs multiscale (highD) PFMs. This was done using temporal ICA, and further replicated using spatial ICA and PCA. This step, as well as the subsequent steps in 2.3.3 focused on *single-subject* modes.

a) Similar to step 2.3.1, $N_{voxel} \times N_{time}$ subject rfMRI timeseries were weighted by each subject PFM (d=25) spatial map, and 5 spatial PCs were extracted per mode per subject. The choice of 5 PCs was determined based on results of step 2.3.1, where PCs >6 were found to explain less than 3% of variance for all the PFMs (Figure S 4).
b) For each PFM the 5 PC maps were concatenated across subjects and group-level PCA was used to extract 5 spatial subcomponents per PFM.
c) These group-level PCs were then linearly regressed onto each subject's PFM-weighted rfMRI timeseries (as described in step 2.3.1a), giving 5 temporal components per subject per mode, which were then concatenated across subjects to obtain 5 long timeseries per PFM.
d) Temporal ICA was applied to the resulting long timeseries to obtain timeseries for 5 temporally-independent subcomponents. These timeseries were split up into temporal chunks to obtain subject-specific tICA subcomponent timeseries.

e) Subject-specific tICA subcomponent timeseries were linearly regressed onto each subject's PFM-weighted rfMRI timeseries (as described in step 2.3.2a) to obtain subject-specific tICA subcomponent spatial maps.

f) By correlating the subject-specific timecourses of each subcomponent with all the original lowD and highD mode timeseries, we applied a winner-takes-all approach to determine how many of the subcomponents originating from large-scale lowD modes were best-represented by lowD vs primary vs secondary mPFMs. An unpaired t-test was used to conduct statistical comparisons, and results were Bonferroni-corrected for multiple comparisons.

2.3.3 Do secondary mPFMs emerge out of areas of spatial overlap and/or temporal co-activation?

1. This step aimed to test whether subcomponents identified in step 2.3.2 were exclusive to one lowD mode, or if they resided in the *temporal space* shared by multiple lowD modes:
    a. Similar to step 2.3.2a, $N_{voxel} \times N_{time}$ subject fMRI timeseries were weighted by each subject PFM (d=25) spatial map, obtaining 25 weighted fMRI datasets per subject.
    b. Linear regression was used to regress out the timecourses of every other (24) PFM from each PFM's weighted fMRI. Then 5 PCs were extracted per mode per subject.
    c. Steps 2.3.2b-f were repeated to test if subcomponents originating from the temporally-exclusive space of lowD modes were best-represented by lowD vs primary vs secondary mPFMs.
2. This step aimed to test whether subcomponents identified in step 2.3.2 were exclusive to a single lowD mode, or if they resided in the *spatial* space shared by multiple lowD PFMs:
    a. Linear regression was used to regress out the spatial maps of every other (24) PFM from each PFM's spatial map.
    b. $N_{voxel} \times N_{time}$ subject fMRI timeseries were weighted by the resulting subject PFM (d=25) spatial map after regression, obtaining 25 weighted fMRI datasets per subject.
    c. Linear regression was used to regress the spatial maps of every other (24) PFM from each PFM's weighted fMRI. Then 5 PCs were extracted per mode per subject.
    d. Steps 2.3.2b-f were repeated to test if subcomponents originating from the spatially-exclusive (non-overlapping) space occupied by lowD modes were best-represented by lowD vs primary vs secondary mPFMs.

2.4 Simulations

Based on real data (Methods section 2.3 and Results section 3.1), we found that Secondary mPFMs were likely due to the existence of temporally-distinct subcomponents within the Primary mPFMs, i.e., the ensemble would capture integrated-segregated modes of brain connectivity. Here we simulated datasets where overlapping distributed and localised RSNs exist in the brain, such that localised RSNs are spatial sub-nodes of distributed RSNs but have distinct timecourses. Using these simulations, we tested how well PFM decomposition can recover the co-existence of global-local RSNs in the brain, and compared its performance to spatial ICA as the standard technique. Details of simulations were as follows:

- Data: 10 resting state fMRI datasets were simulated, each consisting of 50 subjects and 2 runs per subject.
- 12 resting state modes were simulated in each dataset, and they were created from scratch in each dataset.
- 6 of these modes were distributed, covering >=2 brain areas, 6 of these modes were localised, covering 1 brain area. Distributed modes were allowed to spatially overlap with each other, such that on average, 1.3 modes included a given voxel.
- Localised modes were simulated such that they were primarily a spatial sub-node of one distributed mode, while spatially overlapping with two or multiple distributed modes.
- Data properties per subject and run: 10,000 voxels, 300 time points at a TR of 0.72s. Data was created as detailed below:

We defined a set of spatial maps $\mathbf{P_g}$ at the group level. Each mode consisted of one or more randomly-selected contiguous blocks of voxels (i.e., parcels). Signal weights (per voxel per mode) were drawn from a Gamma distribution. Subject-specific spatial maps, $\mathbf{P_s}$, were defined based on the group maps by adding background Gaussian noise and applying spatial warps. Subject maps were generated to be spatially misaligned in reference to the group, with any a given subject mode having on average 83% overlap with the group-average (i.e., 17% spatial misalignment). Time courses were generated independently for each session, subject and mode to mimic the unconstrained resting state data (but following the between-mode temporal correlation structure described next). We assigned a hierarchical link between the group and subjects' temporal correlation matrices, following a Wishart distribution. This was done to obtain a consistent functional connectivity pattern between subjects and the group. Time courses were initially simulated as semi-Gaussian neural time courses with amplified frequencies < 0.1Hz, and subsequently convolved with a random draw from the FLOBS haemodynamic basis functions (Woolrich et al., 2004) to mimic realistic BOLD signal. Finally, data were created using outer product of subject mode spatial maps and timecourses, and random noise was added to the outer

product to create space-time data matrices. More details of simulation parameters are available in (Harrison et al., 2020).

These simulated modes were then estimated using PFM decomposition and ICA-Dual Regression, and results were compared with the ground truth. We initialised both models based on the same set of spatial bases from MIGP, to ensure that the observed differences were not due to initialisation of the probabilistic models.

2.5 mPFMs vs HCP-MMP

We applied sPROFUMO to cortical rfMRI data of 446 HCP subjects (matching the number of subjects for which HCP-MMP was available) and estimated 360 PFMs. The two decompositions were compared at two levels: First, we computed group correspondence between mPFMs and HCP-MMP. We binarised group-level mPFMs' spatial maps by hard thresholding at 0.3 and computed their dice similarity to group-level parcels in HCP-MMP. Second, we examined subject-level correspondence by binarising subject-level mPFMs' spatial maps by hard thresholding, and finding dice similarity between binarised mPFMs and subject-level parcels in HCP-MMP. The modes were paired based on group-level matching, and this ordering was used in both group-level and subject-level visualisations.

2.6 Out-of-sample prediction pipeline

We investigated the link between the spatial/temporal properties of mPFMs and a range of phenotypes in UKB and HCP. This was done using out-of-sample predictions in UKB and Canonical Correlation Analysis (CCA) in HCP. These accuracies were used to compare mPFMs vs conventional large-scale RSNs, primary vs secondary mPFMs, and mPFMs vs spatial ICA and Schaefer parcellation of same dimensionality. Depending on the question at hand in each subsection, we conducted predictions based on all or some of the following features: a) Mode spatial maps, b) Mode temporal network matrices (TNETs or temporal NetMats; partial correlation matrix between mode timecourses), c) Mode spatial NetMats (SNETS, the full spatial correlation matrix between mode maps).

2.6.1 Imaging and non-imaging derived phenotypes (IDPs and nIDPs)

We used 893 Imaging-derived and non-Imaging-Derived phenotypes (IDPs and nIDPs) from UK Biobank as targets in the prediction pipeline. These included IDPs related to Grey Matter, White Matter and task fMRI, as well as nIDPs related to cognitive scores and Cardiovascular health metrics. A brief summary of these phenotypes is included in this subsection, and a full list is included in Table S 7-S 13.

411 **Grey Matter**: 148 IDPs related to GM area and 148 IDPs related to GM thickness, generated with
412 Freesurfer by parcellation of the white surface using Destrieux (a2009s) parcellation. More
413 details about these IDPs (e.g. histograms) can be found on the UKB website:
414 https://biobank.ctsu.ox.ac.uk/crystal/label.cgi?id=197.

415 **White matter**: 3 IDPs extracted based on the volume of white matter hyperintensities using
416 BIANCA. 162 IDPs extracted based on mean intensity for dtifit outputs (FA, MD, MO3) and NODDI
417 outputs (ICVF, OD, ISOVF) for the 27 tracts segmented with the probabilistic tractography
418 analysis. 288 IDPs extracted based on mean intensity for dtifit outputs (FA, MD, MO) and NODDI
419 outputs (ICVF, OD, ISOVF) for the 48 tracts segmented with the TBSS-like analysis.

420 **Task fMRI**: UKB includes data for one task, which involves brain responses to images of
421 emotionally-valenced faces and shapes. 1 IDP, called 'contrast' in the results section, was
422 estimated as 90th percentile (z-statistic) of activity for response to faces vs shapes
423 (https://biobank.ctsu.ox.ac.uk/crystal/field.cgi?id=25766).

424 **Cognitive**: Starting from 1172 nIDPs related to cognitive tests, we narrowed these down to 68
425 using two criteria. Firstly, one of the authors prefiltered the outcome measures manually, by only
426 including active measures that were indicative of subject's performance. For example, among
427 outputs directly related to "Reaction Time", which entails viewing two cards (A and B) and
428 pressing buttons when identical, "Mean time to correctly identify matches", was included
429 whereas "Index for card A in round" was excluded. Secondly, we only selected tests that had a
430 non-NaN value in at least 25% of the subjects. The final list of 68 tests belonged to the following
431 categories: *Reaction time, Trail making, Matrix pattern completion, Numeric memory,*
432 *Prospective memory, Pairs matching, Symbol digit substitution* and *Fluid intelligence*. More
433 details about these nIDPs are available here:
434 https://biobank.ctsu.ox.ac.uk/crystal/label.cgi?id=100026.

435 **Cardiovascular health metrics**: Starting from 992 nIDPs related to cardiovascular health, we
436 narrowed these down to 77 using two criteria. Firstly, one of the authors prefiltered the outcome
437 measures manually, by excluding metrics that were not directly related to health. For example,
438 "LV stroke volume (2.0)" and "Systolic blood pressure, manual reading (0.0)" were included
439 whereas "Completion status of test (0.0)" and "Program category (0.0)" were excluded. Secondly,
440 we only selected tests that had a non-NaN value in at least 25% of the subjects. More details
441 about these nIDPs are available in the following links:
442 https://biobank.ctsu.ox.ac.uk/crystal/label.cgi?id=100011,
443 https://biobank.ctsu.ox.ac.uk/crystal/label.cgi?id=104 and
444 https://biobank.ctsu.ox.ac.uk/crystal/label.cgi?id=100012

2.6.2 Confound removal in UK Biobank

When conducting brain-behaviour analyses in big data, e.g., association or out-of-sample predictions, imaging confounds can significantly distort the interpretability of the results (Alfaro-Almagro et al., 2020; Snoek et al., 2019). For example, if a common confounding factor such as head size or head motion is correlated with both brain features and behavioural targets, this is likely to inflate our prediction accuracies (Snoek et al., 2019). Therefore, we opted to de-confound the data before conducting any brain-behaviour analyses. This was done by linearly regressing out the confounds from both imaging and non-imaging variables.

For UK Biobank, we started with a comprehensive set of 602 confounds proposed by Alfaro-Almagro et al. (2020). We reduced this set by: a) selecting conventional confounds including age, age squared, sex, age × sex, site, head size and head motion; b) reducing the remaining confounds using Principal Component Analysis. The top PCs that explained >85% of the variance were kept. These two steps yielded 82 variables. Importantly, we applied deconfounding taking into account cross-validation folds in order to avoid leakage of information from test to train data, as proposed by (Snoek et al., 2019).

2.6.3 Elastic-Net prediction and cross-validation

We used Python 3.6.5 and scikit-learn 0.19.1 (Pedregosa et al., 2011) to set up the prediction pipeline. We used ElasticNet regression and nested 5-fold cross validations to conduct out-of-sample predictions, with 20/80 test/train ratio. Next, we Gaussianised each predictor and target variable across subjects using quantile transformation (*QuantileTransformer*). In each cross-validation loop, the training set was used to compute the quantile transforms which were then applied to both train and test data. Similarly, deconfounding was also conducted within each cross-validation loop, where training set was used to estimate the regression parameters (or "betas"), which were then applied to de-confound both train and test data. Finally, we used *ElasticNetCV* to predict target variables in the test set. Hyperparameters related to ElasticNet; i.e., ratio of Lasso to Lasso+Ridge regularisation (*L1 ratio* varied between [0.1,0.5,0.7,0.9,0.95,0.99,1.0] with 10 alphas per *L1 ratio*) we optimised using nested cross-validations within the training set. Finally, we computed correlations between estimated and actual values of the target variables across subjects in the test set, which will be reported as prediction accuracies.

2.6.4 Feature spaces: large-scale vs multiscale PFMs

The comparison between conventional large-scale RSNs and mPFMs aimed to determine any added benefit of the latter, and additionally to determine the added benefit of having both types

of primary and secondary modes in a decomposition, as opposed to one type only. For this purpose, we focussed on interactions between the modes, i.e., spatial/temporal correlations, and their prediction power for phenotypes in UKB. Specifically, the comparison was made between prediction power of correlations in three scenarios: large-scale RSNs vs mPFMs; between primary and secondary mPFMs vs within primary mPFMs; between primary and secondary mPFMs vs within secondary mPFMs.

To summarise mode interactions, two types of feature matrices were used, Spatial NetMats (SNETs) and partial temporal NetMats (TNETs). TNETs were computed based on precision matrices and with a Tikhonov regularisation parameter 0.01. SNETs and TNETs are of size $N_{mode} \times N_{mode}$ for each subject in each PFM decomposition.

Based on highD decomposition (d=100), we divided each of the NetMats into three sub-matrices

1. $SNET_{pri2pri}$: $N_{pri} \times N_{pri}$ and $TNET_{pri2pri}$: $N_{pri} \times N_{pri}$, with $N_{pri} = 25$
2. $SNET_{sec2sec}$: $N_{sec} \times N_{sec}$ and $TNET_{sec2sec}$: $N_{sec} \times N_{sec}$, with $N_{sec} = 75$
3. $SNET_{pri2sec}$: $N_{pri} \times N_{sec}$ and $TNET_{pri2sec}$: $N_{pri} \times N_{sec}$
4. Based on large-scale lowD decomposition (d=25), we obtained $SNET_{low_dim}$: $N_{lowD} \times N_{lowD}$ and $TNET_{low_dim}$: $N_{lowD} \times N_{lowD}$

We applied the Fisher r-to-Z transformation, and unwrapped these matrices into 1D vectors. Since Pri2Pri, Sec2Sec and LowD2LowD matrices (but not Pri2Sec) are symmetric, only the upper diagonal elements were kept before unwrapping. Next, we used SVD to dimension-reduce each of these eight $N_{subject} \times N_{edges}$ matrices to feature matrices of size $N_{subject} \times 200$, which were fed into ElasticNet. The dimensionality reduction was done to obtain the same number of features to use in predictions, reducing the possibility of one mode type outperforming due to the number of features. Five-fold cross-validations were performed and accuracies computed as detailed in section 2.6.3. Cross-validations were repeated k times for each phenotype category, k being: GM: 20, WM: 20, task fMRI-condition: 200 (since there is only 1 phenotype in this category), task fMRI-contrast: 200, cognitive: 20, cardiovascular metrics: 20.

For each IDP/nIDP category, accuracies were computed for each phenotype, and pooled across the number of phenotypes in that category ($N_{phenotype}$) and number of repeats (k), yielding two vectors of size k* $N_{phenotype}$ to be compared between different mode types.

2.6.5 Feature spaces: mPFMs vs standard techniques:

2.6.5.1 mPFMs and spatial ICA features:

The comparison of mPFMs to highD ICA aimed to examine any added benefit of mPFMs over state-of-the-art RSN decompositions for capturing individualistic traits and phenotype predictions. For this purpose, three types of feature matrices were used in predictions, spatial

maps (SMAPs), SNETs, and TNETs. HighD decompositions with 100 RSNs from UKB data were used in these analyses. This dimensionality was optimised based on a left-out phenotype (age): we compared age prediction using 50, 100, 150, and 200 dimensional decompositions and found 100 to be optimal for both mPFMs and highD ICA. It is worth noting that d=100 was found to be optimal here for UK Biobank data. The optimal choice may vary for datasets with a smaller number of subjects, or datasets with longer recording sessions per subject. Feature spaces were estimated using a pipeline similar to the Uni-mode predictions described in our previous study (Farahibozorg et al., 2021):

- For SMAPs, estimating 100 modes resulted in a matrix of $N_{subject} \times N_{voxel} \times 100$. We performed dimensionality reduction across the second dimension (per mode) using sparse dictionary learning (Mairal et al., 2010) and obtained feature matrix of size ($N_{subject} \times 500$) × 100. Next, we used the feature space of each PFM separately in phenotype predictions.
- For SNETs and TNETs, we started with matrices of size $N_{subject} \times N_{mode} \times N_{mode}$. After applying the Fisher r-to-Z transformation per subject, we used rows of correlation matrices (i.e., one per mode) for each subject without additional dimensionality reduction, thus obtaining 100 feature matrices of size $N_{subject} \times 99$ from SNETs and 100 feature matrices of size $N_{subject} \times 99$ from TNETs that were used separately for predictions. TNETs were computed using partial correlations with Tikhonov regularisation.

2.6.5.2 Schaefer Parcellation features:

The standard Schaefer parcellation in MNI space, Schaefer2018_100Parcels_17Networks_order_FSLMNI152_2mm.nii.gz, from https://github.com/ThomasYeoLab/CBIG/tree/master/stable_projects/brain_parcellation/Schaefer2018_LocalGlobal/Parcellations/MNI was applied to pre-processed fMRI timeseries of UK Biobank subjects; timecourses of voxels within each parcel were averaged to obtain one timecourse per parcel/mode. Subsequently, similar to the PFM/sICA procedure explained above, TNET features were computed from these timecourses using Fisher r-to-Z transformed partial correlations with Tikhonov regularisation. These resulted in 100 feature matrices of size $N_{subject} \times 99$ which were used in uni-mode predictions.

2.6.5.3 Summary of features from mPFM/sICA/Schaefer:

We used features of each mode separately in predictions.

On the one hand, for PFM and sICA, SMAP, SNET and TNET feature matrices were concatenated horizontally, yielding 100 features matrices of size $N_{subject} \times 799$ to use in cross-validated ElasticNet predictions, as elaborated in section 2.6. Accuracies were computed for each mode and pooled across modes to compare PFMs and ICA-DR.

On the other hand, mPFMs and spatial ICA both yield spatial and temporal features, whereas hard parcellations such as Schaefer generate TNET features only. This is due to the fact that hard parcellations yield binarised spatial topographies for the modes/parcels, with fixed hard boundaries between the parcels. Therefore, the number of features derived from the Schaefer parcellation will be lower. To balance this number across the three methods, we used cross-validated feature selection based on correlation to the target in the training set to select a subset of features from mPFM/sICA, matching the number of features (99 per mode) from the Schaefer parcellation.

For each phenotype category (e.g., cardiovascular, cognitive etc.), accuracies were computed for each mode and each phenotype, and pooled across the number of phenotypes in that category ($N_{phenotype}$) and number of modes (100), yielding three vectors of length 100* $N_{phenotype}$ to compare mPFMs with ICA-DR and the Schaefer Parcellation.

2.7 Canonical Correlation Analysis

Canonical Correlation Analysis (CCA) was used to find a single multivariate mapping between a set of PFM features and a set of non-imaging variables. Each CCA component estimates a linear combination of PFM features and a linear combination of phenotypes, such that the transformed outputs are maximally correlated. In order words, CCA components project data onto common axes of subject variability that co-vary between brain and behaviour. In previous studies of rfMRI functional connectivity using spatial ICA, CCA has identified a positive-negative axis of brain-behaviour associations in HCP data (Smith et al., 2015). Here, using HCP and focussing on the same set of phenotypes, we conduced CCA aiming to examine the behavioural relevance of mPFMs in the context of existing literature. 1001 out of 1003 subjects that were included in both fMRI and phenotype recordings were included in this analysis.

2.7.1 Non-imaging phenotypes

We used 158 phenotypes for the CCA. These were sub-selected from a wider range of phenotypes using a set of criteria described in (Smith et al., 2015), and included metrics related to cognition such as various measures of fluid intelligence, executive function, language, episodic memory, working memory, attention; metrics of emotion such as life satisfaction, friendship, loneliness; metrics of psychiatric and life function such as depression, anxiety, aggression; metrics of alertness related to sleep, five metrics of personality related to agreeableness, extraversion, neuroticism, conscientiousness and openness, metrics of physical health related to sensory-motor function, and lifestyle metrics related to substance use such as alcohol, tobacco and drug. A full list of these phenotypes is presented in Table S 14.

2.7.2 Confound removal in HCP

Similar to section 2.6.2, we de-confounded both feature and phenotype matrices before conducting CCA. For this purpose, we used 13 imaging confounds that are included regularly in HCP studies: acquisition reconstruction software version; age; age squared; sex; age × sex; sex × age squared; race, ethnicity; height; weight; a summary statistic quantifying average subject head motion during the resting-state fMRI acquisitions; the cube-root of total brain volume (including ventricles), as estimated by FreeSurfer; the cube-root of total intracranial volume, as estimated by FreeSurfer.

2.7.3 mPFM feature spaces

For CCA, we need a single feature matrix of size $N_{subject} \times N_{feature}$ estimated from brain imaging to be paired with a single phenotype matrix of size $N_{subject} \times N_{phenotype}$ from behavioural/non-imaging traits. Therefore, we required a further condensed feature space compared to what described earlier in trait prediction. We extracted feature summaries separately from SMAPs, SNETs, and TNETs, and ran CCA using each feature type independently. Feature summaries were created as follows:

- For the spatial maps (SMAPs), estimating 150 modes resulted in a $\mathbf{P}_{grand}$ matrix of ($N_{subject} \times N_{voxel}$)x 150, yielding ~35 million features per subject. To extract a few hundred features that can meaningfully capture the essence of subject SMAPs, we used unsupervised learning by applying FMRIB's Linked ICA for big data (BigFLICA) (Gong et al., 2021; Groves et al., 2011). This is an ICA framework originally proposed for multimodal data fusion, and was applied here in two steps. First, subject SMAPs were dimension-reduced across voxels using sparse dictionary learning, yielding a matrix of size ($N_{subject} \times 1000$) × 150. Next, each PFM was used as a separate "modality" within bigFLICA to obtain a feature matrix of size $N_{subject} \times 500$. Using FLICA here allows us to preserve subject-specific variations in each mPFM and efficiently summarise them across numerous modes to obtain a set of independent features to characterise the sources of population variations in mPFMs.
- For the spatial and partial temporal correlation matrices (SNET, TNET), we started with matrices of size $N_{mode} \times N_{mode}$ for each subject. After applying the Fisher r-to-Z transformation, we flattened these matrices by taking the above-diagonal elements. With 150 PFMs, this resulted in two feature matrices, one for SNET another for TNET, each of size $N_{subject} \times (150*149/2) = N_{subject} \times 11175$.

Each of these three matrices were used separately in CCA estimation, and results were combined post-hoc, as elaborated in the next subsection.

2.7.4 Conducting CCA

CCA analysis consisted of the following steps:

1. One side of CCA received mPFM feature matrices as input: SMAPs: 1001 × 500 ($N_{subject}$ × N_{FLICA}); SNETs: 1001 × 11175 ($N_{subject}$ × $N_{NET_ELEMENT}$); TNETs: 1001 × 11175 ($N_{subject}$ × $N_{NET_ELEMENT}$). Preprocessing steps similar to (Smith et al., 2015) were applied to these matrices, which included normalisation and dimensionality reduction using SVD. Each feature matrix was reduced to a matrix of size 1001 × 50. The dimension reduction helps to avoid an over-determined (rank deficient) CCA solution.
2. The other side of CCA received phenotype matrices as input: 1001 × 158. Preprocessing steps similar to (Smith et al., 2015) were applied to these matrices, which included inverse Gaussian transformation and dimensionality reduction using SVD. The phenotype matrix was reduced to a matrix of size 1001 × 50.
3. Imaging confounds were regressed out of both phenotype and PFM feature matrices, as elaborated in section 2.6.2.
4. CCA was conducted for three pair-wise comparisons: 1) PFM SMAP vs phenotypes; 2) PFM SNET vs phenotypes; 3) PFM TNET vs phenotypes. This was done using the 'canoncorr' function in Matlab.
5. CCA yields a linear transformation of PFM feature matrix (X) and phenotype feature matrix (Y) so as to maximise their correlation; i.e. Y*A=U ~ X*B=V, where U and V are the linearly-transformed versions of ICA-DR and PFM feature matrices, respectively.
6. By finding correlations between columns of U and V for the top 50 CCA components, we estimated shared variances for each pairwise comparison.
7. We finally tested how many of the CCA components were significantly correlated. For this purpose, we conducted multi-level block permutations (Winkler et al., 2015) which takes family structure of HCP data into account. In each iteration, X in step 5 was kept fixed, and rows of Y were randomly permuted (i.e., permuting subjects, while keeping family members together), and step 6 was repeated. Across 10,000 permutations, we constructed a null distribution for the correlations of the top 50 CCA components. The correlation value corresponding to the top 5% of the null distribution for the *first* CCA component was used as threshold for p-value<0.05 significance level for all the CCA components. This yields a significance threshold that is Family-Wise Error-rate (FWE) corrected for multiple comparisons (across CCA components).

3 Results

Figure 2 illustrates four examples of the well-known large-scale RSNs, estimated using ICA in HCP's standard pipelines. Underneath, we have shown one primary mPFM (the closest spatial match) and two example secondary mPFMs, matching subsystems of these RSNs. Default Modes (DMN 1&2) closely resemble the two DMNs that have been reported in the previous literature (Andrews-Hanna et al., 2010), and distinct from each other in various ways, e.g.,: DMN1 features a para-hippocampal node that is missing in DMN2; DMN1 encompasses posterior IPL and Angular Gyrus, whereas DMN2 encompasses anterior IPL; lateral temporal areas such as the anterior temporal lobes are more pronounced in DMN2; DMN1 encompasses ventromedial PFC, whereas DMN2 encompasses dorsomedial PFC. The secondary mPFMs under each DMN highlight a few of their subnodes that contribute to distinct functionalities. Similarly, the illustrated Language network has been reported in previous meta-analytic analyses of language processing in the brain (Lipkin et al., 2022), featuring a predominantly left-lateralised network that covers temporal, inferior parietal, inferior frontal and pre-motor regions. Each of the secondary mPFMs then highlight subparts of this network. Finally, one of the brain's motor networks located at the central sulcus and two of its subsystems are shown on the right hand side of the figure. In the next section, we dive into understanding what these mPFMs are, and why they appear as separate, yet overlapping, entities in high-dimensional decomposition of resting-state fMRI (rfMRI) data using sPROFUMO.

We applied sPROFUMO to rfMRI data from 1003 HCP subjects and separately to 4999 UK Biobank subjects. Results predominantly focus on a 25-mode low-dimensional (lowD) PFM decomposition and its comparison to higher-dimensional (highD, >100) PFM decompositions, to investigate the functional relevance and the added benefit of mPFMs. Table S 1 includes a summary of the terminology used in the paper.

The following subsections aim to unravel the properties of the brain function that give rise to mPFMs, validating them based on simulations and real data, as well as demonstrating the utility of mPFMs for capturing individualistic traits and prediction of cognition and health. HCP data, compared to UK Biobank, provides higher data quality and longer fMRI recordings (1 hour vs ~6 minutes) per subject. Thus, HCP data was primarily used for untangling the functional relevance of mPFMs, and how they yield a more comprehensive summary of brain activity and connectivity. UK Biobank, given its larger sample size and more extensive phenotyping, was primarily used for evaluating mPFMs' performance in out-of-sample prediction of phenotypes. Both datasets were used for reproducibility analyses.

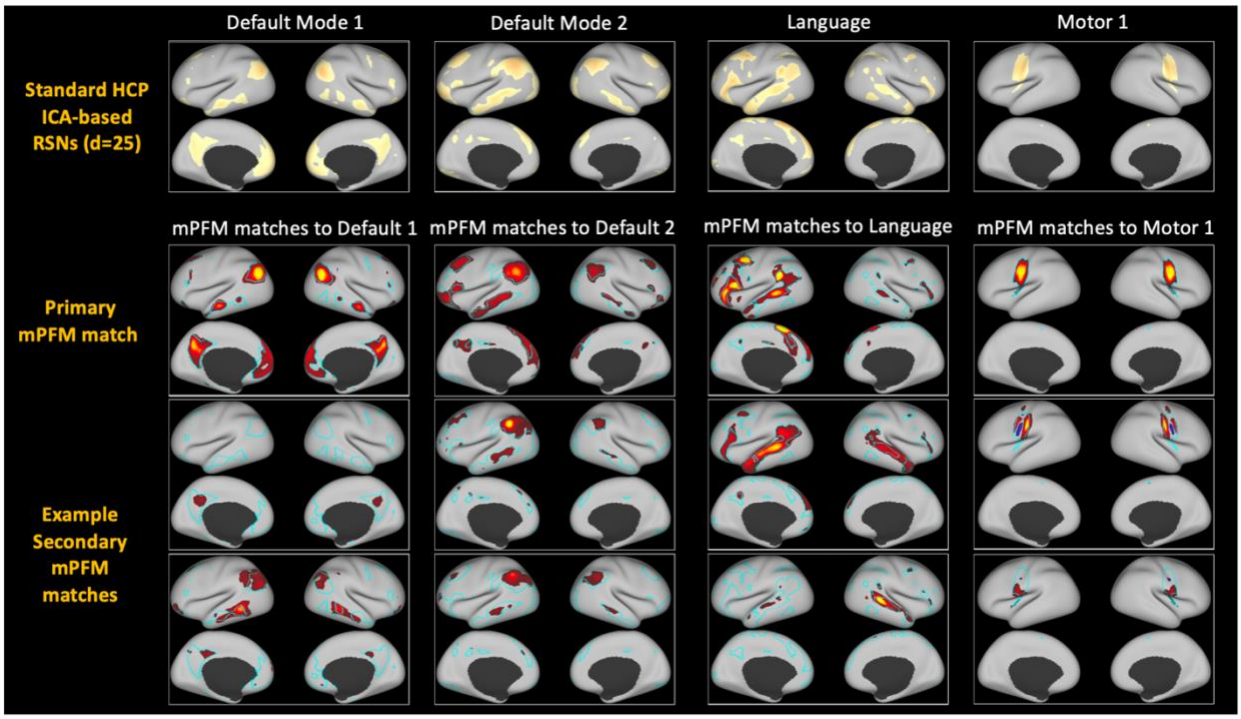

Figure 2 Four well-known RSNs and example primary and secondary mPFM matches to each. Top panel shows conventional RSNs, included in HCP's standard Resting State Networks (RSNs) releases, and estimated using ICA (d=25). mPFMs are a subset of the RSNs that are obtained from a high-dimensional (d=150) decomposition of cortical rfMRI data from 1003 HCP subjects. mPFMs are overlayed on cyan-coloured borders of the ICA-based RSNs for easier visual comparison. As can be noted, primary mPFMs are, in principle, the well-known RSNs: Default Mode 1 (DMN1), Default Mode 2, Language and Motor 1 networks. Secondary mPFMs depicted underneath show subsystems of these networks.

3.1 The functional relevance of mPFMs

Focussing on 25 and 150 sPROFUMO RSNs from rfMRI data of 1003 HCP subjects, we investigated the properties of the brain function that underlie mPFMs, and their added benefit over large-scale (distributed-only) decompositions for capturing multiscale information processing in the brain. We use the following terminology in the rest of the paper (see Figure 1): a) large-scale RSNs (yellow): 25 modes from low-dimensional decomposition; b) "primary" mPFMs (orange): 25 distributed modes from high-dimensional PFM decomposition that best-match the low-dimensional RSNs; c) "secondary" mPFMs (blue): remaining 125 modes from high-dimensional PFM decomposition.

3.1.1 Multiple distinct subcomponents within each conventional RSN

Each functional mode is typically characterised by two key characteristics: a) the spatial topography across brain voxels (i.e., a spatial map), from which one can derive spatial correlations or overlaps with other modes; b) a timecourse, and the resulting temporal correlations with other modes (i.e., functional connectivity).

This representation assumes that voxels within a mode are highly correlated with each other, such that they can be summarised with one consensus timecourse. In other words, it assumes that temporal variability across voxels within each RSN's spatial map is negligible. To determine whether lowD decomposition complies with this assumption, we asked: how many components are needed to explain a majority of variance in the voxel-wise activity within each large-scale RSN? We used Principal Component Analysis (PCA), which projects data onto new orthogonal axes of variation, and allows to identify distinct subcomponents based on the proportion of variance explained (details in 2.3.1). We found that (Figure S 4) the first Principal Component could only explain up to 53.2% of variance of temporal variability, with the second to fifth Principal Components each explaining ~5-10% of variance.

Next, we applied follow-up temporal ICA on data of each mode from lowD decomposition, as detailed in 2.3.2. Temporal ICA projects data onto statistically independent axes of variations and allow us to identify *temporally independent subcomponents* more interpretably. We restricted the number of subcomponents to 5 per mode; i.e., 5 subcomponents for each of the 25 modes. Figure 3a shows examples of these subcomponents for the Default Mode and Language Networks. We found that these subcomponents were spatially and temporally distinct, with one subcomponent being spatiotemporally best-represented by the original large-scale decomposition compared with the rest of the subcomponents (Figure S 5). Figure 3b shows the temporal correlation of each of the best-matching subcomponents to the large-scale RSN that they originated from. Interestingly, we found that even the temporal representation of the best-represented subcomponent ranged from Pearson correlation coefficient = 0.22±0.13 to 0.70±0.09 across the 25 large-scale RSNs (i.e., in many cases there is not a strong temporal match). We replicated this finding using two alternative subcomponent identification techniques, spatial ICA and PCA (Figure S 6).

These results provide evidence that voxel-wise temporal variability within large-scale RSNs results in *temporally distinct* subcomponents that are not well-represented by one timecourse, as assigned in conventional RSN decompositions. Instead, multiple subcomponents are needed to explain this voxel-wise temporal variability.

3.1.2 Temporal subcomponents were best represented by mPFMs

We next tested if the 5×25 subcomponents described in the previous subsection were better represented by mPFMs, compared with lowD large-scale RSNs alone. For this purpose, we applied a winner-takes-all approach: the timecourse of each of the subcomponents was correlated with: a) large-scale RSNs; b) primary mPFMs and c) secondary mPFMs. As shown in Figure 3d, we found that only 12 of the 125 subcomponents were temporally best represented by the large-scale RSNs, 9 of which were significantly higher (Bonferroni-corrected, Table S 2). The remaining 28

and 85 subcomponents were better represented by the primary and secondary mPFMs, respectively (28/28 and 83/85 statistically significant). Therefore, the ensemble of multiscale modes in mPFMs can capture the temporally distinct subcomponents within conventional large-scale RSNs.

743

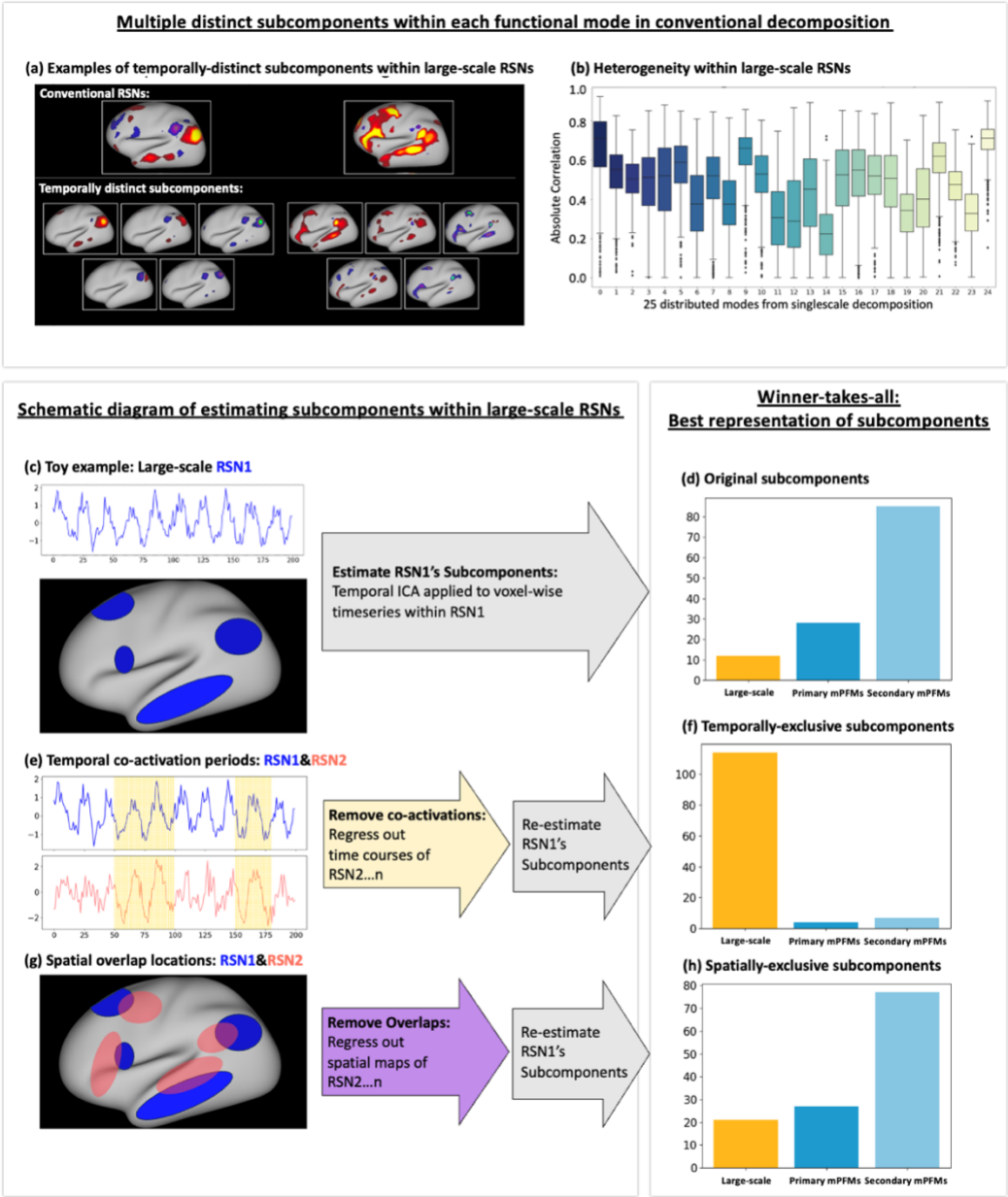

Figure 3 Temporally distinct subcomponents within conventional large-scale RSNs give rise to multiscale Probabilistic Functional Modes (mPFMs). A low-dimensional PFM decomposition consisting of 25 modes was estimated from cortical rfMRI data of 1003 HCP subjects, yielding conventional large-scale RSNs. Follow-up temporal ICA was applied to voxel-wise timeseries within each RSN to identify temporally-distinct subcomponents. a) Examples of two conventional RSNs and 5 subcomponents within each RSN. b) Temporal correlation of the best-matching subcomponents to the large-scale RSNs that they originated from. c) A schematic illustration of subcomponent estimation. d) A winner-takes-all approach was applied, and it was found that of the 25x5 subcomponents, only 12 were best represented by the large-scale RSNs that they originated from, whereas 28 and 85 subcomponents were better represented by primary and by secondary mPFMs, respectively. e) A schematic illustration of estimating temporally-exclusive subcomponents. f) A winner-takes-all approach was applied, and it was found that of the 25x5 temporally-exclusive subcomponents, 114 (i.e., the majority) were best represented by large-scale RSNs, 4 by primary and 7 by secondary mPFMs. g) A schematic illustration of estimating spatially-exclusive subcomponents. h) A winner-takes-all approach was applied, and it was found that of the 25x5 spatially-exclusive subcomponents, 21 were best represented by large-scale RSNs, 27 by primary and 77 by secondary mPFMs. These results show that temporally-distinct subcomponents within large-scale RSNs give rise to secondary PFMs, and that periods of temporal co-activation between subnetworks of large-scale RSNs give rise to secondary PFMs.

3.1.3 Secondary mPFMs were due to temporal co-activation periods of multiple large-scale RSNs

Resting state networks in the brain might be spatially and temporally correlated. Therefore, it is plausible that the subcomponents that give rise to the secondary mPFMs are either due to spatially overlapping regions or due to distinct temporal co-activation periods of multiple large-scale RSNs. To test this hypothesis, we compared the original subcomponents (described in the previous subsection) with two alternative scenarios: a) temporally-exclusive subcomponents: estimated from each large-scale RSN after the timecourses of the other 24 RSNs were regressed out of its timecourse; b) spatially-exclusive subcomponents: estimated from each large-scale RSN after spatial maps of the other 24 RSNs were regressed out of its spatial map. Conceptual illustrations of these two analyses are included in Figure 3e and g, respectively.

As shown in Figure 3f, we found that a majority of the temporally-exclusive subcomponents (114 out of 125, with 109 statistically significant, Table S 3) were temporally best represented by the *lowD decomposition*, unlike the original subcomponents that were best represented by mPFMs. Conversely, as shown in Figure 3h, the majority of the spatially-exclusive subcomponents (104 out of 125, 100 statistically significant, Table S 4) were instead best represented by primary and secondary mPFMs, i.e., similar to original subcomponents. All results were Bonferroni-corrected for multiple comparisons.

Put together, these analyses suggest that temporal co-activation periods of two or multiple large-scale RSNs from lowD decomposition give rise to new mixed-scale RSNs that are temporally distinct from the existing RSNs. These emerge as secondary mPFMs. In other words, secondary mPFMs are likely due to transient dynamic connectivity between subnetworks of large-scale modes, that are eliminated by temporal averaging (across voxels and/or timepoints) in low-dimensional decompositions.

3.1.4 Functional connectivity modelling using mPFMs

Having examined timecourse modelling in mPFMs, we next investigated the functional connectivity (i.e., temporal correlations) between the modes to compare large-scale RSNs (lowD, d=25) with mPFMs (d=100 and 150). This was done using 4999 subjects in UK Biobank (Figure S 7a), and separately using 1003 subjects in HCP (Figure S 7b). Figure S 7-right shows the histograms of between-subject consistency of functional connectivity. We observed that in both datasets, between-subject consistency was increased with dimensionality. In HCP, this was increased from 0.19±0.09 for d=25 to 0.46±0.06 and 0.59±0.05 for d=100 and 150, respectively. In UK Biobank, this was increased from 0.07±0.07 for d=25 to 0.43±0.13 and 0.70±0.12 for d=100 and 150, respectively. Therefore, mPFMs yielded more consistent functional connectivity values than

lowD. Interestingly, subject-specific functional connectivity became sparser for mPFMs (Figure S 7-left). This sparsity can be understood in light of the previous subcomponent analysis: by attempting to merge the time courses of multiple temporally distinct subcomponents, conventional RSNs can yield spurious connectivity between large-scale modes, which in reality can be due to one or a few of the underlying subcomponents. mPFMs capture the subcomponents using secondary modes, thus reducing such spurious connectivity, and yielding sparser subject-specific functional connectivity. As a result, functional connectivity in mPFMs becomes less contaminated by modelling inaccuracies, and more consistent across individuals.

3.1.5 Cross-scale interactions of mPFMs predict individualistic traits

mPFMs can directly characterise the *spatiotemporal correlations between modes across multiple scales of information processing in the brain*. We evaluated the utility of temporal and spatial correlations between mPFMs for out-of-sample prediction of a range of phenotypes in UK Biobank. This was performed using volumetric rfMRI data from 4999 UK Biobank subjects and 893 phenotypes divided into 6 categories: region-wise cortical area (148 phenotypes) and thickness (148), White Matter (WM) tracts' microstructure (451), task fMRI response to the contrast of emotional faces and shapes (1), cognition (68) and cardiovascular health (77), see 2.6.1 for details of phenotypes, feature extraction and prediction pipeline. Figure 4a&b, show functional and spatial connectivity of large-scale RSNs (d=25) and mPFMs (d=100) for an example subject. As demonstrated in these panels, cross-subject consistency of both functional and spatial connectivity was higher in mPFMs compared with large-scale RSNs.

First, we compared the prediction performance of mPFMs with large-scale RSNs, shown in Figure 4c. We found that spatial connectivity of mPFMs, after matching the number of features as elaborated in 2.6.4, significantly outperformed (Bonferroni corrected, Table S 5) that of large-scale decomposition in predicting all phenotype categories except cognitive scores. Similarly, functional connectivity of mPFMs significantly outperformed that of large-scale decomposition in predicting region-wise cortical area and task fMRI phenotypes. This shows that subject-specific spatiotemporal correlations between mPFMs carry additional information about each person's traits that is otherwise missed in large-scale decomposition.

807

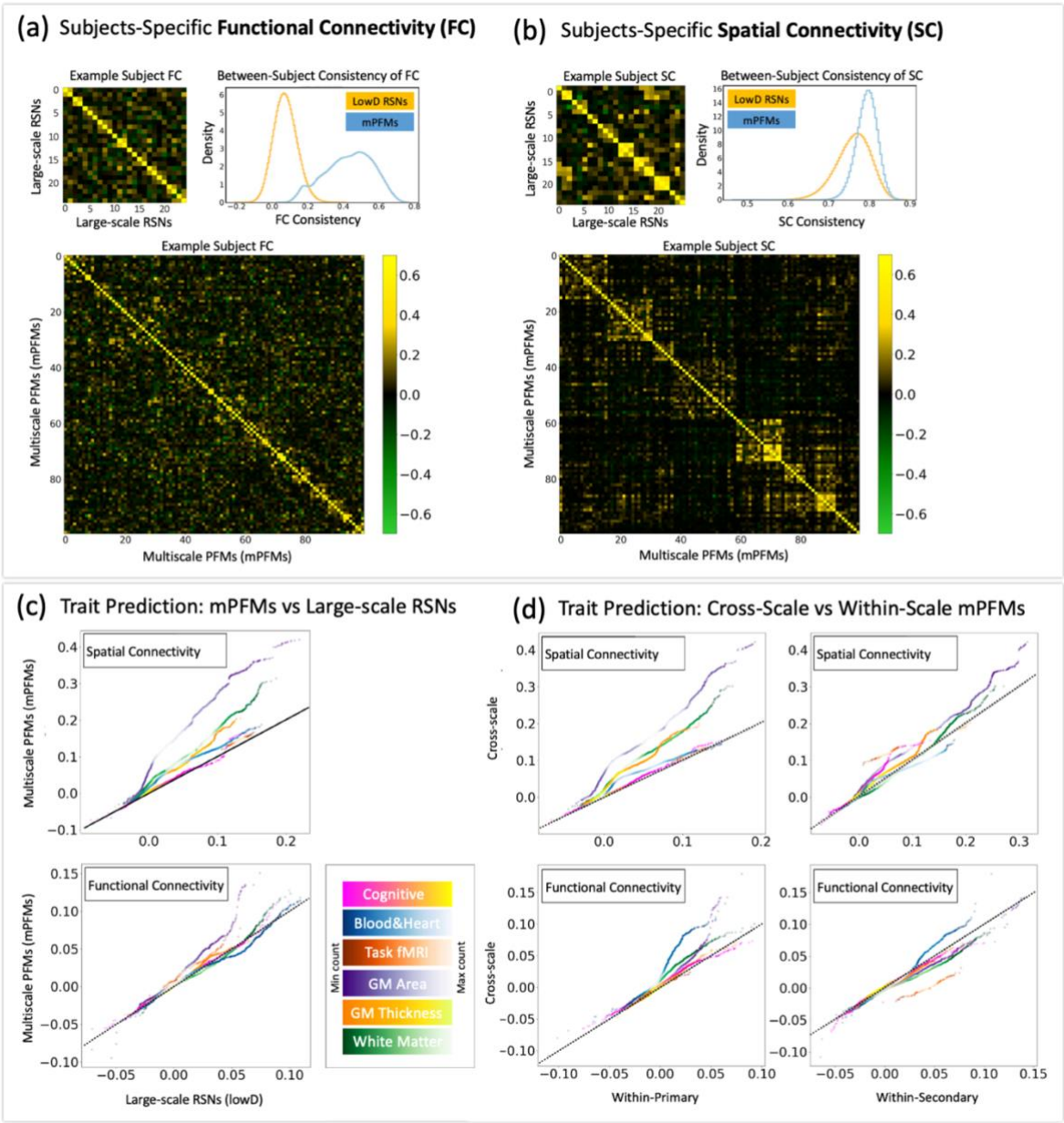

Figure 4 Spatiotemporal interaction of mPFMs captures subject variability that is predictive of personalised traits in UK Biobank. a) Examples of temporal correlations (functional connectivity) between the modes, which were used as one set of features for trait prediction. Functional connectivity matrices are shown separately for conventional large-scale RSNs from lowD (25 mode) decomposition and mPFMs (highD, 100 mode) decomposition. Histograms show that between-subject consistency of functional connectivity was notably higher in mPFMs. b) Examples of spatial connectivity between the modes which were used as another set of features for trait prediction. Spatial correlation matrices are shown separately for conventional large-scale RSNs from (lowD, 25 mode) decomposition and mPFMs (highD, 100 mode) decomposition. Histograms show that between-subject consistency of spatial connectivity was notably higher in mPFMs. c) Comparing the prediction performance of mPFMs with large-scale RSNs for a range of 893 phenotypes across 6 categories. Spatial (top) and temporal (bottom) correlations were used as predictors. d) Cross-scale correlations refer to spatial and temporal interactions between primary and secondary mPFMs, an aspect of the brain's functional connectivity that conventional distributed-only RSNs or localised-only parcellations cannot capture. The prediction performance of this cross-scale spatial (top) and temporal (bottom) correlations were compared to that of within primary or within secondary modes, and a clear improvement was observed, especially in spatial domain. Density scatter plots are used to compare prediction performances, and colours denote densities, as shown in colourbars. Dotted black lines denote equal performance.

Next, we divided spatiotemporal interactions of mPFMs into three subtypes: i) within primary mPFMs; ii) within secondary mPFMs, and iii) between primary and secondary mPFMs. Aiming to evaluate the added benefit of cross-scale interactions (iii), we compared its prediction performance to within-scale interactions (i.e., i&ii). As shown in Figure 4d-top, cross-scale spatial connectivity significantly outperformed (Bonferroni corrected, Table S 5) within-secondary for region-wise cortical area and thickness, task fMRI and cognitive phenotype categories, and they additionally outperformed within-primary for all the phenotype categories, except cognitive scores. Cross-scale functional connectivity, as shown in Figure 4d-bottom, performed on par with or worse than within-secondary, but significantly outperformed (Bonferroni corrected, Table S 5) within-primary for region-wise cortical area, thickness, cardiovascular health metrics.

These results demonstrate that spatial and temporal connectivity of the ensemble of mPFMs yields improved biomarkers for individualistic traits in two ways: first, compared with conventional large-scale RSNs, and second, subject variability captured in spatiotemporal correlations *across* multiple scales is more phenotypically relevant compared with within a single scale.

3.2 Validation of mPFMs

Our results so far have focussed on resolving the origins and functional relevance of mPFMs in modelling brain connectivity. Here we set to validate this new decomposition based on simulations, its reproducibility between HCP and UK Biobank datasets and similarity to HCP's Multimodal Parcellation.

3.2.1 Simulations

Based on real data in section 3.1, we found evidence that secondary mPFMs likely emerge to explain temporally-distinct subcomponents within the conventional large-scale RSNs. In other words, we showed that the ensemble of multiscale modes captures the integrated-segregated functional brain connectivity. Here we simulated datasets, where distributed and localised modes co-existed, such that localised modes were spatial sub-modes of the distributed modes, with distinct timecourses. Using these simulations, we tested how well PFM decomposition can recover the co-existence of multiscale modes, and compared its performance to spatial ICA, which is the current standard technique for RSN modelling. Across 10 simulated datasets, each included 12 modes, 6 distributed and 6 localised. Figure 5a shows examples of group-average simulated modes in one of the datasets. We compared three elements of the estimated modes with the ground truth: group spatial maps, subject spatial maps, and subject time courses, shown in Figure 5 b, c, d, respectively. For the group spatial maps, the accuracy of PFM decomposition for distributed and localised modes was 0.93 and 0.78, respectively, which was reduced to 0.74

and 0.47 in spatial ICA modes. For the subject spatial maps, the accuracy of PFM decomposition for distributed and localised modes was 0.87 and 0.70, respectively, which was nearly halved (0.48 and 0.31) in spatial ICA modes. For the subject timecourses, the accuracy of PFM decomposition for distributed and localised modes was 0.95 and 0.64, respectively, which was reduced to 0.78 and 0.33 in spatial ICA modes. Therefore, at both group and subject level, PFMs can recover the simulated multiscale modes with overall (average) accuracy of ~0.8 or higher.

3.2.2 Reproducibility across datasets and data formats

The next step of validation involved cross-dataset reproducibility tests. We applied sPROFUMO separately to three rfMRI datasets and obtained 150 modes per run: 1) 1003 HCP subjects, volumetric whole-brain rfMRI (HCP VOL); 2) 1003 HCP subjects, CIFTI cortical rfMRI (HCP CORT); 3) 4999 UKB subjects, volumetric whole-brain rfMRI (UKB VOL). All these scenarios yielded mPFMs consisting of primary and secondary modes. We paired these mPFMs across runs based on spatial correlation of group-level spatial maps.

Firstly, by comparing HCP VOL and UKB VOL (Figure 5e), we found 100 modes to show very good matching (absolute correlation >= 0.7), 40 modes to show good matching (absolute correlation between 0.5 and 0.7), and 10 modes to show mediocre matching (absolute correlation <= 0.5). Some of the modes in the latter category were due to differences in structured noise across datasets (Figure S 8). The top 40 best matching mPFMs across these datasets are shown in Figure 5g.

Secondly, by comparing HCP VOL and HCP CORT (Figure 5f), we found 74 modes to show very good matching (absolute correlation >= 0.7), 28 modes to show good matching (absolute correlation between 0.5 and 0.7), and 48 modes to show mediocre matching (absolute correlation <= 0.5). Some of the modes in the latter category were due to structured noise differences across data-types, and the remainder were due to inherent differences between volumetric and CIFTI fMRI reconstructions.

3.2.2.1 Reproducibility across multiple dimensionalities

We additionally tested for reproducibility of the mPFMs across multiple dimensionalities: 50, 100 and 150. This was done separately for UK Biobank and HCP datasets, with each dimensionality starting from a different independent initialisation. Results are shown in **Figure S 9**, starting from low-dimensional d=25 decomposition, we compared 25 versus 50, 50 versus 100 and 100 versus 150. The modes were sorted and paired based on spatial correlations between the group-level spatial maps. For the UK Biobank volumetric fMRI, we found mean±STD of the paired correlations to be: 0.90 ± 0.11 (d = 25 vs 50), 0.84 ± 0.14 (d = 50 vs 100), 0.82 ± 0.14 (d = 100 vs 150). For the HCP Cortical CIFTI fMRI we found mean±STD of the paired correlations to be: 0.88 ± 0.08 (d = 25 vs 50), 0.90 ± 0.11 (d = 50 vs 100), 0.88 ± 0.11 (d = 100 vs 150). These strong pairings show that,

as we increase the dimensionality, the existing modes at lower dimensions are largely maintained, and new modes are added to the ensemble, indicating an eventual addition of more and more secondary PFMs.

3.2.3 Validating secondary mPFMs using HCP's Multi-Modal Parcellation

As the next step, we validated mPFMs based on the HCP's Multimodal Parcellation v1.0 (HCP-MMP), aiming to validate the secondary mPFMs. We compared mPFMs against the HCP-MMP, which includes 360 cortical parcels. This parcellation includes both group-average and subject-specific parcels, allowing us to validate secondary mPFMs at both levels. We applied sPROFUMO to cortical rfMRI data of 446 HCP subjects (matching the number of subjects in HCP-MMP) and estimated 360 modes. Similar to before, this yielded a multiscale ensemble of modes (i.e., mPFMs).

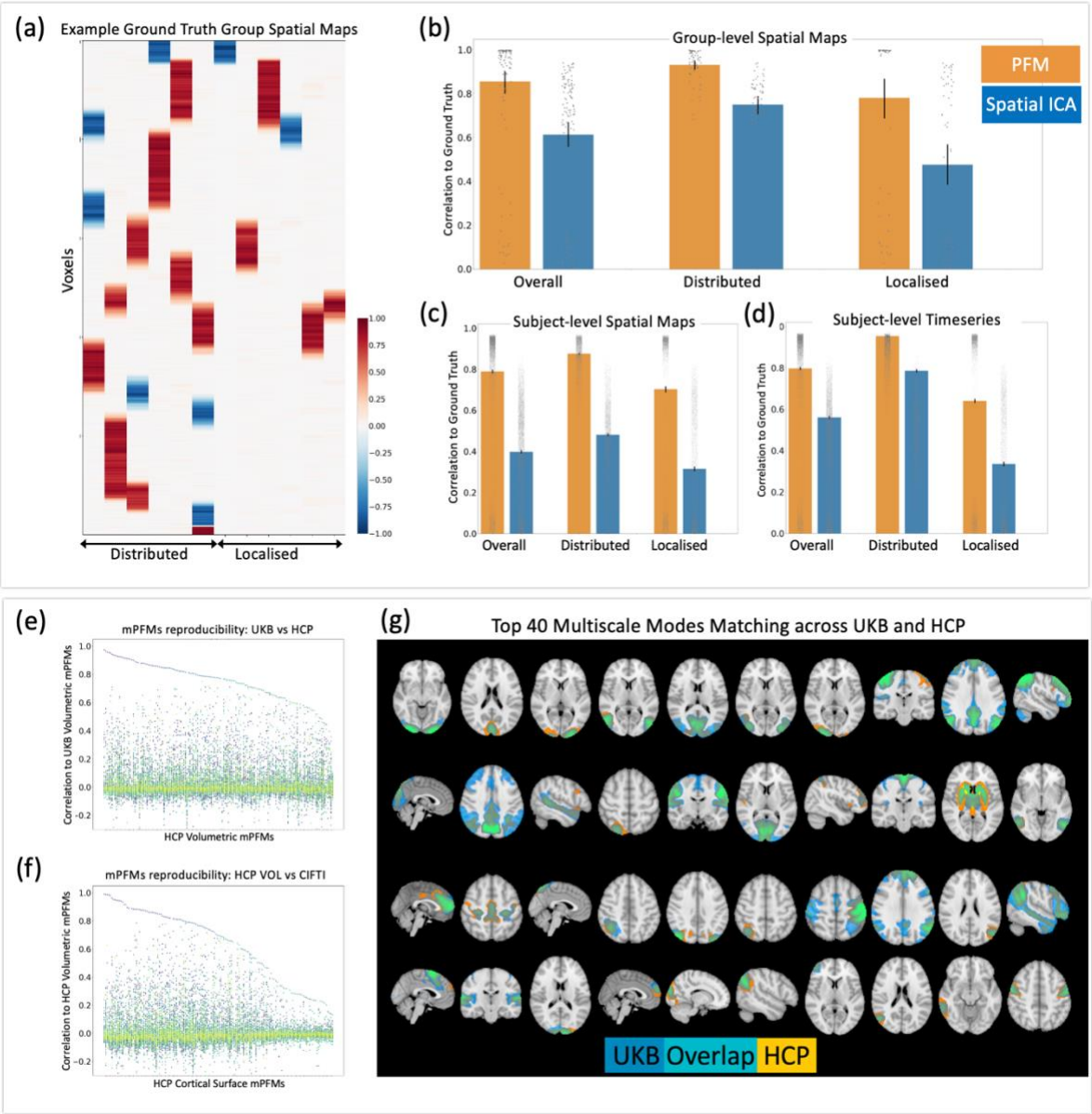

Figure 5 Validation of mPFMs using synthetic data (top panel) and reproducibility in real data (bottom panel). Simulations: a) Examples of simulated group-average spatial maps (one of ten simulated datasets). 12 modes were simulated, 6 mimicking distributed modes, 6 mimicking localised subcomponents of these modes. Aiming to validate findings of section 3.1.1, we observed that PFM decomposition can well capture: b) group spatial maps; c) subject spatial maps; d) subject timecourses of these co-existing modes. The performance was ~50% higher than that of spatial ICA at group level and up to 100% higher at subject level. Real data reproducibility: PFM decomposition with 150 modes was compared at the group level between: e) UK Biobank (UKB) and HCP volumetric fMRI; each HCP mode is shown on the x axis and correlations to all UKB modes on the y axis. Clear one-to-one pairing can be seen for the top 100 modes. f) A similar comparison was conducted between HCP volumetric and surface-based fMRI data types. g) The top 40 matching mPFMs between HCP and UKB (both volumetric fMRI) are illlustrated. HCP modes are shown in yellow, UKB in blue and overlap in green/cyan.

We found an interesting group correspondence (Figure S 10a) between secondary mPFMs and HCP-MMP, especially 95 parcel-like mPFMs showed a strong correspondence to HCP-MMP (dice

similarity >=0.5), illustrated in Figure S 10c. 73 mPFMs showed less than 0.2 dice similarity to any HCP-MMP parcels; these included conventional low-dimensional modes and the less well-known variants of distributed modes, and typically occupied multiple distant sub-regions. We further found that modes with high group-level correspondence also showed high subject-level correspondence (Figure S 10b), with the top 95 matching modes at the group-level also showing correlation of 0.47±0.19 (average ± standard deviation across subjects and modes) at the subject-level. These evaluations add to the previous validation steps, by demonstrating that those of the mPFMs that are parcel-like have clear spatial similarities to the state-of-the-art individualised parcellations, both at the group- and the subject-level.

In summary, this section on mPFM validation sheds light on robustness and interpretation of the multiscale representation of the RSNs. Specifically, in section 3.2.1, we simulated scenarios resembling multiscale modes; i.e., temporally distinct pairs of modes where one mode is a spatial sub-node of the other, thus mimicking the primary and secondary modes in real data. We showed that these can be recovered using PFM decomposition much more accurately than ICA decomposition. Next, in section 3.2.2, we demonstrated reproducibility of mPFMs across two large-scale datasets (HCP and UKB), two standard fMRI data formats (volumetric and CIFTI), and a range of dimensionalities (25, 50, 100, 150) in both HCP and UKB. These results demonstrate that the multiscale representation is a stable finding and not a byproduct of certain data properties or choices of preprocessing or dimensionality. Finally, in section 3.2.3, we demonstrated that secondary mPFMs can be reliably linked to the well-established HCP-MMP parcellation. Considering that the multiscale representation is (as shown in earlier in **Figure 1**, and section 3.2.2.1), in principle, an ensemble of the large-scale RSNs (primary mPFMs), with the addition of the fine-grained secondary mPFMs, this analysis validated that the secondary mPFMs (especially the parcel-like) are meaningful functional entities, and not due to fMRI smoothness or noise properties.

3.3 mPFMs yield enhanced biomarkers of personalised traits

We have so far investigated the brain properties underlying mPFMs, and their added benefit over conventional large-scale RSNs. Here we focus on the usefulness of individual-specific mPFMs for providing biomarkers for traits and disease. This was done using two steps: 1- population covariations between mPFMs and behavioural traits in HCP; 2- out of sample prediction of phenotypes in UK Biobank. CCA and phenotype prediction results serve two distinct, complementary goals. mPFMs provide a new type of RSN representation, which is subject-specific and multiscale. Therefore, we aimed to test: 1- how interpretably they can be linked to non-imaging variables (CCA analysis in HCP), and 2- compared to existing techniques, how accurately they can predict non-imaging variables (phenotype prediction in UK Biobank).

3.3.1 Positive-negative axis of population covariation between mPFMs and behaviour

The goal of CCA analysis was mainly qualitative, and to illustrate how interpretably mPFMs can be linked to personalised traits. In order to achieve this, we referred to the previous literature and investigated whether mPFMs can capture one of the key findings in brain-behaviour literature: a positive-negative mode of population covariation that has been discovered to link brain features to cognition and lifestyle. Positive traits such as language comprehension or life satisfaction are placed at one end of this axis, and negative traits such as antisocial behaviour or substance use at the other end. This was originally reported by Smith et al., (2015) using HCP data, and has been replicated on other large datasets since (Goyal et al., 2022).

Canonical Correlation Analysis (CCA), allows to investigate associations between a set of features from brain data and a set of personalised traits (i.e., phenotypes), all in a single integrated multivariate analysis. Using CCA, we identified modes of population co-variation between brain and phenotypes, as pairs of canonical components along which the phenotypes and mPFMs co-varied in a similar way across subjects. We used the same analysis pipeline as Smith et al., (2015), as detailed in section 2.7.

To examine the link between personalised traits and spatiotemporal characteristics of mPFMs, CCA was conducted on features related to spatial maps, spatial and functional connectivity of mPFMs. Figure S 11a shows results of correlation between brain and phenotypes after CCA transformation. As shown in Figure 6a, we identified the raw behavioural metrics that were most strongly linked to the top CCA components. This was done by correlating the behavioural metrics with the topmost CCA components related to spatial maps, spatial and functional connectivity, and taking the average correlation across the three. This unravelled a single axis of population variation: at one end of this axis, we found behavioural traits positively associated with the prominent CCA component. These included metrics related to cognition and emotion such as performance on language tests, executive function, self-regulation, and life satisfaction. At the other end of this axis, we found behavioural traits negatively associated with the prominent CCA component. These included metrics related to psychiatric disorders and substance use such as rule breaking, antisocial behaviour, alcohol, tobacco and cannabis use. Therefore, mPFMs, similar to existing literature, provide a simple interpretable mapping between brain features and positive versus negative traits.

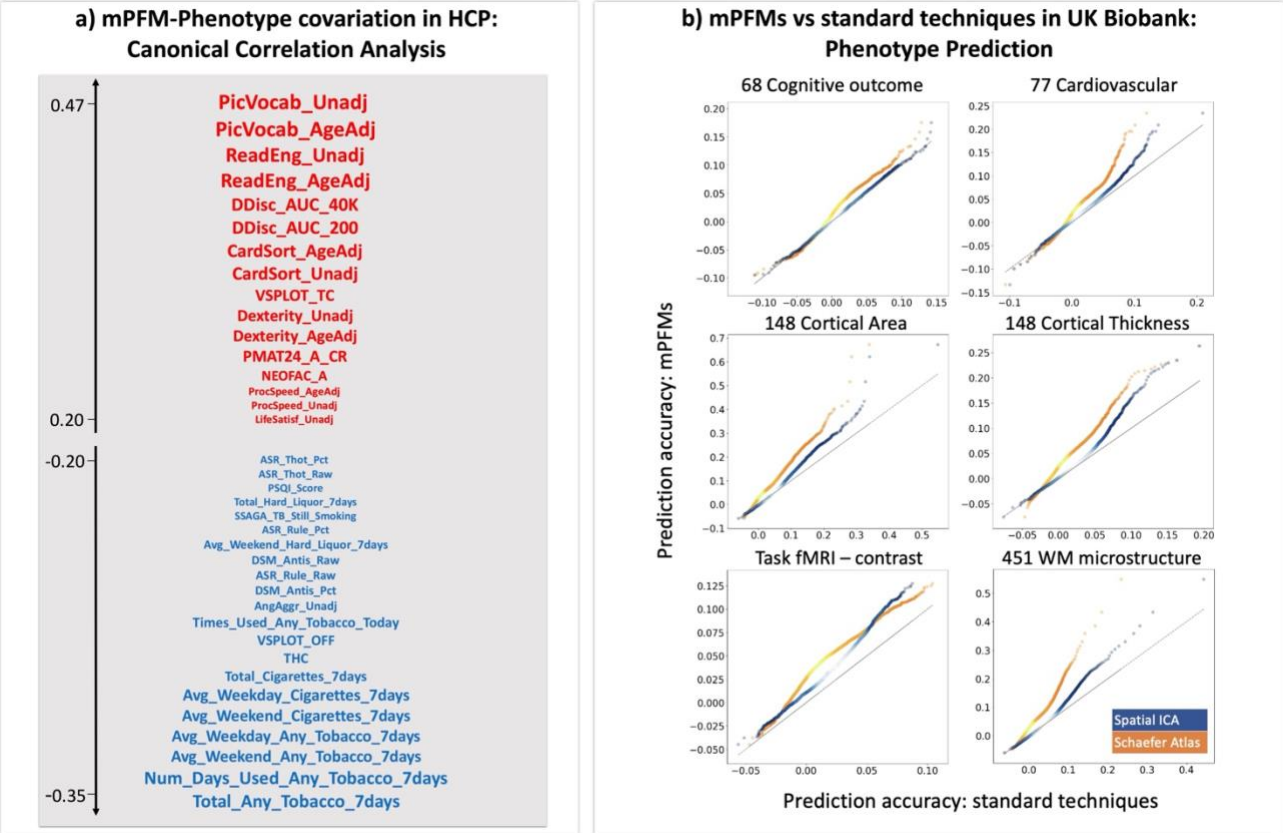

Figure 6 mPFMs yield biomarkers for behaviour and traits: a) shared population covariation between mPFMs and behaviour in HCP: A positive-negative axis of population covariation identified based on top CCA components of spatial and temporal mPFMs features. b) mPFMs yield enhanced fMRI biomarkers of traits and health compared with current standard techniques in UKB. 893 phenotypes across six categories were used. Comparisons between trait prediction accuracies of mPFMs (y-axis) with spatial ICA and the Schaefer parcellation (x-axis blue and orange, respectively) are shown. Correlation coefficient between actual and predicted phenotypes in the test set was used as metric of accuracy. Density scatter plots are used to compare mPFM's prediction accuracies to that of the standard techniques, each dot corresponds to the prediction accuracy of a mode for a phenotype/trait, and brighter colours (e.g. light yellow and light blue) denote higher densities. Dotted grey lines denote equal performance. 100-mode decompositions were used for all the methods, and prediction accuracies were computed per mode, in order to obtain a more detailed summary of the modes. mPFMs outperformed the standard techniques for all the categories.

3.3.2 Enhanced predictions of personalised traits in UK Biobank

Next, we evaluated the potential of mPFMs for out-of-sample prediction of traits. For this purpose, we compared mPFMs to two of the most commonly used techniques in the literature: spatial ICA, the current main RSN modelling technique in the core UK Biobank pipeline (Alfaro-Almagro et al., 2018); and the Schaefer atlas, a high dimensional hard parcellation derived from the Yeo parcellation (Schaefer et al., 2018). For this purpose, we used the 100-dimensional Schaefer parcellation and further applied the other two methods to volumetric rfMRI data of 4999 UK Biobank subjects, estimating 100 mPFMs and 100 ICs. We used features of each functional mode (feature matrix, X) to make predictions about phenotypes (target, y), details in

2.6. We used ElasticNet regression with 5-fold nested cross-validation, and predicted 893 phenotypes divided into 6 categories: region-wise cortical area and thickness (296), White Matter (WM) tracts' microstructure (451), task fMRI response to the contrast of emotional faces vs shapes (1), cognition (68) and cardiovascular health metrics (77). Correlation coefficient between actual and predicted phenotypes in the test sets was used as metric of accuracy. Results are shown in Figure 6b, with prediction accuracies of mPFMs shown on the y-axis, and accuracies of spatial ICA modes on the x-axis. We found that mPFMs significantly outperformed (Bonferroni corrected, Table S 6) Schaefer parcellation in prediction of all the phenotype categories, and significantly outperformed spatial ICA in prediction of all the phenotype categories except cognitive traits. These results were replicated using higher dimensionalities, as shown in **Figure S 12**.

These results, taken together with the findings of section 3.1.5, demonstrate three key benefits of mPFMs over single-scale representations for phenotype predictions: firstly, the spatiotemporal connectivity between the large-scale and fine-grained PFMs (i.e., cross-scale connectivity) outperforms that of within-scale connectivity - this was illustrated using the PFM modelling framework only, since standard techniques do not yield multiscale representations; secondly, the spatiotemporal connectivity of the full multiscale representation outperforms that of lowD large-scale-only RSNs; thirdly, the full set of features obtained from the multiscale highD representation (including spatial topographies and spatiotemporal connectivity) outperforms that of standard single-scale highD representations obtained from ICA and Schaefer atlas.

4 Discussion

In this paper, we presented Multiscale Probabilistic Functional Modes (mPFMs), a new representation of the brain's functional modes, that improves fMRI modelling and its utility for prediction of cognition and health. mPFMs were identified using stochastic Probabilistic Functional Mode (sPROFUMO) modelling and, unlike any existing representation, include an ensemble of modes across multiple scales. sPROFUMO does not impose or prevent this behaviour by design, instead, it allows the modes to be spatiotemporally correlated, and accumulate Bayesian evidence across individuals in large populations. Therefore, if a multiscale organisation arises from the results, this is driven by the structure in the data. Indeed, as demonstrated in the Results section, the presence of multiple distinct subcomponents within each large-scale (distributed) mode indicates spatiotemporal multiplicity within these modes, which cannot be captured by distributed-only or localised-only representations. Put together, our results demonstrate mPFMs' potential to provide a useful representation of multiscale information processing in the brain, and enhance fMRI-derived biomarkers for traits and disease. We release this representation publicly; we have made the group-level maps openly available,

1001 whereas subject-level maps will be released in the future as part of ongoing work to rebuild a
1002 "v2" UKB pipeline (Alfaro-Almagro et al., 2018).

1003 Multiscale modes are compatible with the notion of hierarchal information processing in the
1004 brain; that brain function is organised in a hierarchical manner, with each level of the hierarchy
1005 underlying increasingly complex and abstract processing (Cordes et al., 2002; Deco and
1006 Kringelbach, 2017; Friston, 2008; Vidaurre et al., 2017). Hierarchical processing is enabled via
1007 interactions between local unimodal information processing and distributed higher level
1008 cognitive function. Such interactions have been well-established, e.g., in visual search tasks,
1009 where interactions between sub-regions of the visual cortex and fronto-parietal attention
1010 networks allows for top-down modulation of visual processing by attention, as well as bottom-
1011 up modulation of attention by the incoming stimulus(Buschman and Miller, 2007; Gazzaley and
1012 Nobre, 2012). Interactions between brain systems at different levels of these hierarchies cannot
1013 be directly modelled using existing functional mode representations, considering that they can
1014 either be used to measure interactions between distributed modes (lowD decompositions) or
1015 localised regions (highD parcellations). As a result, the interactions *between* these two layers will
1016 have to be modelled using post-hoc techniques, such as hierarchical clustering (Smith et al.,
1017 2013a) or module detection(Sporns and Betzel, 2016). Even then, the resulting connectivity will
1018 depend on the assumptions of the post-hoc techniques, and change depending on the choice of
1019 parameters, rather than being integrated within the original brain function mapping.

1020 In contrast, mPFM inherently incorporates such multiscale connectivity within fMRI
1021 decomposition. Crucially, as we demonstrated in the Results, these interactions provide stronger
1022 biomarkers of cognition and health compared with connectivity within large-scale or within
1023 localised modes alone. This complements previous brain-behaviour research (Bolt et al., 2018),
1024 that compared the association of individual brain regions and their respective whole-brain
1025 networks with task performance. The authors showed that for tasks such as working memory
1026 and relational, areal-level associations to task performance will become redundant when large-
1027 scale network covariation in accounted for, unlike arithmetic task, where areal-level and
1028 network-level associations showed distinctively significant roles. This study, similar to ours,
1029 highlights the need for multiscale representations.

1030 Additionally, by allowing the modes to be spatially overlapping, mPFMs can capture functional
1031 multiplicity in the brain's organisation, where a single region contributes to different functions
1032 depending on the context. Until recently, spatially overlapping functional modes have been
1033 largely absent from resting state literature. Specifically, studies have commonly relied on either
1034 spatial ICA (Beckmann et al., 2005; Calhoun and Adali, 2012), that imposes spatial independence
1035 between the modes, or hard parcellations, that enforce rigid boundaries between the modes
1036 (Craddock et al., 2012; Yeo et al., 2011). Therefore, the literature has relied on spatial bases that

minimise or prevent spatial overlap. However, growing evidence is now highlighting the importance of spatial overlap in modelling cross-individual variability in brain function (Bijsterbosch et al., 2019). Importantly, recent successful characterisation of brain connectivity using gradients or connectopy mapping (Haak et al., 2018; Margulies et al., 2016) has brought functional multiplicity to the fore of resting state connectivity. These techniques have unravelled multiple overlapping connectivity patterns within brain regions, e.g., retinotopic organisation of the visual cortex, and follow-up studies have highlighted the cognitive and clinical relevance of these overlapping connectopy maps. As such, allowing for spatial overlap, in addition to and not instead of temporal correlations between the modes, is proving imperative for obtaining a more comprehensive representation of brain function. mPFMs are, on the one hand, similar to conventional spatial ICA modes, in that they capture large-scale prominent networks in the brain. One the other hand, mPFMs additionally allow multiple overlapping subcomponents within these networks to co-exist. As such, mPFMs can be viewed as a new category of brain function mapping that stands somewhere in-between conventional RSNs and the more recent gradients of connectivity. In this context, future studies would be valuable, to compare mPFMs to gradients of connectivity. Additionally, comparison to the recently-introduced tensor-based decomposition methods such as NASCAR (Li et al., 2023) and Tucker (Han et al., 2024) would be valuable. Unlike spatial/temporal ICA, and similar to PROFUMO, these methods allow the estimation of RSNs that are both spatially and temporally correlated.

In addition to their functional relevance, mPFMs also provided enhanced biomarkers for prediction of individualistic traits in UK Biobank. To benchmark of mPFMs against standard ICA-based or hard parcellation techniques, we used a range of imaging-derived and non-imaging derived phenotypes such as cortical geometry, white matter microstructure, brain response to cognitive tasks, behavioural cognitive scores, and cardiovascular health outcome. The overall prediction accuracies of mPFMs were found to be significantly higher than standard techniques for various phenotype categories. One key reason for this higher prediction accuracy is likely the more comprehensive spatiotemporal feature set in mPFMs. On the one hand, hard parcellations yield binary spatial maps for each mode/parcel, and no spatial overlaps between the modes, thus summarising individual-specific brain features solely in temporal space. Somewhat similarly, Spatial ICA, due to its assumption of spatial independence between the modes, minimises spatial overlap by design (Farahibozorg et al., 2021; Kiviniemi et al., 2009; Smith et al., 2013b). Consequently, both approaches yield singlescale decompositions of functional modes. Therefore, large-scale modes, and their spatiotemporal correlation with each other and with localised modes, cannot be fully characterised by the standard techniques. Therefore, mPFMs yield a richer set of spatial and temporal features that, even after matching the number of features between the methods, can outperform the standard techniques in reflecting individualistic traits in brain function.

5 Acknowledgements

We would like to thank Dr. Matthew F Glasser for providing subject-specific HCP-MMP (v1.0) parcellations and helpful input on an earlier version of the manuscript. We are grateful to HCP and UK Biobank for making these invaluable resources available, and to the UK Biobank and HCP participants for dedicating their time to make these data possible.

Funding:

- Royal Academy of Engineering under the Research Fellowship programme RF2122-21-310 (SRF)
- Wellcome Trust grants 215573/Z/19/Z (SMS and MWW), 106183/Z/14/Z (MWW). Wellcome Centre for Integrative Neuroimaging is supported by core funding from the Wellcome Trust (203139/Z/16/Z).
- MRC Mental Health Pathfinder grant MC_PC_17215 (SMS)
- National Institute of Health grants R01 MH128286 and R01 MH132962 (JDB)
- The computational aspects of this research were partly carried out at Oxford Biomedical Research Computing (BMRC), that is funded by the NIHR Oxford BRC with additional support from the Wellcome Trust Core Award Grant Number 203141/Z/16/Z. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

For the purpose of open access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission.

Author contribution:

Conceptualisation: SRF, SJH, MWW, SMS; Methodology: SRF, SJH, JDB, MWW, SMS; Investigation: SRF, MWW, SMS, Software: SRF, SJH, SMS; Validation: SRF; Data curation, Formal Analysis and Visualisation: SRF; Resources: MWW, SMS; Funding acquisition: SRF, MWW, SMS; Writing—original draft: SRF; Writing—review & editing: SRF, SJH, JDB, MWW, SMS.

Competing interests:

Authors declare that they have no competing interests related to this manuscript. At the time of contributing to this research, SJH was affiliated with Universities of Oxford & Otago, he is currently employed by ADInstruments.

Data availability:

- UK Biobank data is available upon registration and applying for data access from UK Biobank website: http://www.ukbiobank.ac.uk/register-apply.

- Human Connectome Project data is available upon registration from HCP website: https://db.humanconnectome.org.
- Group average mPFMs obtained from HCP and UK Biobank are available here: https://www.fmrib.ox.ac.uk/ukbiobank/PROFUMO.

Code availability:

- Code for sPROFUMO is currently available from the following repository, it will be made available in an upcoming FSL release: https://git.fmrib.ox.ac.uk/profumo/profumo/-/tree/sprofumo-cpp-clean.
- Code used for simulations is available from the following public repository: https://git.fmrib.ox.ac.uk/profumo/PFM_Simulations.

6 References

Alfaro-Almagro, F., Jenkinson, M., Bangerter, N.K., Andersson, J.L.R., Griffanti, L., Douaud, G., Sotiropoulos, S.N., Jbabdi, S., Hernandez-Fernandez, M., Vallee, E., Vidaurre, D., Webster, M., McCarthy, P., Rorden, C., Daducci, A., Alexander, D.C., Zhang, H., Dragonu, I., Matthews, P.M., Miller, K.L., Smith, S.M., 2018. Image processing and Quality Control for the first 10,000 brain imaging datasets from UK Biobank. Neuroimage 166, 400–424. https://doi.org/10.1016/j.neuroimage.2017.10.034

Alfaro-Almagro, F., McCarthy, P., Afyouni, S., Andersson, J.L.R., Bastiani, M., Miller, K.L., Nichols, T.E., Smith, S.M., 2020. Confound modelling in UK Biobank brain imaging. Neuroimage 224, 117002. https://doi.org/10.1016/j.neuroimage.2020.117002

Andrews-Hanna, J.R., Reidler, J.S., Sepulcre, J., Poulin, R., Buckner, R.L., 2010. Functional-Anatomic Fractionation of the Brain's Default Network. Neuron 65, 550–562. https://doi.org/10.1016/j.neuron.2010.02.005

Bastian, M., Heymann, S., Jacomy, M., 2009. Gephi: An Open Source Software for Exploring and Manipulating Networks. Proceedings of the International AAAI Conference on Web and Social Media 3, 361–362. https://doi.org/10.1609/icwsm.v3i1.13937

Beckmann, C.F., DeLuca, M., Devlin, J.T., Smith, S.M., 2005. Investigations into resting-state connectivity using independent component analysis. Philosophical Transactions of the Royal Society B: Biological Sciences 360, 1001–1013. https://doi.org/10.1098/rstb.2005.1634

Beckmann, C.F., Smith, S.M., 2004. Probabilistic Independent Component Analysis for Functional Magnetic Resonance Imaging. IEEE Trans Med Imaging 23, 137–152. https://doi.org/10.1109/TMI.2003.822821

Bellec, P., Rosa-Neto, P., Lyttelton, O.C., Benali, H., Evans, A.C., 2010. Multi-level bootstrap analysis of stable clusters in resting-state fMRI. Neuroimage 51, 1126–1139. https://doi.org/10.1016/j.neuroimage.2010.02.082

Bijsterbosch, J.D., Beckmann, C.F., Woolrich, M.W., Smith, S.M., Harrison, S.J., 2019. The relationship between spatial configuration and functional connectivity of brain regions revisited. Elife 8. https://doi.org/10.7554/eLife.44890

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. https://doi.org/10.1002/mrm.1910340409

Bolt, T., Prince, E.B., Nomi, J.S., Messinger, D., Llabre, M.M., Uddin, L.Q., 2018. Combining region- and network-level brain-behavior relationships in a structural equation model. Neuroimage 165, 158–169. https://doi.org/10.1016/j.neuroimage.2017.10.007

Buckner, R.L., Vincent, J.L., 2007. Unrest at rest: Default activity and spontaneous network correlations. Neuroimage 37, 1091–1096. https://doi.org/10.1016/j.neuroimage.2007.01.010

Buschman, T.J., Miller, E.K., 2007. Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. Science, 315, 1860–1862. https://doi.org/10.1126/science.1138071

Calhoun, V.D., Adali, T., 2012. Multisubject independent component analysis of fMRI: A decade of intrinsic networks, default mode, and neurodiagnostic discovery. IEEE Rev Biomed Eng 5, 60–73. https://doi.org/10.1109/RBME.2012.2211076

Calhoun, V.D., Adali, T., Pearlson, G.D., Pekar, J.J., 2001. A method for making group inferences from functional MRI data using independent component analysis. Hum Brain Mapp 14, 140–151. https://doi.org/10.1002/hbm.1048

Calhoun, V.D., Kiehl, K.A., Pearlson, G.D., 2008. Modulation of temporally coherent brain networks estimated using ICA at rest and during cognitive tasks. Hum Brain Mapp 29, 828–838. https://doi.org/10.1002/hbm.20581

Cordes, D., Haughton, V., Carew, J.D., Arfanakis, K., Maravilla, K., 2002. Hierarchical clustering to measure connectivity in fMRI resting-state data. Magn Reson Imaging 20, 305–317. https://doi.org/10.1016/S0730-725X(02)00503-9

Craddock, R.C., James, G.A., Holtzheimer, P.E., Hu, X.P., Mayberg, H.S., 2012. A whole brain fMRI atlas generated via spatially constrained spectral clustering. Hum Brain Mapp 33, 1914–1928. https://doi.org/10.1002/hbm.21333

Dadi, K., Varoquaux, G., Machlouzarides-Shalit, A., Gorgolewski, K.J., Wassermann, D., Thirion, B., Mensch, A., 2020. Fine-grain atlases of functional modes for fMRI analysis. Neuroimage 221. https://doi.org/10.1016/j.neuroimage.2020.117126

Damoiseaux, J.S., Rombouts, S.A.R.B., Barkhof, F., Scheltens, P., Stam, C.J., Smith, S.M., Beckmann, C.F., 2006. Consistent resting-state networks across healthy subjects. Proc Natl Acad Sci U S A 103, 13848–13853. https://doi.org/10.1073/pnas.0601417103

Deco, G., Kringelbach, M.L., 2017. Hierarchy of Information Processing in the Brain: A Novel 'Intrinsic Ignition' Framework. Neuron 94, 961–968. https://doi.org/10.1016/j.neuron.2017.03.028

Elliott, L.T., Sharp, K., Alfaro-Almagro, F., Shi, S., Miller, K.L., Douaud, G., Marchini, J., Smith, S.M., 2018. Genome-wide association studies of brain imaging phenotypes in UK Biobank. Nature 562, 210–216. https://doi.org/10.1038/s41586-018-0571-7

Farahibozorg, S.-R., Bijsterbosch, J.D., Gong, W., Jbabdi, S., Smith, S.M., Harrison, S.J., Woolrich, M.W., 2021. Hierarchical modelling of functional brain networks in population and individuals from big fMRI data. Neuroimage. https://doi.org/10.1016/j.neuroimage.2021.118513

Finn, E.S., Shen, X., Scheinost, D., Rosenberg, M.D., Huang, J., Chun, M.M., Papademetris, X., Constable, R.T., 2015. Functional connectome fingerprinting: Identifying individuals using patterns of brain connectivity. Nat Neurosci 18, 1664–1671. https://doi.org/10.1038/nn.4135

Friston, K., 2008. Hierarchical models in the brain. PLoS Comput Biol 4, e1000211. https://doi.org/10.1371/journal.pcbi.1000211

Gazzaley, A., Nobre, A.C., 2012. Top-down modulation: Bridging selective attention and working memory. Trends Cogn Sci 16, 129–135. https://doi.org/10.1016/j.tics.2011.11.014

Glasser, M.F., Coalson, T.S., Robinson, E.C., Hacker, C.D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C.F., Jenkinson, M., Smith, S.M., Van Essen, D.C., 2016. A multi-modal parcellation of human cerebral cortex. Nature 536, 171–178. https://doi.org/10.1038/nature18933

Glasser, M.F., Sotiropoulos, S.N., Wilson, J.A., Coalson, T.S., Fischl, B., Andersson, J.L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J.R., Van Essen, D.C., Jenkinson, M., 2013. The minimal preprocessing pipelines for the Human Connectome Project. Neuroimage 80, 105–124. https://doi.org/10.1016/j.neuroimage.2013.04.127

Gong, W., Beckmann, C.F., Smith, S.M., 2021. Phenotype discovery from population brain imaging. Med Image Anal. https://doi.org/10.1016/j.media.2021.102050

Goyal, N., Moraczewski, D., Bandettini, P.A., Finn, E.S., Thomas, A.G., 2022. The positive-negative mode link between brain connectivity, demographics and behaviour: a pre-registered replication of Smith et al. (2015). R Soc Open Sci. https://doi.org/10.1098/rsos.201090

Groves, A.R., Beckmann, C.F., Smith, S.M., Woolrich, M.W., 2011. Linked independent component analysis for multimodal data fusion. Neuroimage 54, 2198–2217. https://doi.org/10.1016/j.neuroimage.2010.09.073

Haak, K. V., Marquand, A.F., Beckmann, C.F., 2018. Connectopic mapping with resting-state fMRI. Neuroimage. https://doi.org/10.1016/j.neuroimage.2017.06.075

Han, Y., Lin, Q.-H., Kuang, L.-D., Zhao, B.-H., Gong, X.-F., Cong, F., Wang, Y.-P., Calhoun, V.D., 2024. A core tensor sparsity enhancement method for solving Tucker-2 model of multi-subject fMRI data. Biomed Signal Process Control 95, 106471. https://doi.org/10.1016/j.bspc.2024.106471

Harrison, S.J., Bijsterbosch, J.D., Segerdahl, A.R., Fitzgibbon, S.P., Farahibozorg, S.-R., Duff, E., Smith, S.M., Woolrich, M.W., 2020. Modelling subject variability in the spatial and temporal characteristics of functional modes. Neuroimage 222, 117226. https://doi.org/10.1016/j.neuroimage.2020.117226

Herculano-Houzel, S., 2009. The human brain in numbers: A linearly scaled-up primate brain. Front Hum Neurosci. https://doi.org/10.3389/neuro.09.031.2009

Jiang, R., Calhoun, V.D., Fan, L., Zuo, N., Jung, R., Qi, S., Lin, D., Li, J., Zhuo, C., Song, M., Fu, Z., Jiang, T., Sui, J., 2020. Gender Differences in Connectome-based Predictions of Individualized Intelligence Quotient and Sub-domain Scores. Cerebral Cortex 30, 888–900. https://doi.org/10.1093/cercor/bhz134

Kiviniemi, V., Starck, T., Remes, J., Long, X., Nikkinen, J., Haapea, M., Veijola, J., Moilanen, I., Isohanni, M., Zang, Y.F., Tervonen, O., 2009. Functional segmentation of the brain cortex using high model order group PICA. Hum Brain Mapp 30, 3865–3886. https://doi.org/10.1002/hbm.20813

Kong, R., Li, J., Orban, C., Sabuncu, M.R., Liu, H., Schaefer, A., Sun, N., Zuo, X.N., Holmes, A.J., Eickhoff, S.B., Yeo, B.T.T., 2019. Spatial Topography of Individual-Specific Cortical Networks Predicts Human Cognition, Personality, and Emotion. Cerebral Cortex 29, 2533–2551. https://doi.org/10.1093/cercor/bhy123

Li, J., Liu, Y., Wisnowski, J.L., Leahy, R.M., 2023. Identification of overlapping and interacting networks reveals intrinsic spatiotemporal organization of the human brain. Neuroimage 270, 119944. https://doi.org/10.1016/j.neuroimage.2023.119944

Lipkin, B., Tuckute, G., Affourtit, J., Small, H., Mineroff, Z., Kean, H., Jouravlev, O., Rakocevic, L., Pritchett, B., Siegelman, M., Hoeflin, C., Pongos, A., Blank, I.A., Struhl, M.K., Ivanova, A., Shannon, S., Sathe, A., Hoffmann, M., Nieto-Castañón, A., Fedorenko, E., 2022. Probabilistic atlas for the language network based on precision fMRI data from >800 individuals. Sci Data 9, 529. https://doi.org/10.1038/s41597-022-01645-3

Mairal, J., Bach, F., Ponce, J., Sapiro, G., 2010. Online learning for matrix factorization and sparse coding. Journal of Machine Learning Research 11, 19–60.

Margulies, D.S., Ghosh, S.S., Goulas, A., Falkiewicz, M., Huntenburg, J.M., Langs, G., Bezgin, G., Eickhoff, S.B., Castellanos, F.X., Petrides, M., Jefferies, E., Smallwood, J., 2016. Situating the default-mode network along a principal gradient of macroscale cortical organization. Proc Natl Acad Sci U S A 113, 12574–12579. https://doi.org/10.1073/pnas.1608282113

McKeown, M.J., Makeig, S., Brown, G.G., Jung, T.P., Kindermann, S.S., Bell, A.J., Sejnowski, T.J., 1998. Analysis of fMRI data by blind separation into independent spatial components. Hum Brain Mapp 6, 160–188. https://doi.org/10.1002/(SICI)1097-0193(1998)6:3<160::AID-HBM5>3.0.CO;2-1

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, É., 2011. Scikit-learn : Machine Learning in Python. Journal of Machine Learning Research 12, 2825–2830.

Pervaiz, U., Vidaurre, D., Woolrich, M.W., Smith, S.M., 2020. Optimising network modelling methods for fMRI. Neuroimage 211. https://doi.org/10.1016/j.neuroimage.2020.116604

Power, J.D., Cohen, A.L., Nelson, S.M., Wig, G.S., Barnes, K.A., Church, J.A., Vogel, A.C., Laumann, T.O., Miezin, F.M., Schlaggar, B.L., Petersen, S.E., 2011. Functional Network Organization of the Human Brain. Neuron. https://doi.org/10.1016/j.neuron.2011.09.006

Raichle, M.E., MacLeod, A.M., Snyder, A.Z., Powers, W.J., Gusnard, D.A., Shulman, G.L., 2001. A default mode of brain function. Proc Natl Acad Sci U S A 98, 676–682. https://doi.org/10.1073/pnas.98.2.676

Robinson, E.C., Jbabdi, S., Glasser, M.F., Andersson, J., Burgess, G.C., Harms, M.P., Smith, S.M., Van Essen, D.C., Jenkinson, M., 2014. MSM: A new flexible framework for multimodal surface matching. Neuroimage 100, 414–426. https://doi.org/10.1016/j.neuroimage.2014.05.069

Schaefer, A., Kong, R., Gordon, E.M., Laumann, T.O., Zuo, X.N., Holmes, A.J., Eickhoff, S.B., Yeo, B.T.T., 2018. Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity mri. Cerebral Cortex 28, 3095–3114. https://doi.org/10.1093/cercor/bhx17

Smith, S.M., Beckmann, C.F., Andersson, J., Auerbach, E.J., Bijsterbosch, J., Douaud, G., Duff, E., Feinberg, D.A., Griffanti, L., Harms, M.P., Kelly, M., Laumann, T., Miller, K.L., Moeller, S., Petersen, S., Power, J., Salimi-Khorshidi, G., Snyder, A.Z., Vu, A.T., Woolrich, M.W., Xu, J., Yacoub, E., Uğurbil, K., Van Essen, D.C., Glasser, M.F., 2013a. Resting-state fMRI in the Human Connectome Project. Neuroimage 80, 144–168. https://doi.org/10.1016/j.neuroimage.2013.05.039

Smith, S.M., Fox, P.T., Miller, K.L., Glahn, D.C., Fox, P.M., Mackay, C.E., Filippini, N., Watkins, K.E., Toro, R., Laird, A.R., Beckmann, C.F., 2009. Correspondence of the brain's functional architecture during activation and rest. Proc Natl Acad Sci U S A 106, 13040–13045. https://doi.org/10.1073/pnas.0905267106

Smith, S.M., Hyvärinen, A., Varoquaux, G., Miller, K.L., Beckmann, C.F., 2014. Group-PCA for very large fMRI datasets. Neuroimage 101, 738–749. https://doi.org/10.1016/j.neuroimage.2014.07.051

Smith, S.M., Nichols, T.E., Vidaurre, D., Winkler, A.M., Behrens, T.E.J., Glasser, M.F., Ugurbil, K., Barch, D.M., Van Essen, D.C., Miller, K.L., 2015. A positive-negative mode of population covariation links brain connectivity, demographics and behavior. Nat Neurosci. https://doi.org/10.1038/nn.4125

Smith, S.M., Vidaurre, D., Beckmann, C.F., Glasser, M.F., Jenkinson, M., Miller, K.L., Nichols, T.E., Robinson, E.C., Salimi-Khorshidi, G., Woolrich, M.W., Barch, D.M., Uğurbil, K., Van Essen, D.C., 2013b. Functional connectomics from resting-state fMRI. Trends Cogn Sci. https://doi.org/10.1016/j.tics.2013.09.016

Snoek, L., Miletić, S., Scholte, H.S., 2019. How to control for confounds in decoding analyses of neuroimaging data. Neuroimage 184, 741–760. https://doi.org/10.1016/j.neuroimage.2018.09.074

Sporns, O., Betzel, R.F., 2016. Modular Brain Networks. Annu Rev Psychol 67, 613–640. https://doi.org/10.1146/annurev-psych-122414-033634

Yeo, B.T.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M., Roffman, J.L., Smoller, J.W., Zöllei, L., Polimeni, J.R., Fisch, B., Liu, H., Buckner, R.L., 2011. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J Neurophysiol 106, 1125–1165. https://doi.org/10.1152/jn.00338.2011

Vidaurre, D., Smith, S.M., Woolrich, M.W., 2017. Brain network dynamics are hierarchically organized in time. Proceedings of the National Academy of Sciences 114, 201705120. https://doi.org/10.1073/pnas.1705120114

Winkler, A.M., Webster, M.A., Vidaurre, D., Nichols, T.E., Smith, S.M., 2015. Multi-level block permutation. Neuroimage. https://doi.org/10.1016/j.neuroimage.2015.05.092

Woolrich, M.W., Behrens, T.E.J., Smith, S.M., 2004. Constrained linear basis sets for HRF modelling using Variational Bayes. Neuroimage 21, 1748–1761. https://doi.org/10.1016/j.neuroimage.2003.12.024

7 Supplementary Materials

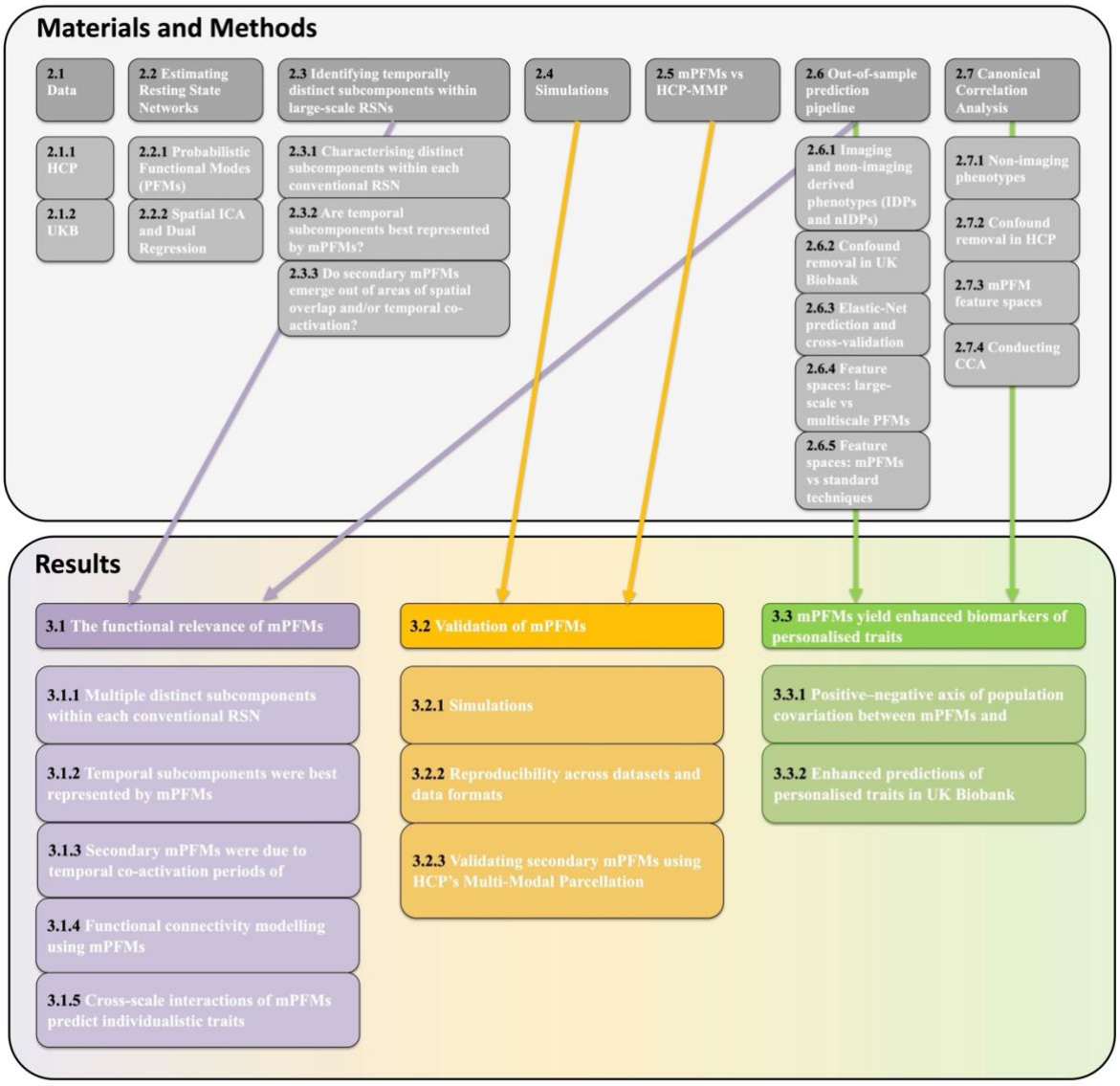

Figure S 1 A Flowchart of different subsections of Methods and Results, as well as the key analysis steps feeding into each Results subsection. Note that since Data and Estimating RSNs feed into multiple Results sections, they are not linked via arrows to avoid unnecessary business of the graph.

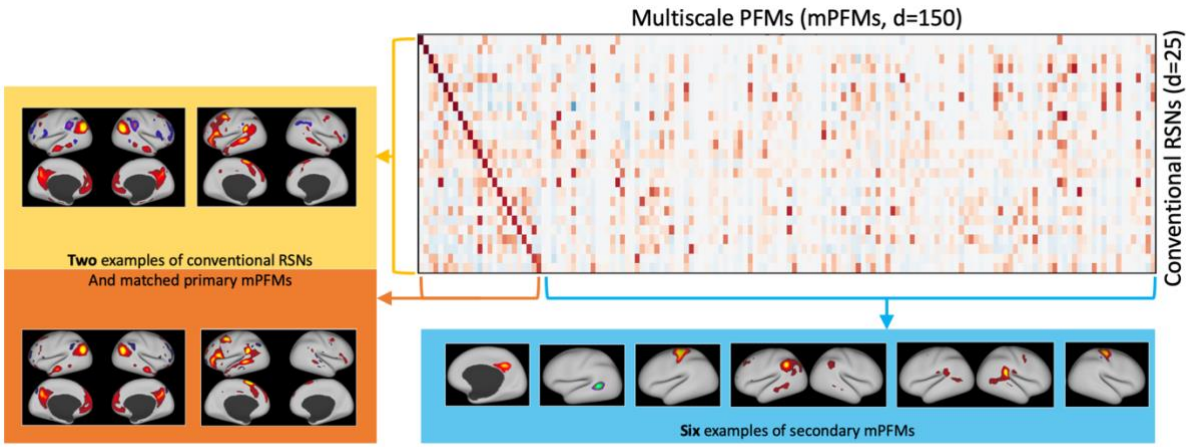

Figure S 2 Supplement to Figure 1: This figure provides an alternative visualisation of the comparison between multiscale Probabilistic Functional Modes (mPFMs) and conventional large-scale RSNs. This visualisation provides additional clarification that when mPFMs are spatially paired with large-scale RSNs from 25-mode decomposition (yellow), 25 of the mPFMs show a clear one-to-one matching, labelled as Primary mPFMs (orange). Secondary mPFMs (blue) are the remaining 125 mPFMs, that start appearing with increased dimensionality.

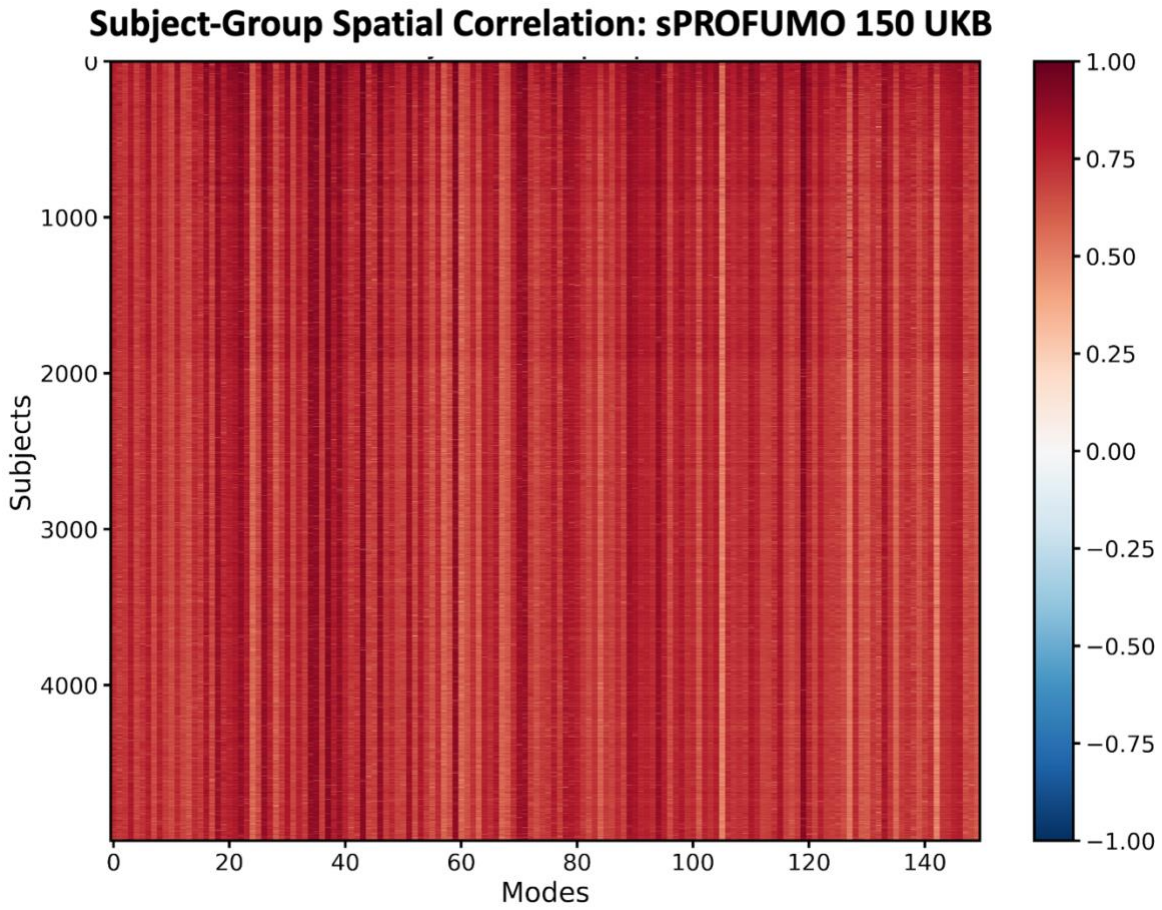

Figure S 3 An example of Subject-group similarity of spatial maps obtained from 150-PFM decomposition of rfMRI data from 5000 subjects in UK Biobank (UKB). This can be used to examine which subject-PFM pairs have been reliably estimated. Bad subject-PFMs can be identified using two heuristic criteria: firstly, if the spatial correlation to the group is very small (<0.1), this indicates a very noisy estimation of that specific PFM for that subject – which would be worth considering removal. Secondly, if the spatial correlation is artefactually high (>0.99), this indicates that posteriors were not reliably estimable at subject level, and therefore that subject-PFM was reverted to group priors – such modes should be removed. In general, in the example shown here, all the modes were estimated reliably for all the subjects. It is worth noting that some PFMs are generally of higher SNR (e.g., Visual or Motor), and they will be more consistent across subjects (columns with darker red colours), whereas other lower-SNR PFMs will be less consistent across subjects (columns with lighter red colours).

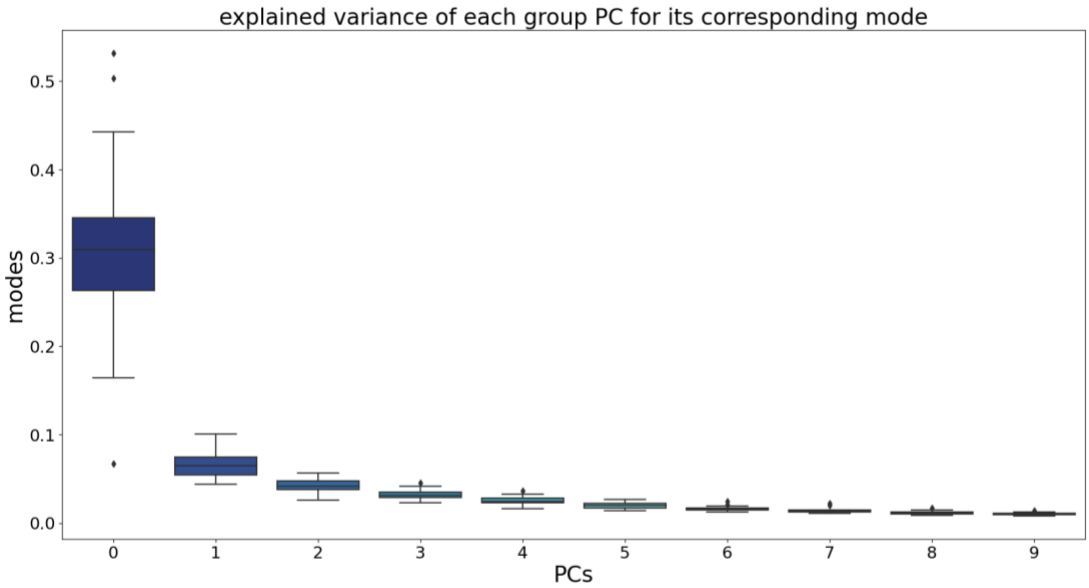

Figure S 4 Supplement to Section 3.1.1: Principal Component Analysis was used to test if one temporal component is sufficient to capture temporal variability within conventional large-scale RSNs from lowD decompositions of rfMRI. LowD refers to 25 modes obtained from PFM decomposition of rfMRI in HCP. The top 10 PCs per lowD mode are illustrated on the x-axis, and their explained variance on the y-axis. Boxplots show median and confidence intervals across 25 modes.

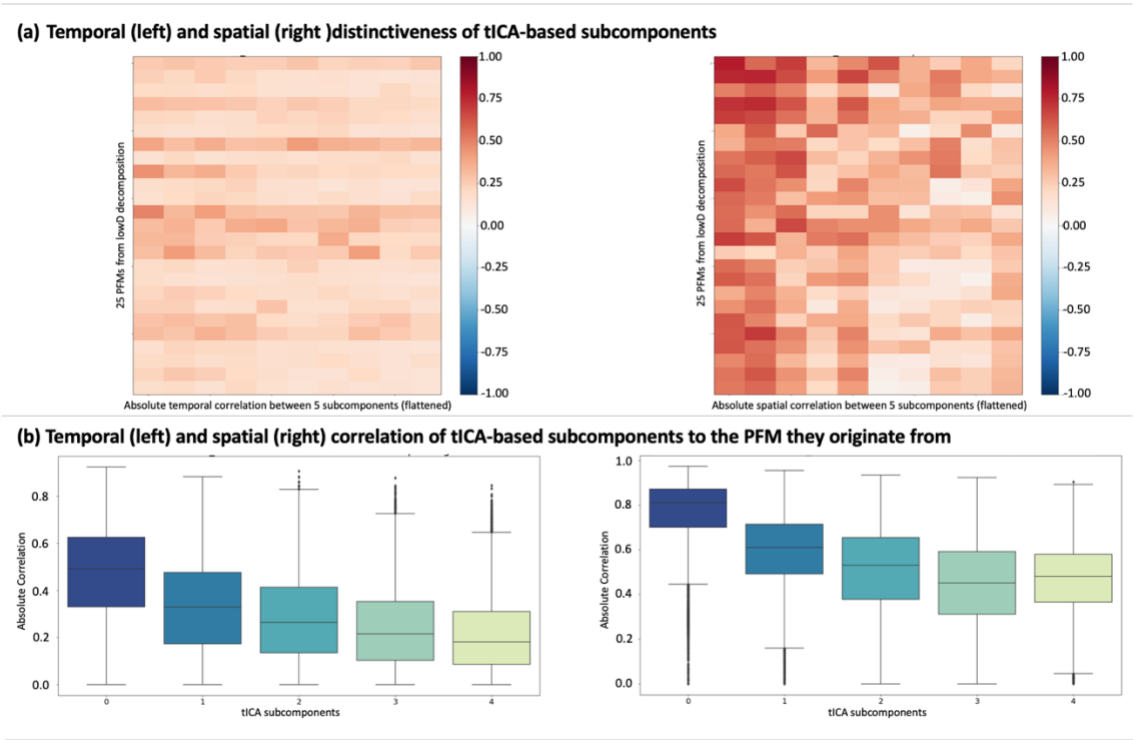

Figure S 5 Supplement to section 3.1.1: a) The degree to which 5x25 subcomponents estimated using temporal ICA are spatially (right) and temporally (left) correlated with each other (i.e. testing for distinctiveness). The 25 rows correspond to the 25 modes, and the flattened above-diagonal of 5x5 correlation matrices (correlations between subcomponents estimated within-subject and then averaged over subjects) appear on the x-axis. b) Temporal (left) and spatial (right) correlation of 5x25 subcomponents with their respective lowD modes. Boxplots show median and confidence intervals across 25 modes.

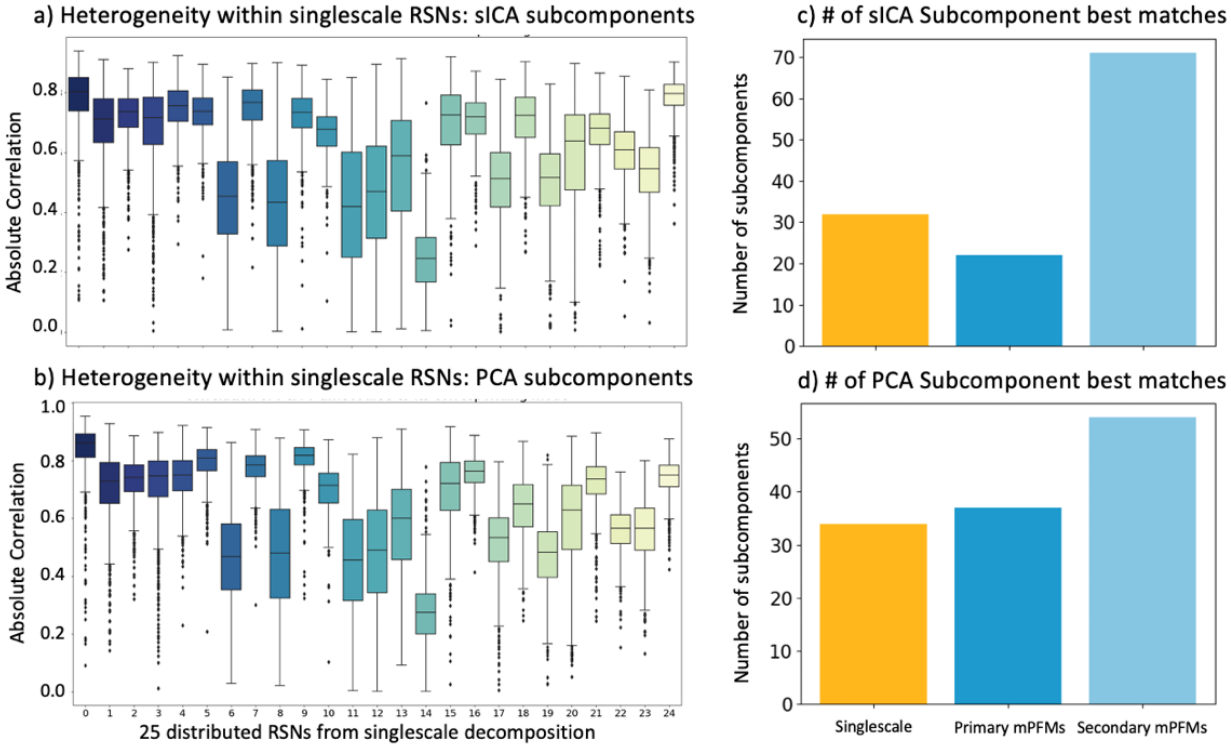

Figure S 6 Supplement to Section 3.1.1: A low-dimensional (lowD) PFM decomposition consisting of 25 modes was estimated from cortical rfMRI data of 1003 HCP subjects, yielding conventional large-scale RSNs, and follow-up spatial ICA (top) and PCA (bottom) applied to voxel-wise timeseries within these RSNs to identify temporally-distinct subcomponents. This analysis was conducted to confirm that subcomponent identification results in section 3.1.1 are not specific to the choice of temporal ICA as the subcomponent identification technique. Temporal correlation of the best-matching a) sICA and b) PCA subcomponents to the large-scale RSNs that they originated from; A winner-takes-all approach was applied: c) of the 25x5 sICA subcomponents 32, 22 and 71 were best represented by large-scale decomposition, primary and secondary mPFMs, respectively; d) of the 25x5 PCA subcomponents 34, 37 and 54 were best represented by large-scale decomposition, primary and secondary mPFMs, respectively.

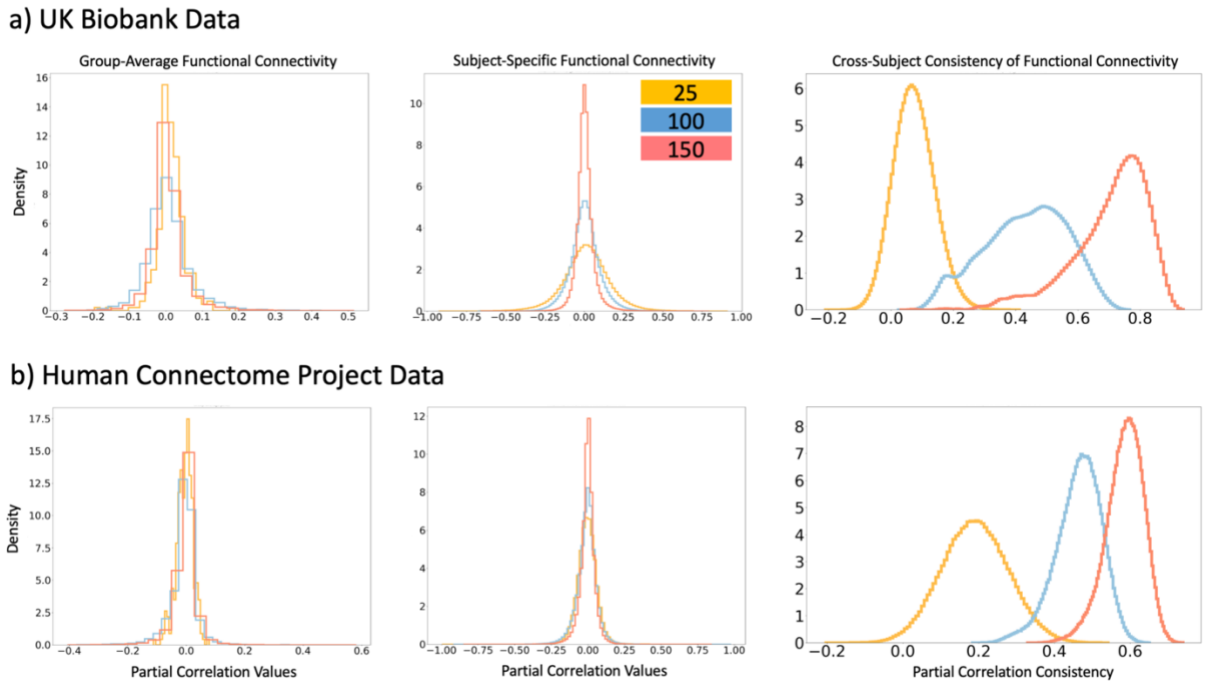

Figure S 7 Supplement to section 3.1.4: Functional connectivity values between modes, estimated using regularised partial correlations. Three PFM dimensionalities are compared: 25, 100 and 150 based on a) 4999 subjects in UK Biobank data (volumetric fMRI) and b) 1003 subjects in Human Connectome Project data (cortical CIFTI). Left: distributions of group average Functional connectivity values. Middle: distributions of subject specific functional connectivity values. Right: Cross subject consistencies of Functional connectivity, which is calculated as Pearson correlation coefficient between vectorised functional connectivity matrices across subjects. Subject-specific functional connectivity becomes sparser and more consistent acorss subjects in mPFMs, leading to less sparse group-average functional connectivity.

28 HCP modes with reproducibility score <0.6 in UKB

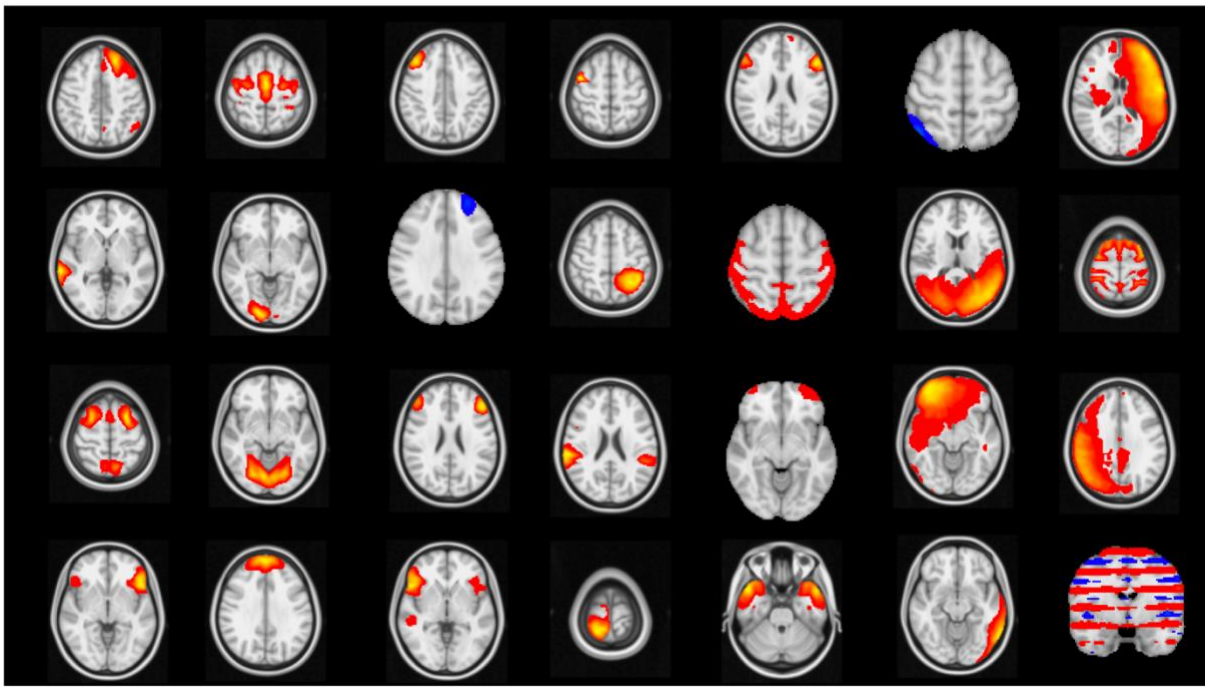

Figure S 8 Supplement to Figure 5: 28 Volumetric HCP modes with reproducibility scores of <0.6 in UKB. Several of these modes were artefactual.

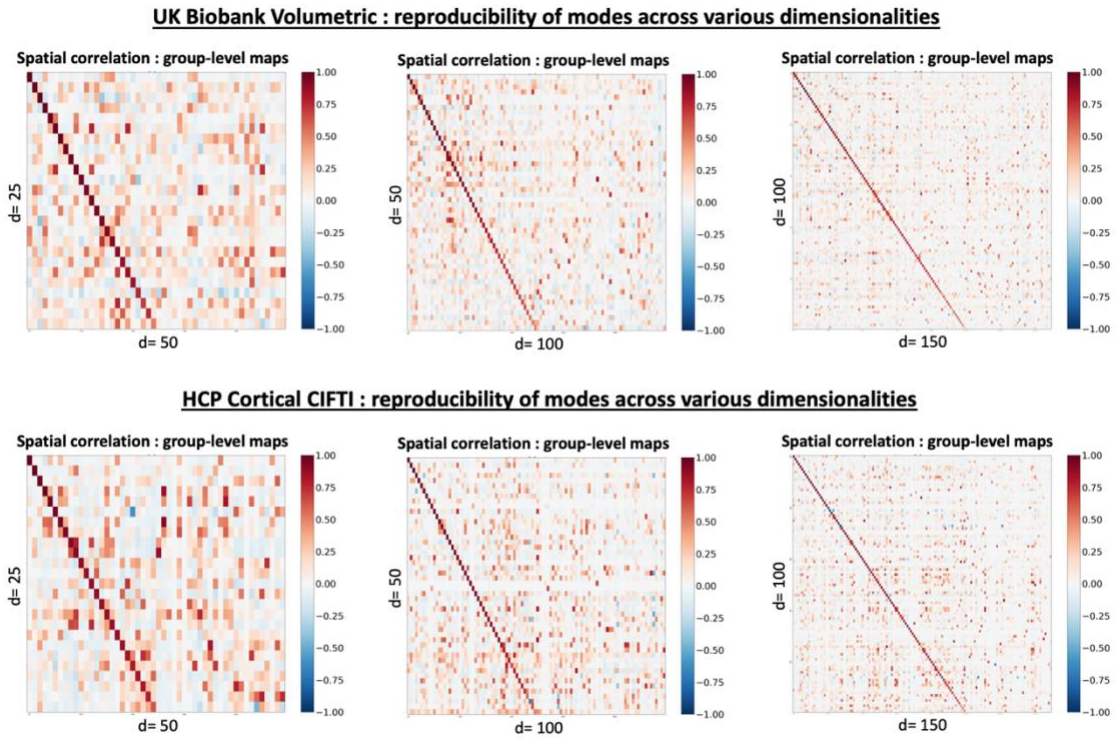

Figure S 9 Supplement to section 3.2.2.1. Reproducibility of mPFMs across multiple dimensionalities in two independent datasets, UK Biobank and HCP: starting from low-dimensional d=25 decomposition, we compared 25 vs 50, 50 vs 100 and 100 vs 150. The modes are sorted and paired based on spatial correlations between the group-level spatial maps. Strong diagonal elements show that, as we increase the dimensionality, the existing modes at lower dimensions are largely maintained, and new modes are added to the ensemble. This indicated an eventual addition of more and more secondary PFMs. Mean±STD of the diagonal values: for the top panel, UK Biobank volumetric fMRI: **25 vs 50**: 0.90 ± 0.11, **50 vs 100**: 0.84 ± 0.14, **100 vs 150**: 0.82 ± 0.14. For the bottom panel, HCP Cortical CIFTI fMRI: **25 vs 50**: 0.88 ± 0.08, **50 vs 100**: 0.90 ± 0.11, **100 vs 150**: 0.88 ± 0.11.

1397

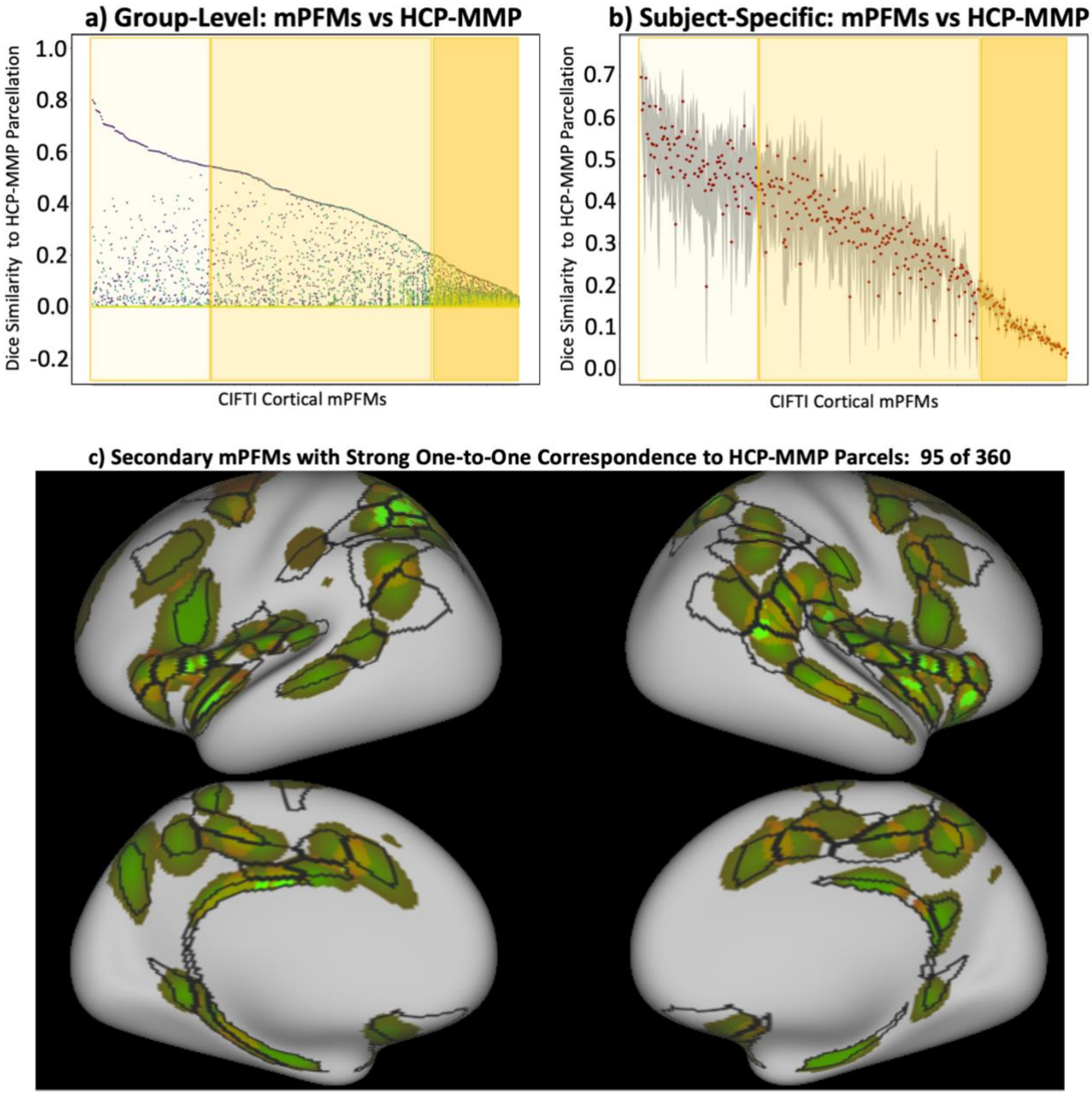

Figure S 10 Supplement to section 3.2.3: a) group-level HCP-MMP parcellation with 360 parcels compared to a 360-mPFM decomposition of HCP. 95 modes were parcel-like (lightest shade of yellow) and showed a clear one-to-one match to HCP-MMP parcels. The remaining large-scale and mixed-scale mPFMs are marked with dark and medium shades of yellow, respectively. 73 mPFMs showed less than 0.2 dice similarity to any HCP-MMP parcels (highlighted with the darkest shade of yellow); these included conventional low-dimensional modes and the less well-known variants of large-scale modes, and typically occupied multiple distant sub-regions. b) subject-level HCP MMP (with the same ordering as Fig5a). Median showed in red, and grey error margins show 25 to 75 percentile across subjects. Good pairing for 95 parcel-like modes were observed here as well. c) Top 95 matched pairs of modes between mPFM and HCP-MMP. Black contours denote group-level HCP-MMP parcels and green-brown patches denote thresholded group-level PFMs.

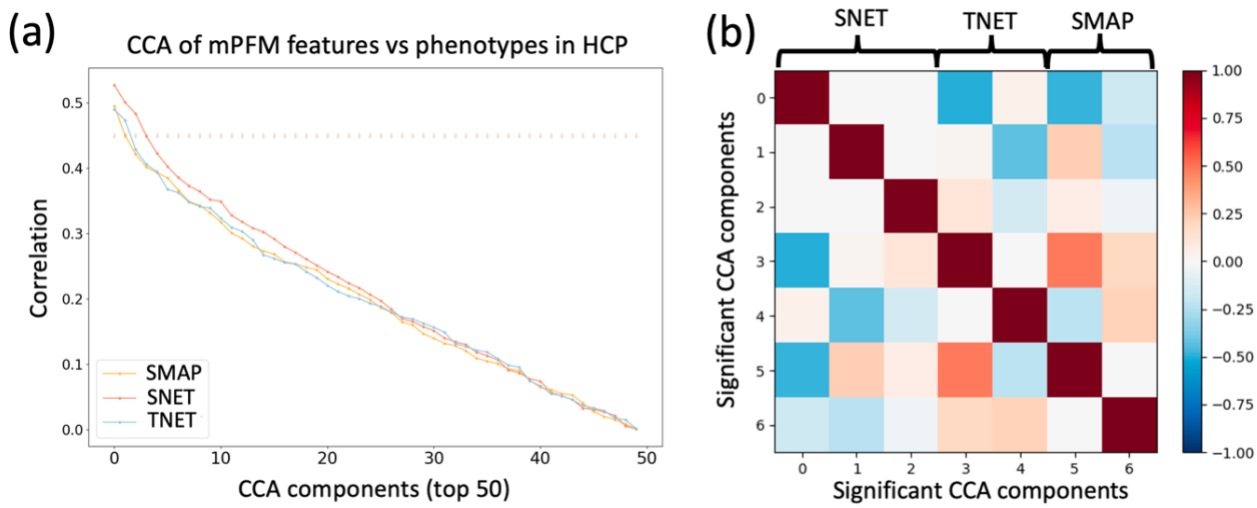

Figure S 11 Supplement to Figure 6a: shared population covariation between mPFMs and behaviour in HCP: a) Canonical Correlation Analysis (CCA) was used to compute modes of population co-variation between behavioural traits and spatial maps (SMAPs)/Spatial correlations (SNETs)/temporal correlations (TNETs) of mPFMs. Statistical significance of CCA components was determined using multi-level block permutations that takes family structure of HCP data into account (Winkler et al., 2015). Correlation values for the top CCA components of SMAP, SNET and TNET were found to be 0.49, 0.52 and 0.49, respectively. Using the recently-introduced GEMMR tool (Helmer et al., 2024), we verified that given these correlation values, our utilised sample size (1001) is sufficient to yield at least 85% power and at most 10% error for the association strength, weight, scores and loadings; b) 2, 3 and 2 significant CCA components were identified for SMAPs, SNETs and TNETs, respectively, and their correlations are shown. we found maximum correlations between transformed pairs SNET-SMAP, SNET-TNET, and SMAP-TNET to be -0.498, -0.484 and 0.483, respectively, indicating that the phenotype-transformed SMAPs, SNETs, and TNETs capture distinct aspects of subject variability.

1425

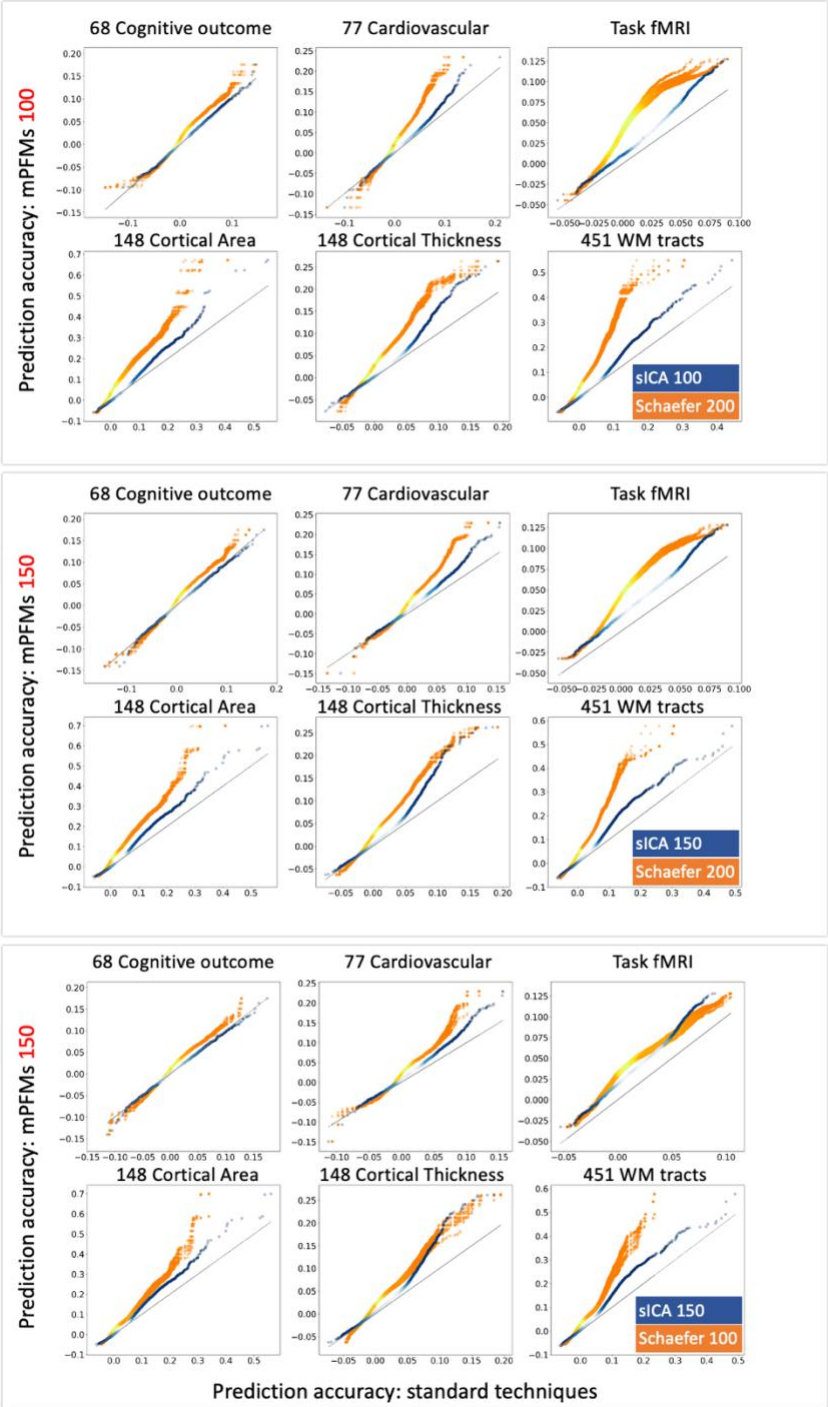

Figure S 12 Supplement to Figure 6b: replicating trait prediction results of mPFMs vs standard techniques using different choices of dimensionality. Top panel: comparing mPFM 100 with spatial ICA (sICA, blue) 100 and Schaefer (orange) 200; middle panel: comparing mPFM 150 with sICA 150 and Schaefer 200; bottom panel: comparing mPFM 150 with sICA 150 and Schaefer 100. In all three scenarios, similar to the original result where d=100 was used for all three methods, mPFMs significantly outperformed the standard techniques in trait predictions. This was especially evident in comparison to Schaefer parcellation. Note that when the number of modes is mismatched between mPFMs/sICA and Schaefer atlas, we have randomly selected a subset of the modes/parcels (e.g., 100 out of 150 in the bottom panel), and repeated that sub-selection 10 times for plotting (hence 10 orange graphs vs 1 blue, given that sICA and mPFM dimensionalities are always matched).

Table S 1 A summary of terminology used throughout the paper.

Terminology	Description
Functional Modes	Modes of functional brain connectivity, where one or multiple brain regions work in synchrony such that they can be described with a single timecourse.
PFMs	Probabilistic Functional Modes: functional modes estimated using sPROFUMO software, which estimates the modes simultaneously for big populations and individuals, thus yielding *personalised* functional modes.
RSNs	Resting State Networks: functional modes estimated from resting state fMRI (rfMRI). Resting state refers to brain activity from task-free recordings, where participants are instructed to not think about anything specific while lying in the scanner.
Large-scale or Distributed RSNs/PFMs	25 modes from low-dimensional (lowD) decomposition of rfMRI data using PFMs, unless specified otherwise (e.g., in Figure 2).
Multiscale PFMs (mPFMs)	Modes from high-dimensional (highD) decomposition of rfMRI data using PFMs. Depending on dataset/application, we have used different number of modes as highD decomposition. Specifically, for HCP data we have used 50, 100, 150 or 360 modes, for UK Biobank data we have used 150 or 100 modes. Justifications are included the corresponding Methods subsections.
Primary mPFMs	25 distributed modes from high-dimensional PFM decomposition of rfMRI that best-match the large-scale RSNs.
Secondary mPFMs	New modes from high-dimensional PFM decomposition that emerge as we increase the dimensionality to >=100.
Spatial Maps, or SMAPs	Spatial organisation of functional modes across the brain voxels.
Functional Connectivity	A matrix of size $N_{mode} \times N_{mode}$ including partial temporal correlations between timecourses of functional modes, also referred to as TNET. L2/Tikhonov regularisation is used in estimation of partial correlations.
Spatial Connectivity	A matrix of size $N_{mode} \times N_{mode}$ including spatial correlation coefficients between SMAPs of functional modes, also referred to as SNET. This is an indicator of *spatial overlap* between the modes.

Table S 2 T-value and p-values related to subcomponent analysis: original subcomponents. Statistically significant p-values (Bonferroni-corrected threshold to account for multiple comparisons) are shown in bold font. Supplement to **Figure 3**c&d.

Winner: Large-Scale		Winner: Primary mPFMs		Winner: Secondary mPFMs			
t-values	**p-values**	**t-values**	**p-values**	**t-values**	**p-values**	**t-values**	**p-values**
0.46	6.42e-01	19.81	**5.57e-74**	22.96	**4.70e-94**	23.08	**7.82e-95**
40.45	**7.04e-213**	17.29	**8.47e-59**	14.19	**8.30e-42**	16.34	**2.07e-53**
1.86	6.35e-02	22.53	**2.74e-91**	17.81	**7.82e-62**	29.73	**8.81e-140**
7.60	**6.64e-14**	14.21	**7.29e-42**	4.64	**4.01e-06**	17.66	**5.87e-61**
36.36	**3.06e-185**	10.26	**1.54e-23**	28.59	**6.15e-132**	-0.61	5.43e-01
37.73	**1.35e-194**	5.33	**1.21e-07**	31.04	**9.76e-149**	24.86	**1.12e-106**
15.01	**4.24e-46**	44.60	**3.17e-240**	19.73	**1.70e-73**	15.69	**8.63e-50**
-2.64	8.33e-03	15.12	**1.07e-46**	10.46	**2.31e-24**	10.89	**3.61e-26**
7.04	**3.49e-12**	14.02	**6.60e-41**	24.87	**1.03e-106**	-0.01	9.94e-01
7.98	**4.06e-15**	57.85	**1.13e-321**	10.30	**1.00e-23**	4.77	**2.08e-06**
35.48	**3.05e-179**	23.84	**7.72e-100**	5.45	**6.42e-08**	11.08	**5.24e-27**
15.05	**2.63e-46**	14.47	**3.29e-43**	12.29	**1.95e-32**	40.87	**1.13e-215**
		4.88	**1.26e-06**	4.77	**2.08e-06**	32.10	**5.01e-156**
		16.06	**8.24e-52**	16.35	**1.92e-53**	18.47	**8.46e-66**
		7.10	**2.36e-12**	51.75	**2.76e-285**	26.74	**2.64e-119**
		37.67	**3.76e-194**	9.99	**1.76e-22**	26.53	**7.10e-118**
		10.92	**2.62e-26**	27.84	**8.20e-127**	16.47	**4.19e-54**
		3.23	**1.27e-03**	52.06	**3.32e-287**	35.81	**1.68e-181**
		15.55	**4.96e-49**	17.32	**5.24e-59**	39.36	**1.46e-205**
		17.40	**1.91e-59**	44.44	**3.60e-239**	61.30	**0.00e+00**
		6.98	**5.22e-12**	10.31	**9.63e-24**	29.06	**3.88e-135**
		33.82	**7.52e-168**	31.74	**1.39e-153**	17.77	**1.28e-61**
		37.02	**9.35e-190**	24.91	**5.39e-107**	29.35	**3.66e-137**
		26.00	**2.47e-114**	15.64	**1.69e-49**	11.63	**1.99e-29**
		27.83	**9.80e-127**	24.38	**1.78e-103**	13.61	**7.77e-39**
		16.07	**6.90e-52**	21.47	**2.08e-84**	34.87	**4.61e-175**
		19.13	**9.34e-70**	22.20	**3.81e-89**	9.84	**7.32e-22**
		38.03	**1.39e-196**	38.65	**8.15e-201**	54.03	**3.83e-299**
				61.69	**0.00e+00**	23.29	**2.93e-96**
				27.37	**1.24e-123**	25.14	**1.54e-108**
				27.32	**2.92e-123**	14.85	**2.95e-45**
				49.52	**1.59e-271**	74.34	**0.00e+00**
				88.06	**0.00e+00**	6.00	**2.82e-09**
				44.50	**1.58e-239**	14.35	**1.34e-42**
				3.59	**3.42e-04**	19.80	**6.30e-74**
				11.52	**6.05e-29**	24.71	**1.17e-105**
				9.78	**1.22e-21**	19.73	**1.75e-73**
				22.43	**1.26e-90**	13.38	**1.12e-37**
				11.77	**4.55e-30**	13.04	**5.38e-36**
				40.02	**5.40e-210**	62.63	**0.00e+00**
				34.89	**3.22e-175**	16.04	**1.02e-51**
				14.68	**2.54e-44**	37.97	**3.27e-196**
						49.84	**1.71e-273**

Table S 3 T-value and p-values related to subcomponent analysis: temporally-exclusive subcomponents. Statistically significant p-values (Bonferroni-corrected threshold to account for multiple comparisons) are shown in bold font. Supplement to **Figure 3**e&f.

Winner: Large-Scale				Winner: Primary mPFMs		Winner: Secondary mPFMs	
t-values	**p-values**	**t-values**	**p-values**	**t-values**	**p-values**	**t-values**	**p-values**
30.98	**2.52e-148**	46.74	**6.02e-254**	1.52	1.28e-01	6.41	**2.20e-10**
28.70	**1.00e-132**	37.99	**2.40e-196**	3.12	**1.88e-03**	15.88	**8.34e-51**
15.48	**1.20e-48**	4.87	**1.31e-06**	4.21	**2.78e-05**	1.55	1.22e-01
4.11	**4.30e-05**	34.01	**3.38e-169**	1.13	2.59e-01	6.07	**1.82e-09**
41.43	**2.13e-219**	20.80	**3.71e-80**			11.45	**1.27e-28**
1.62	1.06e-01	35.79	**2.27e-181**			11.26	**8.68e-28**
15.50	**9.33e-49**	10.34	**7.19e-24**			1.69	9.20e-02
10.65	**3.63e-25**	10.02	**1.32e-22**				
67.89	**0.00e+00**	1.27	2.04e-01				
16.01	**1.50e-51**	62.64	**0.00e+00**				
26.71	**4.26e-119**	37.78	**6.91e-195**				
7.64	**5.18e-14**	37.84	**2.51e-195**				
42.21	**1.46e-224**	25.98	**3.70e-114**				
15.03	**3.57e-46**	34.11	**7.63e-170**				
24.92	**4.92e-107**	42.65	**1.83e-227**				
6.33	**3.61e-10**	10.44	**2.72e-24**				
22.85	**2.21e-93**	22.93	**6.88e-94**				
60.43	**0.00e+00**	58.42	**0.00e+00**				
25.70	**2.91e-112**	25.22	**4.90e-109**				
5.16	**3.02e-07**	26.83	**6.54e-120**				
27.78	**2.27e-126**	22.02	**5.46e-88**				
8.02	**3.00e-15**	8.93	**1.96e-18**				
50.27	**3.33e-276**	54.66	**6.28e-303**				
33.88	**2.63e-168**	3.63	**2.99e-04**				
27.60	**3.55e-125**	3.79	**1.58e-04**				
17.19	**3.05e-58**	2.38	1.75e-02				
13.68	**3.60e-39**	72.20	**0.00e+00**				
71.00	**0.00e+00**	44.52	**1.04e-239**				
43.53	**3.04e-233**	21.80	**1.47e-86**				
34.39	**9.00e-172**	23.88	**3.76e-100**				
30.32	**8.00e-144**	4.15	**3.59e-05**				
35.19	**2.98e-177**	58.48	**0.00e+00**				
38.02	**1.57e-196**	24.23	**1.77e-102**				
33.46	**2.09e-165**	8.54	**4.82e-17**				
28.81	**1.82e-133**	105.56	**0.00e+00**				
23.01	**2.10e-94**	29.69	**1.62e-139**				
21.48	**1.81e-84**	23.61	**2.41e-98**				
49.70	**1.10e-272**	34.32	**2.69e-171**				
30.70	**2.12e-146**	27.49	**2.07e-124**				
20.28	**6.98e-77**	69.98	**0.00e+00**				
16.30	**3.74e-53**	10.83	**6.31e-26**				
24.46	**5.31e-104**	9.19	**2.15e-19**				
63.10	**0.00e+00**	14.30	**2.33e-42**				
3.10	1.97e-03	4.85	**1.40e-06**				
12.88	**3.22e-35**	38.13	**2.64e-197**				
18.38	**3.02e-65**	23.52	**1.00e-97**				
12.50	**2.12e-33**	11.53	**5.92e-29**				
78.92	**0.00e+00**	74.21	**0.00e+00**				
35.71	**8.46e-181**	18.10	**1.39e-63**				
19.95	**7.70e-75**	42.58	**5.12e-227**				
30.69	**2.35e-146**	31.93	**6.42e-155**				
25.75	**1.29e-112**	38.12	**3.17e-197**				
19.17	**4.85e-70**	91.42	**0.00e+00**				
38.45	**1.86e-199**	29.99	**1.49e-141**				
21.55	**5.77e-85**	30.16	**9.74e-143**				
23.77	**2.17e-99**	27.38	**1.14e-123**				
2.54	1.12e-02	22.64	**5.45e-92**				

Table S 4 T-value and p-values related to subcomponent analysis: spatially-exclusive subcomponents. Statistically significant p-values (Bonferroni-corrected threshold to account for multiple comparisons) are shown in bold font. Supplement to **Figure 3**g&h.

Winner: Large-Scale		**Winner: Primary mPFMs**		**Winner: Secondary mPFMs**			
t-values	**p-values**	**t-values**	**p-values**	**t-values**	**p-values**	**t-values**	**p-values**
5.69	**1.66e-08**	6.72	**2.99e-11**	9.19	**2.26e-19**	34.20	**1.78e-170**
5.38	**9.08e-08**	5.23	**2.08e-07**	7.34	**4.28e-13**	26.13	**3.38e-115**
6.52	**1.13e-10**	15.49	**1.18e-48**	26.05	**1.25e-114**	2.76	5.98e-03
13.65	**5.09e-39**	15.99	**2.08e-51**	34.10	**8.66e-170**	40.76	**5.93e-215**
0.13	8.96e-01	21.76	**2.96e-86**	27.08	**1.23e-121**	18.72	**2.89e-67**
26.21	**9.44e-116**	14.93	**1.17e-45**	43.91	**1.10e-235**	23.28	**3.76e-96**
61.72	**0.00e+00**	10.59	**6.59e-25**	5.19	**2.51e-07**	13.26	**4.26e-37**
22.23	**2.76e-89**	18.96	**1.03e-68**	4.75	**2.30e-06**	53.60	**1.43e-296**
16.48	**3.43e-54**	3.97	**7.59e-05**	49.56	**9.49e-272**	57.32	**1.46e-318**
0.59	5.55e-01	38.36	**7.99e-199**	12.44	**3.84e-33**	31.82	**4.05e-154**
5.10	**4.14e-07**	8.41	**1.43e-16**	31.73	**1.77e-153**	9.44	**2.56e-20**
3.46	**5.65e-04**	19.28	**1.02e-70**	3.03	2.49e-03	14.04	**5.04e-41**
3.33	**9.01e-04**	5.47	**5.58e-08**	55.94	**1.62e-310**	34.32	**2.67e-171**
1.65	9.96e-02	52.83	**6.65e-292**	14.90	**1.68e-45**	26.62	**1.73e-118**
8.39	**1.63e-16**	22.88	**1.58e-93**	4.61	**4.54e-06**	49.63	**3.14e-272**
11.01	**1.13e-26**	12.22	**4.24e-32**	22.67	**3.51e-92**	13.06	**4.35e-36**
19.96	**7.23e-75**	3.86	**1.19e-04**	20.11	**7.89e-76**	26.66	**8.40e-119**
21.98	**1.08e-87**	15.40	**3.47e-48**	1.22	2.21e-01	44.86	**7.14e-242**
11.82	**2.91e-30**	4.54	**6.24e-06**	45.25	**1.96e-244**	26.79	**1.24e-119**
15.23	**2.86e-47**	27.83	**9.10e-127**	15.46	**1.67e-48**	21.68	**8.53e-86**
15.47	**1.43e-48**	27.79	**1.69e-126**	15.53	**6.65e-49**	5.08	**4.61e-07**
		17.66	**6.01e-61**	0.24	8.10e-01	31.60	**1.20e-152**
		27.53	**1.12e-124**	22.53	**2.79e-91**	30.37	**3.89e-144**
		15.86	**1.07e-50**	45.45	**1.05e-245**	52.73	**2.57e-291**
		21.31	**2.11e-83**	3.94	**8.72e-05**	14.26	**3.83e-42**
		18.86	**3.69e-68**	36.17	**5.61e-184**	55.50	**7.10e-308**
		23.85	**6.66e-100**	18.62	**1.05e-66**	13.58	**1.05e-38**
				12.19	**5.60e-32**	12.37	**8.18e-33**
				38.93	**1.22e-202**	11.22	**1.35e-27**
				45.34	**5.24e-245**	25.39	**3.53e-110**
				6.91	**8.36e-12**	24.99	**1.66e-107**
				9.67	**3.17e-21**	18.15	**7.23e-64**
				6.76	**2.32e-11**	25.25	**3.02e-109**
				18.88	**2.96e-68**	75.50	**0.00e+00**
				4.75	**2.31e-06**	8.23	**5.80e-16**
				18.66	**6.15e-67**	38.35	**9.55e-199**
				20.49	**3.29e-78**	14.31	**2.02e-42**
				10.18	**3.28e-23**	27.85	**6.72e-127**
						33.14	**3.38e-163**

Table S 5 T-value and Bonferroni-corrected (36 comparisons) p-values to compare phenotype-prediction performance of Spatial and Temporal/Functional connectivity features related to: a) LowD large-scale RSNs vs highD mPFMs and b) within-scale vs cross-scale mPFMs. Pri: Primary, Sec: Secondary, Statistically significant values are shown in bold font. Supplement to **Figure 4**.

	T-Value (corrected P-Value)					
	Spatial connectivity			**Temporal connectivity**		
	lowD > mPFMs	**Pri2Pri > Pri2Sec**	**Sec2Sec > Pri2Sec**	**lowD > mPFMs**	**Pri2Pri > Pri2Sec**	**Sec2Sec > Pri2Sec**
Cognitive	-1.464 (0.148)	-1.243 (0.219)	**-3.526 (7.92E-04)**	-0.6182 (0.539)	-1.5678 (0.122)	-0.487 (0.628)
GMa	**-21.657 (1.37E-47)**	**-21.017 (3.89E-46)**	**-14.978 (2.21E-31)**	**-4.6115 (8.88E-06)**	**-2.6633 (8.64E-03)**	**3.3587 (1.01E-03)**
GMt	**-9.226 (2.81E-16)**	**-12.042 (1.14E-23)**	**-4.916 (2.34E-06)**	-1.0236 (0.31)	-0.1672 (0.867)	**2.0203 (0.045)**
WM	**-26.469 (6.31E-94)**	**-25.468 (2.19E-89)**	-0.583 (0.56)	1.0486 (0.295)	**-9.8976 (5.09E-21)**	**7.1892 (2.75E-12)**
Task-Contrast	**-15.190 (4.15E-35)**	**-7.165 (1.48E-11)**	**-56.461 (1.94E-124)**	**-7.0706 (2.56E-11)**	-0.6513 (0.516)	**35.8817 (7.73E-89)**
Cardiovascular	**-7.272 (2.68E-10)**	**-4.296 (5.09E-05)**	0.663 (0.51)	0.2176 (0.828)	**-5.4334 (6.59E-07)**	-0.5626 (0.575)

1446

1447 Table S 6 T-values and Bonferroni-corrected (6 comparisons) p-values to compare phenotype-prediction performance of mPFMs vs spatial
1448 ICA (sICA) and Schaefer Parcellation of the same dimensionality (100 modes). Statistically significant values have been shown in bold font.
1449 Supplement to **Figure 6**.

	T-value			Corrected P-Value		
	mPFM > sICA	mPFM > Schaefer	sICA > Schaefer	mPFM vs sICA	mPFM vs Schaefer	sICA vs Schaefer
Cognitive	1.261	**7.385**	**7.252**	0.212	**3.61E-10**	**6.21E-10**
GMa	**15.137**	**26.991**	**23.107**	**8.57E-32**	**7.90E-59**	**8.59E-51**
GMt	**13.964**	**33.923**	**31.499**	**9.64E-29**	**2.09E-71**	**3.06E-67**
WM	**12.675**	**51.843**	**57.347**	**1.07E-31**	**6.96E-192**	**5.19E-209**
Task-Contrast	**14.572**	**14.496**	**5.086**	**1.93E-19**	**2.38E-19**	**5.75E-06**
Cardiovascular	**7.974**	**17.409**	**17.33**	**1.24E-11**	**3.07E-28**	**4.04E-28**

1450

1451

1452 Table S 7 Names of UK Biobank Imaging Derived Phenotypes (IDPs) related to GM area: 148. Supplement to **Figure 6**.

IDP #	IDP Name	IDP #	IDP Name
1	aparc-a2009s_lh_area_G+S-frontomargin	**75**	aparc-a2009s_rh_area_G+S-frontomargin
2	aparc-a2009s_lh_area_G+S-occipital-inf	**76**	aparc-a2009s_rh_area_G+S-occipital-inf
3	aparc-a2009s_lh_area_G+S-paracentral	**77**	aparc-a2009s_rh_area_G+S-paracentral
4	aparc-a2009s_lh_area_G+S-subcentral	**78**	aparc-a2009s_rh_area_G+S-subcentral
5	aparc-a2009s_lh_area_G+S-transv-frontopol	**79**	aparc-a2009s_rh_area_G+S-transv-frontopol
6	aparc-a2009s_lh_area_G+S-cingul-Ant	**80**	aparc-a2009s_rh_area_G+S-cingul-Ant
7	aparc-a2009s_lh_area_G+S-cingul-Mid-Ant	**81**	aparc-a2009s_rh_area_G+S-cingul-Mid-Ant
8	aparc-a2009s_lh_area_G+S-cingul-Mid-Post	**82**	aparc-a2009s_rh_area_G+S-cingul-Mid-Post
9	aparc-a2009s_lh_area_G-cingul-Post-dorsal	**83**	aparc-a2009s_rh_area_G-cingul-Post-dorsal
10	aparc-a2009s_lh_area_G-cingul-Post-ventral	**84**	aparc-a2009s_rh_area_G-cingul-Post-ventral
11	aparc-a2009s_lh_area_G-cuneus	**85**	aparc-a2009s_rh_area_G-cuneus
12	aparc-a2009s_lh_area_G-front-inf-Opercular	**86**	aparc-a2009s_rh_area_G-front-inf-Opercular
13	aparc-a2009s_lh_area_G-front-inf-Orbital	**87**	aparc-a2009s_rh_area_G-front-inf-Orbital
14	aparc-a2009s_lh_area_G-front-inf-Triangul	**88**	aparc-a2009s_rh_area_G-front-inf-Triangul
15	aparc-a2009s_lh_area_G-front-middle	**89**	aparc-a2009s_rh_area_G-front-middle
16	aparc-a2009s_lh_area_G-front-sup	**90**	aparc-a2009s_rh_area_G-front-sup
17	aparc-a2009s_lh_area_G-Ins-lg+S-cent-ins	**91**	aparc-a2009s_rh_area_G-Ins-lg+S-cent-ins
18	aparc-a2009s_lh_area_G-insular-short	**92**	aparc-a2009s_rh_area_G-insular-short
19	aparc-a2009s_lh_area_G-occipital-middle	**93**	aparc-a2009s_rh_area_G-occipital-middle
20	aparc-a2009s_lh_area_G-occipital-sup	**94**	aparc-a2009s_rh_area_G-occipital-sup
21	aparc-a2009s_lh_area_G-oc-temp-lat-fusifor	**95**	aparc-a2009s_rh_area_G-oc-temp-lat-fusifor
22	aparc-a2009s_lh_area_G-oc-temp-med-Lingual	**96**	aparc-a2009s_rh_area_G-oc-temp-med-Lingual
23	aparc-a2009s_lh_area_G-oc-temp-med-Parahip	**97**	aparc-a2009s_rh_area_G-oc-temp-med-Parahip
24	aparc-a2009s_lh_area_G-orbital	**98**	aparc-a2009s_rh_area_G-orbital
25	aparc-a2009s_lh_area_G-pariet-inf-Angular	**99**	aparc-a2009s_rh_area_G-pariet-inf-Angular
26	aparc-a2009s_lh_area_G-pariet-inf-Supramar	**100**	aparc-a2009s_rh_area_G-pariet-inf-Supramar
27	aparc-a2009s_lh_area_G-parietal-sup	**101**	aparc-a2009s_rh_area_G-parietal-sup
28	aparc-a2009s_lh_area_G-postcentral	**102**	aparc-a2009s_rh_area_G-postcentral
29	aparc-a2009s_lh_area_G-precentral	**103**	aparc-a2009s_rh_area_G-precentral
30	aparc-a2009s_lh_area_G-precuneus	**104**	aparc-a2009s_rh_area_G-precuneus
31	aparc-a2009s_lh_area_G-rectus	**105**	aparc-a2009s_rh_area_G-rectus
32	aparc-a2009s_lh_area_G-subcallosal	**106**	aparc-a2009s_rh_area_G-subcallosal
33	aparc-a2009s_lh_area_G-temp-sup-G-T-transv	**107**	aparc-a2009s_rh_area_G-temp-sup-G-T-transv
34	aparc-a2009s_lh_area_G-temp-sup-Lateral	**108**	aparc-a2009s_rh_area_G-temp-sup-Lateral
35	aparc-a2009s_lh_area_G-temp-sup-Plan-polar	**109**	aparc-a2009s_rh_area_G-temp-sup-Plan-polar
36	aparc-a2009s_lh_area_G-temp-sup-Plan-tempo	**110**	aparc-a2009s_rh_area_G-temp-sup-Plan-tempo
37	aparc-a2009s_lh_area_G-temporal-inf	**111**	aparc-a2009s_rh_area_G-temporal-inf
38	aparc-a2009s_lh_area_G-temporal-middle	**112**	aparc-a2009s_rh_area_G-temporal-middle
39	aparc-a2009s_lh_area_Lat-Fis-ant-Horizont	**113**	aparc-a2009s_rh_area_Lat-Fis-ant-Horizont
40	aparc-a2009s_lh_area_Lat-Fis-ant-Vertical	**114**	aparc-a2009s_rh_area_Lat-Fis-ant-Vertical
41	aparc-a2009s_lh_area_Lat-Fis-post	**115**	aparc-a2009s_rh_area_Lat-Fis-post
42	aparc-a2009s_lh_area_Pole-occipital	**116**	aparc-a2009s_rh_area_Pole-occipital
43	aparc-a2009s_lh_area_Pole-temporal	**117**	aparc-a2009s_rh_area_Pole-temporal
44	aparc-a2009s_lh_area_S-calcarine	**118**	aparc-a2009s_rh_area_S-calcarine
45	aparc-a2009s_lh_area_S-central	**119**	aparc-a2009s_rh_area_S-central
46	aparc-a2009s_lh_area_S-cingul-Marginalis	**120**	aparc-a2009s_rh_area_S-cingul-Marginalis
47	aparc-a2009s_lh_area_S-circular-insula-ant	**121**	aparc-a2009s_rh_area_S-circular-insula-ant
48	aparc-a2009s_lh_area_S-circular-insula-inf	**122**	aparc-a2009s_rh_area_S-circular-insula-inf
49	aparc-a2009s_lh_area_S-circular-insula-sup	**123**	aparc-a2009s_rh_area_S-circular-insula-sup
50	aparc-a2009s_lh_area_S-collat-transv-ant	**124**	aparc-a2009s_rh_area_S-collat-transv-ant
51	aparc-a2009s_lh_area_S-collat-transv-post	**125**	aparc-a2009s_rh_area_S-collat-transv-post
52	aparc-a2009s_lh_area_S-front-inf	**126**	aparc-a2009s_rh_area_S-front-inf
53	aparc-a2009s_lh_area_S-front-middle	**127**	aparc-a2009s_rh_area_S-front-middle
54	aparc-a2009s_lh_area_S-front-sup	**128**	aparc-a2009s_rh_area_S-front-sup
55	aparc-a2009s_lh_area_S-interm-prim-Jensen	**129**	aparc-a2009s_rh_area_S-interm-prim-Jensen
56	aparc-a2009s_lh_area_S-intrapariet+P-trans	**130**	aparc-a2009s_rh_area_S-intrapariet+P-trans
57	aparc-a2009s_lh_area_S-oc-middle+Lunatus	**131**	aparc-a2009s_rh_area_S-oc-middle+Lunatus
58	aparc-a2009s_lh_area_S-oc-sup+transversal	**132**	aparc-a2009s_rh_area_S-oc-sup+transversal
59	aparc-a2009s_lh_area_S-occipital-ant	**133**	aparc-a2009s_rh_area_S-occipital-ant
60	aparc-a2009s_lh_area_S-oc-temp-lat	**134**	aparc-a2009s_rh_area_S-oc-temp-lat
61	aparc-a2009s_lh_area_S-oc-temp-med+Lingual	**135**	aparc-a2009s_rh_area_S-oc-temp-med+Lingual
62	aparc-a2009s_lh_area_S-orbital-lateral	**136**	aparc-a2009s_rh_area_S-orbital-lateral
63	aparc-a2009s_lh_area_S-orbital-med-olfact	**137**	aparc-a2009s_rh_area_S-orbital-med-olfact
64	aparc-a2009s_lh_area_S-orbital-H-Shaped	**138**	aparc-a2009s_rh_area_S-orbital-H-Shaped
65	aparc-a2009s_lh_area_S-parieto-occipital	**139**	aparc-a2009s_rh_area_S-parieto-occipital
66	aparc-a2009s_lh_area_S-pericallosal	**140**	aparc-a2009s_rh_area_S-pericallosal
67	aparc-a2009s_lh_area_S-postcentral	**141**	aparc-a2009s_rh_area_S-postcentral
68	aparc-a2009s_lh_area_S-precentral-inf-part	**142**	aparc-a2009s_rh_area_S-precentral-inf-part
69	aparc-a2009s_lh_area_S-precentral-sup-part	**143**	aparc-a2009s_rh_area_S-precentral-sup-part
70	aparc-a2009s_lh_area_S-suborbital	**144**	aparc-a2009s_rh_area_S-suborbital
71	aparc-a2009s_lh_area_S-subparietal	**145**	aparc-a2009s_rh_area_S-subparietal
72	aparc-a2009s_lh_area_S-temporal-inf	**146**	aparc-a2009s_rh_area_S-temporal-inf
73	aparc-a2009s_lh_area_S-temporal-sup	**147**	aparc-a2009s_rh_area_S-temporal-sup
74	aparc-a2009s_lh_area_S-temporal-transverse	**148**	aparc-a2009s_rh_area_S-temporal-transverse

Table S 8 Names of UK Biobank Imaging Derived Phenotypes (IDPs) related to Grey Matter thickness: 148. Supplement to **Figure 6**.

IDP #	IDP Name	IDP #	IDP Name
1	aparc-a2009s_lh_thickness_G+S-frontomargin,	**75**	aparc-a2009s_rh_thickness_G+S-frontomargin,
2	aparc-a2009s_lh_thickness_G+S-occipital-inf,	**76**	aparc-a2009s_rh_thickness_G+S-occipital-inf,
3	aparc-a2009s_lh_thickness_G+S-paracentral,	**77**	aparc-a2009s_rh_thickness_G+S-paracentral,
4	aparc-a2009s_lh_thickness_G+S-subcentral,	**78**	aparc-a2009s_rh_thickness_G+S-subcentral,
5	aparc-a2009s_lh_thickness_G+S-transv-frontopol,	**79**	aparc-a2009s_rh_thickness_G+S-transv-frontopol,
6	aparc-a2009s_lh_thickness_G+S-cingul-Ant,	**80**	aparc-a2009s_rh_thickness_G+S-cingul-Ant,
7	aparc-a2009s_lh_thickness_G+S-cingul-Mid-Ant,	**81**	aparc-a2009s_rh_thickness_G+S-cingul-Mid-Ant,
8	aparc-a2009s_lh_thickness_G+S-cingul-Mid-Post,	**82**	aparc-a2009s_rh_thickness_G+S-cingul-Mid-Post,
9	aparc-a2009s_lh_thickness_G-cingul-Post-dorsal,	**83**	aparc-a2009s_rh_thickness_G-cingul-Post-dorsal,
10	aparc-a2009s_lh_thickness_G-cingul-Post-ventral,	**84**	aparc-a2009s_rh_thickness_G-cingul-Post-ventral,
11	aparc-a2009s_lh_thickness_G-cuneus,	**85**	aparc-a2009s_rh_thickness_G-cuneus,
12	aparc-a2009s_lh_thickness_G-front-inf-Opercular,	**86**	aparc-a2009s_rh_thickness_G-front-inf-Opercular,
13	aparc-a2009s_lh_thickness_G-front-inf-Orbital,	**87**	aparc-a2009s_rh_thickness_G-front-inf-Orbital,
14	aparc-a2009s_lh_thickness_G-front-inf-Triangul,	**88**	aparc-a2009s_rh_thickness_G-front-inf-Triangul,
15	aparc-a2009s_lh_thickness_G-front-middle,	**89**	aparc-a2009s_rh_thickness_G-front-middle,
16	aparc-a2009s_lh_thickness_G-front-sup,	**90**	aparc-a2009s_rh_thickness_G-front-sup,
17	aparc-a2009s_lh_thickness_G-Ins-lg+S-cent-ins,	**91**	aparc-a2009s_rh_thickness_G-Ins-lg+S-cent-ins,
18	aparc-a2009s_lh_thickness_G-insular-short,	**92**	aparc-a2009s_rh_thickness_G-insular-short,
19	aparc-a2009s_lh_thickness_G-occipital-middle,	**93**	aparc-a2009s_rh_thickness_G-occipital-middle,
20	aparc-a2009s_lh_thickness_G-occipital-sup,	**94**	aparc-a2009s_rh_thickness_G-occipital-sup,
21	aparc-a2009s_lh_thickness_G-oc-temp-lat-fusifor,	**95**	aparc-a2009s_rh_thickness_G-oc-temp-lat-fusifor,
22	aparc-a2009s_lh_thickness_G-oc-temp-med-Lingual,	**96**	aparc-a2009s_rh_thickness_G-oc-temp-med-Lingual,
23	aparc-a2009s_lh_thickness_G-oc-temp-med-Parahip,	**97**	aparc-a2009s_rh_thickness_G-oc-temp-med-Parahip,
24	aparc-a2009s_lh_thickness_G-orbital,	**98**	aparc-a2009s_rh_thickness_G-orbital,
25	aparc-a2009s_lh_thickness_G-pariet-inf-Angular,	**99**	aparc-a2009s_rh_thickness_G-pariet-inf-Angular,
26	aparc-a2009s_lh_thickness_G-pariet-inf-Supramar,	**100**	aparc-a2009s_rh_thickness_G-pariet-inf-Supramar,
27	aparc-a2009s_lh_thickness_G-parietal-sup,	**101**	aparc-a2009s_rh_thickness_G-parietal-sup,
28	aparc-a2009s_lh_thickness_G-postcentral,	**102**	aparc-a2009s_rh_thickness_G-postcentral,
29	aparc-a2009s_lh_thickness_G-precentral,	**103**	aparc-a2009s_rh_thickness_G-precentral,
30	aparc-a2009s_lh_thickness_G-precuneus,	**104**	aparc-a2009s_rh_thickness_G-precuneus,
31	aparc-a2009s_lh_thickness_G-rectus,	**105**	aparc-a2009s_rh_thickness_G-rectus,
32	aparc-a2009s_lh_thickness_G-subcallosal,	**106**	aparc-a2009s_rh_thickness_G-subcallosal,
33	aparc-a2009s_lh_thickness_G-temp-sup-G-T-transv,	**107**	aparc-a2009s_rh_thickness_G-temp-sup-G-T-transv,
34	aparc-a2009s_lh_thickness_G-temp-sup-Lateral,	**108**	aparc-a2009s_rh_thickness_G-temp-sup-Lateral,
35	aparc-a2009s_lh_thickness_G-temp-sup-Plan-polar,	**109**	aparc-a2009s_rh_thickness_G-temp-sup-Plan-polar,
36	aparc-a2009s_lh_thickness_G-temp-sup-Plan-tempo,	**110**	aparc-a2009s_rh_thickness_G-temp-sup-Plan-tempo,
37	aparc-a2009s_lh_thickness_G-temporal-inf,	**111**	aparc-a2009s_rh_thickness_G-temporal-inf,
38	aparc-a2009s_lh_thickness_G-temporal-middle,	**112**	aparc-a2009s_rh_thickness_G-temporal-middle,
39	aparc-a2009s_lh_thickness_Lat-Fis-ant-Horizont,	**113**	aparc-a2009s_rh_thickness_Lat-Fis-ant-Horizont,
40	aparc-a2009s_lh_thickness_Lat-Fis-ant-Vertical,	**114**	aparc-a2009s_rh_thickness_Lat-Fis-ant-Vertical,
41	aparc-a2009s_lh_thickness_Lat-Fis-post,	**115**	aparc-a2009s_rh_thickness_Lat-Fis-post,
42	aparc-a2009s_lh_thickness_Pole-occipital,	**116**	aparc-a2009s_rh_thickness_Pole-occipital,
43	aparc-a2009s_lh_thickness_Pole-temporal,	**117**	aparc-a2009s_rh_thickness_Pole-temporal,
44	aparc-a2009s_lh_thickness_S-calcarine,	**118**	aparc-a2009s_rh_thickness_S-calcarine,
45	aparc-a2009s_lh_thickness_S-central,	**119**	aparc-a2009s_rh_thickness_S-central,
46	aparc-a2009s_lh_thickness_S-cingul-Marginalis,	**120**	aparc-a2009s_rh_thickness_S-cingul-Marginalis,
47	aparc-a2009s_lh_thickness_S-circular-insula-ant,	**121**	aparc-a2009s_rh_thickness_S-circular-insula-ant,
48	aparc-a2009s_lh_thickness_S-circular-insula-inf,	**122**	aparc-a2009s_rh_thickness_S-circular-insula-inf,
49	aparc-a2009s_lh_thickness_S-circular-insula-sup,	**123**	aparc-a2009s_rh_thickness_S-circular-insula-sup,
50	aparc-a2009s_lh_thickness_S-collat-transv-ant,	**124**	aparc-a2009s_rh_thickness_S-collat-transv-ant,
51	aparc-a2009s_lh_thickness_S-collat-transv-post,	**125**	aparc-a2009s_rh_thickness_S-collat-transv-post,
52	aparc-a2009s_lh_thickness_S-front-inf,	**126**	aparc-a2009s_rh_thickness_S-front-inf,
53	aparc-a2009s_lh_thickness_S-front-middle,	**127**	aparc-a2009s_rh_thickness_S-front-middle,
54	aparc-a2009s_lh_thickness_S-front-sup,	**128**	aparc-a2009s_rh_thickness_S-front-sup,
55	aparc-a2009s_lh_thickness_S-interm-prim-Jensen,	**129**	aparc-a2009s_rh_thickness_S-interm-prim-Jensen,
56	aparc-a2009s_lh_thickness_S-intrapariet+P-trans,	**130**	aparc-a2009s_rh_thickness_S-intrapariet+P-trans,
57	aparc-a2009s_lh_thickness_S-oc-middle+Lunatus,	**131**	aparc-a2009s_rh_thickness_S-oc-middle+Lunatus,
58	aparc-a2009s_lh_thickness_S-oc-sup+transversal,	**132**	aparc-a2009s_rh_thickness_S-oc-sup+transversal,
59	aparc-a2009s_lh_thickness_S-occipital-ant,	**133**	aparc-a2009s_rh_thickness_S-occipital-ant,
60	aparc-a2009s_lh_thickness_S-oc-temp-lat,	**134**	aparc-a2009s_rh_thickness_S-oc-temp-lat,
61	aparc-a2009s_lh_thickness_S-oc-temp-med+Lingual,	**135**	aparc-a2009s_rh_thickness_S-oc-temp-med+Lingual,
62	aparc-a2009s_lh_thickness_S-orbital-lateral,	**136**	aparc-a2009s_rh_thickness_S-orbital-lateral,
63	aparc-a2009s_lh_thickness_S-orbital-med-olfact,	**137**	aparc-a2009s_rh_thickness_S-orbital-med-olfact,
64	aparc-a2009s_lh_thickness_S-orbital-H-Shaped,	**138**	aparc-a2009s_rh_thickness_S-orbital-H-Shaped,
65	aparc-a2009s_lh_thickness_S-parieto-occipital,	**139**	aparc-a2009s_rh_thickness_S-parieto-occipital,
66	aparc-a2009s_lh_thickness_S-pericallosal,	**140**	aparc-a2009s_rh_thickness_S-pericallosal,
67	aparc-a2009s_lh_thickness_S-postcentral,	**141**	aparc-a2009s_rh_thickness_S-postcentral,
68	aparc-a2009s_lh_thickness_S-precentral-inf-part,	**142**	aparc-a2009s_rh_thickness_S-precentral-inf-part,
69	aparc-a2009s_lh_thickness_S-precentral-sup-part,	**143**	aparc-a2009s_rh_thickness_S-precentral-sup-part,
70	aparc-a2009s_lh_thickness_S-suborbital,	**144**	aparc-a2009s_rh_thickness_S-suborbital,
71	aparc-a2009s_lh_thickness_S-subparietal,	**145**	aparc-a2009s_rh_thickness_S-subparietal,
72	aparc-a2009s_lh_thickness_S-temporal-inf,	**146**	aparc-a2009s_rh_thickness_S-temporal-inf,
73	aparc-a2009s_lh_thickness_S-temporal-sup,	**147**	aparc-a2009s_rh_thickness_S-temporal-sup,
74	aparc-a2009s_lh_thickness_S-temporal-transverse,	**148**	aparc-a2009s_rh_thickness_S-temporal-transverse,

Table S 9 Names of UK Biobank Imaging Derived Phenotypes (IDPs) related to White Matter: 453. Supplement to **Figure 6**.

IDP #	IDP Name	IDP #	IDP Name
1	IDP_T2_FLAIR_BIANCA_WMH_volume	**78**	IDP_dMRI_ProbtrackX_FA_unc_r
2	IDP_T2_FLAIR_BIANCA_periventWMH_volume	**79**	IDP_dMRI_TBSS_MO_Middle_cerebellar_peduncle
3	IDP_T2_FLAIR_BIANCA_deepWMH_volume	**80**	IDP_dMRI_TBSS_MO_Pontine_crossing_tract
4	IDP_dMRI_TBSS_FA_Middle_cerebellar_peduncle	**81**	IDP_dMRI_TBSS_MO_Genu_of_corpus_callosum
5	IDP_dMRI_TBSS_FA_Pontine_crossing_tract	**82**	IDP_dMRI_TBSS_MO_Body_of_corpus_callosum
6	IDP_dMRI_TBSS_FA_Genu_of_corpus_callosum	**83**	IDP_dMRI_TBSS_MO_Splenium_of_corpus_callosum
7	IDP_dMRI_TBSS_FA_Body_of_corpus_callosum	**84**	IDP_dMRI_TBSS_MO_Fornix
8	IDP_dMRI_TBSS_FA_Splenium_of_corpus_callosum	**85**	IDP_dMRI_TBSS_MO_Corticospinal_tract_R
9	IDP_dMRI_TBSS_FA_Fornix	**86**	IDP_dMRI_TBSS_MO_Corticospinal_tract_L
10	IDP_dMRI_TBSS_FA_Corticospinal_tract_R	**87**	IDP_dMRI_TBSS_MO_Medial_lemniscus_R
11	IDP_dMRI_TBSS_FA_Corticospinal_tract_L	**88**	IDP_dMRI_TBSS_MO_Medial_lemniscus_L
12	IDP_dMRI_TBSS_FA_Medial_lemniscus_R	**89**	IDP_dMRI_TBSS_MO_Inferior_cerebellar_peduncle_R
13	IDP_dMRI_TBSS_FA_Medial_lemniscus_L	**90**	IDP_dMRI_TBSS_MO_Inferior_cerebellar_peduncle_L
14	IDP_dMRI_TBSS_FA_Inferior_cerebellar_peduncle_R	**91**	IDP_dMRI_TBSS_MO_Superior_cerebellar_peduncle_R
15	IDP_dMRI_TBSS_FA_Inferior_cerebellar_peduncle_L	**92**	IDP_dMRI_TBSS_MO_Superior_cerebellar_peduncle_L
16	IDP_dMRI_TBSS_FA_Superior_cerebellar_peduncle_R	**93**	IDP_dMRI_TBSS_MO_Cerebral_peduncle_R
17	IDP_dMRI_TBSS_FA_Superior_cerebellar_peduncle_L	**94**	IDP_dMRI_TBSS_MO_Cerebral_peduncle_L
18	IDP_dMRI_TBSS_FA_Cerebral_peduncle_R	**95**	IDP_dMRI_TBSS_MO_Anterior_limb_of_internal_capsule_R
19	IDP_dMRI_TBSS_FA_Cerebral_peduncle_L	**96**	IDP_dMRI_TBSS_MO_Anterior_limb_of_internal_capsule_L
20	IDP_dMRI_TBSS_FA_Anterior_limb_of_internal_capsule_R	**97**	IDP_dMRI_TBSS_MO_Posterior_limb_of_internal_capsule_R
21	IDP_dMRI_TBSS_FA_Anterior_limb_of_internal_capsule_L	**98**	IDP_dMRI_TBSS_MO_Posterior_limb_of_internal_capsule_L
22	IDP_dMRI_TBSS_FA_Posterior_limb_of_internal_capsule_R	**99**	IDP_dMRI_TBSS_MO_Retrolenticular_part_of_internal_capsule_R
23	IDP_dMRI_TBSS_FA_Posterior_limb_of_internal_capsule_L	**100**	IDP_dMRI_TBSS_MO_Retrolenticular_part_of_internal_capsule_L
24	IDP_dMRI_TBSS_FA_Retrolenticular_part_of_internal_capsule_R	**101**	IDP_dMRI_TBSS_MO_Anterior_corona_radiata_R
25	IDP_dMRI_TBSS_FA_Retrolenticular_part_of_internal_capsule_L	**102**	IDP_dMRI_TBSS_MO_Anterior_corona_radiata_L
26	IDP_dMRI_TBSS_FA_Anterior_corona_radiata_R	**103**	IDP_dMRI_TBSS_MO_Superior_corona_radiata_R
27	IDP_dMRI_TBSS_FA_Anterior_corona_radiata_L	**104**	IDP_dMRI_TBSS_MO_Superior_corona_radiata_L
28	IDP_dMRI_TBSS_FA_Superior_corona_radiata_R	**105**	IDP_dMRI_TBSS_MO_Posterior_corona_radiata_R
29	IDP_dMRI_TBSS_FA_Superior_corona_radiata_L	**106**	IDP_dMRI_TBSS_MO_Posterior_corona_radiata_L
30	IDP_dMRI_TBSS_FA_Posterior_corona_radiata_R	**107**	IDP_dMRI_TBSS_MO_Posterior_thalamic_radiation_R
31	IDP_dMRI_TBSS_FA_Posterior_corona_radiata_L	**108**	IDP_dMRI_TBSS_MO_Posterior_thalamic_radiation_L
32	IDP_dMRI_TBSS_FA_Posterior_thalamic_radiation_R	**109**	IDP_dMRI_TBSS_MO_Sagittal_stratum_R
33	IDP_dMRI_TBSS_FA_Posterior_thalamic_radiation_L	**110**	IDP_dMRI_TBSS_MO_Sagittal_stratum_L
34	IDP_dMRI_TBSS_FA_Sagittal_stratum_R	**111**	IDP_dMRI_TBSS_MO_External_capsule_R
35	IDP_dMRI_TBSS_FA_Sagittal_stratum_L	**112**	IDP_dMRI_TBSS_MO_External_capsule_L
36	IDP_dMRI_TBSS_FA_External_capsule_R	**113**	IDP_dMRI_TBSS_MO_Cingulum_cingulate_gyrus_R
37	IDP_dMRI_TBSS_FA_External_capsule_L	**114**	IDP_dMRI_TBSS_MO_Cingulum_cingulate_gyrus_L
38	IDP_dMRI_TBSS_FA_Cingulum_cingulate_gyrus_R	**115**	IDP_dMRI_TBSS_MO_Cingulum_hippocampus_R
39	IDP_dMRI_TBSS_FA_Cingulum_cingulate_gyrus_L	**116**	IDP_dMRI_TBSS_MO_Cingulum_hippocampus_L
40	IDP_dMRI_TBSS_FA_Cingulum_hippocampus_R	**117**	IDP_dMRI_TBSS_MO_Fornix_cres+Stria_terminalis_R
41	IDP_dMRI_TBSS_FA_Cingulum_hippocampus_L	**118**	IDP_dMRI_TBSS_MO_Fornix_cres+Stria_terminalis_L
42	IDP_dMRI_TBSS_FA_Fornix_cres+Stria_terminalis_R	**119**	IDP_dMRI_TBSS_MO_Superior_longitudinal_fasciculus_R
43	IDP_dMRI_TBSS_FA_Fornix_cres+Stria_terminalis_L	**120**	IDP_dMRI_TBSS_MO_Superior_longitudinal_fasciculus_L
44	IDP_dMRI_TBSS_FA_Superior_longitudinal_fasciculus_R	**121**	IDP_dMRI_TBSS_MO_Superior_fronto-occipital_fasciculus_R
45	IDP_dMRI_TBSS_FA_Superior_longitudinal_fasciculus_L	**122**	IDP_dMRI_TBSS_MO_Superior_fronto-occipital_fasciculus_L
46	IDP_dMRI_TBSS_FA_Superior_fronto-occipital_fasciculus_R	**123**	IDP_dMRI_TBSS_MO_Uncinate_fasciculus_R
47	IDP_dMRI_TBSS_FA_Superior_fronto-occipital_fasciculus_L	**124**	IDP_dMRI_TBSS_MO_Uncinate_fasciculus_L
48	IDP_dMRI_TBSS_FA_Uncinate_fasciculus_R	**125**	IDP_dMRI_TBSS_MO_Tapetum_R
49	IDP_dMRI_TBSS_FA_Uncinate_fasciculus_L	**126**	IDP_dMRI_TBSS_MO_Tapetum_L
50	IDP_dMRI_TBSS_FA_Tapetum_R	**127**	IDP_dMRI_ProbtrackX_MO_ar_l
51	IDP_dMRI_TBSS_FA_Tapetum_L	**128**	IDP_dMRI_ProbtrackX_MO_ar_r
52	IDP_dMRI_ProbtrackX_FA_ar_l	**129**	IDP_dMRI_ProbtrackX_MO_atr_l
53	IDP_dMRI_ProbtrackX_FA_ar_r	**130**	IDP_dMRI_ProbtrackX_MO_atr_r
54	IDP_dMRI_ProbtrackX_FA_atr_l	**131**	IDP_dMRI_ProbtrackX_MO_cgc_l
55	IDP_dMRI_ProbtrackX_FA_atr_r	**132**	IDP_dMRI_ProbtrackX_MO_cgc_r
56	IDP_dMRI_ProbtrackX_FA_cgc_l	**133**	IDP_dMRI_ProbtrackX_MO_cgh_l
57	IDP_dMRI_ProbtrackX_FA_cgc_r	**134**	IDP_dMRI_ProbtrackX_MO_cgh_r
58	IDP_dMRI_ProbtrackX_FA_cgh_l	**135**	IDP_dMRI_ProbtrackX_MO_cst_l
59	IDP_dMRI_ProbtrackX_FA_cgh_r	**136**	IDP_dMRI_ProbtrackX_MO_cst_r
60	IDP_dMRI_ProbtrackX_FA_cst_l	**137**	IDP_dMRI_ProbtrackX_MO_fma
61	IDP_dMRI_ProbtrackX_FA_cst_r	**138**	IDP_dMRI_ProbtrackX_MO_fmi
62	IDP_dMRI_ProbtrackX_FA_fma	**139**	IDP_dMRI_ProbtrackX_MO_ifo_l
63	IDP_dMRI_ProbtrackX_FA_fmi	**140**	IDP_dMRI_ProbtrackX_MO_ifo_r
64	IDP_dMRI_ProbtrackX_FA_ifo_l	**141**	IDP_dMRI_ProbtrackX_MO_ilf_l
65	IDP_dMRI_ProbtrackX_FA_ifo_r	**142**	IDP_dMRI_ProbtrackX_MO_ilf_r
66	IDP_dMRI_ProbtrackX_FA_ilf_l	**143**	IDP_dMRI_ProbtrackX_MO_mcp
67	IDP_dMRI_ProbtrackX_FA_ilf_r	**144**	IDP_dMRI_ProbtrackX_MO_ml_l
68	IDP_dMRI_ProbtrackX_FA_mcp	**145**	IDP_dMRI_ProbtrackX_MO_ml_r
69	IDP_dMRI_ProbtrackX_FA_ml_l	**146**	IDP_dMRI_ProbtrackX_MO_ptr_l
70	IDP_dMRI_ProbtrackX_FA_ml_r	**147**	IDP_dMRI_ProbtrackX_MO_ptr_r
71	IDP_dMRI_ProbtrackX_FA_ptr_l	**148**	IDP_dMRI_ProbtrackX_MO_slf_l
72	IDP_dMRI_ProbtrackX_FA_ptr_r	**149**	IDP_dMRI_ProbtrackX_MO_slf_r
73	IDP_dMRI_ProbtrackX_FA_slf_l	**150**	IDP_dMRI_ProbtrackX_MO_str_l
74	IDP_dMRI_ProbtrackX_FA_slf_r	**151**	IDP_dMRI_ProbtrackX_MO_str_r
75	IDP_dMRI_ProbtrackX_FA_str_l	**152**	IDP_dMRI_ProbtrackX_MO_unc_l
76	IDP_dMRI_ProbtrackX_FA_str_r	**153**	IDP_dMRI_ProbtrackX_MO_unc_r
77	IDP_dMRI_ProbtrackX_FA_unc_l	**154**	IDP_dMRI_TBSS_MD_Middle_cerebellar_peduncle

Table S 10 Names of UK Biobank Imaging Derived Phenotypes (IDPs) related to White Matter-continued: 453. Supplement to **Figure 6**.

IDP #	IDP Name	IDP #	IDP Name
155	IDP_dMRI_TBSS_MD_Pontine_crossing_tract	**232**	IDP_dMRI_TBSS_ICVF_Body_of_corpus_callosum
156	IDP_dMRI_TBSS_MD_Genu_of_corpus_callosum	**233**	IDP_dMRI_TBSS_ICVF_Splenium_of_corpus_callosum
157	IDP_dMRI_TBSS_MD_Body_of_corpus_callosum	**234**	IDP_dMRI_TBSS_ICVF_Fornix
158	IDP_dMRI_TBSS_MD_Splenium_of_corpus_callosum	**235**	IDP_dMRI_TBSS_ICVF_Corticospinal_tract_R
159	IDP_dMRI_TBSS_MD_Fornix	**236**	IDP_dMRI_TBSS_ICVF_Corticospinal_tract_L
160	IDP_dMRI_TBSS_MD_Corticospinal_tract_R	**237**	IDP_dMRI_TBSS_ICVF_Medial_lemniscus_R
161	IDP_dMRI_TBSS_MD_Corticospinal_tract_L	**238**	IDP_dMRI_TBSS_ICVF_Medial_lemniscus_L
162	IDP_dMRI_TBSS_MD_Medial_lemniscus_R	**239**	IDP_dMRI_TBSS_ICVF_Inferior_cerebellar_peduncle_R
163	IDP_dMRI_TBSS_MD_Medial_lemniscus_L	**240**	IDP_dMRI_TBSS_ICVF_Inferior_cerebellar_peduncle_L
164	IDP_dMRI_TBSS_MD_Inferior_cerebellar_peduncle_R	**241**	IDP_dMRI_TBSS_ICVF_Superior_cerebellar_peduncle_R
165	IDP_dMRI_TBSS_MD_Inferior_cerebellar_peduncle_L	**242**	IDP_dMRI_TBSS_ICVF_Superior_cerebellar_peduncle_L
166	IDP_dMRI_TBSS_MD_Superior_cerebellar_peduncle_R	**243**	IDP_dMRI_TBSS_ICVF_Cerebral_peduncle_R
167	IDP_dMRI_TBSS_MD_Superior_cerebellar_peduncle_L	**244**	IDP_dMRI_TBSS_ICVF_Cerebral_peduncle_L
168	IDP_dMRI_TBSS_MD_Cerebral_peduncle_R	**245**	IDP_dMRI_TBSS_ICVF_Anterior_limb_of_internal_capsule_R
169	IDP_dMRI_TBSS_MD_Cerebral_peduncle_L	**246**	IDP_dMRI_TBSS_ICVF_Anterior_limb_of_internal_capsule_L
170	IDP_dMRI_TBSS_MD_Anterior_limb_of_internal_capsule_R	**247**	IDP_dMRI_TBSS_ICVF_Posterior_limb_of_internal_capsule_R
171	IDP_dMRI_TBSS_MD_Anterior_limb_of_internal_capsule_L	**248**	IDP_dMRI_TBSS_ICVF_Posterior_limb_of_internal_capsule_L
172	IDP_dMRI_TBSS_MD_Posterior_limb_of_internal_capsule_R	**249**	IDP_dMRI_TBSS_ICVF_Retrolenticular_part_of_internal_capsule_R
173	IDP_dMRI_TBSS_MD_Posterior_limb_of_internal_capsule_L	**250**	IDP_dMRI_TBSS_ICVF_Retrolenticular_part_of_internal_capsule_L
174	IDP_dMRI_TBSS_MD_Retrolenticular_part_of_internal_capsule_R	**251**	IDP_dMRI_TBSS_ICVF_Anterior_corona_radiata_R
175	IDP_dMRI_TBSS_MD_Retrolenticular_part_of_internal_capsule_L	**252**	IDP_dMRI_TBSS_ICVF_Anterior_corona_radiata_L
176	IDP_dMRI_TBSS_MD_Anterior_corona_radiata_R	**253**	IDP_dMRI_TBSS_ICVF_Superior_corona_radiata_R
177	IDP_dMRI_TBSS_MD_Anterior_corona_radiata_L	**254**	IDP_dMRI_TBSS_ICVF_Superior_corona_radiata_L
178	IDP_dMRI_TBSS_MD_Superior_corona_radiata_R	**255**	IDP_dMRI_TBSS_ICVF_Posterior_corona_radiata_R
179	IDP_dMRI_TBSS_MD_Superior_corona_radiata_L	**256**	IDP_dMRI_TBSS_ICVF_Posterior_corona_radiata_L
180	IDP_dMRI_TBSS_MD_Posterior_corona_radiata_R	**257**	IDP_dMRI_TBSS_ICVF_Posterior_thalamic_radiation_R
181	IDP_dMRI_TBSS_MD_Posterior_corona_radiata_L	**258**	IDP_dMRI_TBSS_ICVF_Posterior_thalamic_radiation_L
182	IDP_dMRI_TBSS_MD_Posterior_thalamic_radiation_R	**259**	IDP_dMRI_TBSS_ICVF_Sagittal_stratum_R
183	IDP_dMRI_TBSS_MD_Posterior_thalamic_radiation_L	**260**	IDP_dMRI_TBSS_ICVF_Sagittal_stratum_L
184	IDP_dMRI_TBSS_MD_Sagittal_stratum_R	**261**	IDP_dMRI_TBSS_ICVF_External_capsule_R
185	IDP_dMRI_TBSS_MD_Sagittal_stratum_L	**262**	IDP_dMRI_TBSS_ICVF_External_capsule_L
186	IDP_dMRI_TBSS_MD_External_capsule_R	**263**	IDP_dMRI_TBSS_ICVF_Cingulum_cingulate_gyrus_R
187	IDP_dMRI_TBSS_MD_External_capsule_L	**264**	IDP_dMRI_TBSS_ICVF_Cingulum_cingulate_gyrus_L
188	IDP_dMRI_TBSS_MD_Cingulum_cingulate_gyrus_R	**265**	IDP_dMRI_TBSS_ICVF_Cingulum_hippocampus_R
189	IDP_dMRI_TBSS_MD_Cingulum_cingulate_gyrus_L	**266**	IDP_dMRI_TBSS_ICVF_Cingulum_hippocampus_L
190	IDP_dMRI_TBSS_MD_Cingulum_hippocampus_R	**267**	IDP_dMRI_TBSS_ICVF_Fornix_cres+Stria_terminalis_R
191	IDP_dMRI_TBSS_MD_Cingulum_hippocampus_L	**268**	IDP_dMRI_TBSS_ICVF_Fornix_cres+Stria_terminalis_L
192	IDP_dMRI_TBSS_MD_Fornix_cres+Stria_terminalis_R	**269**	IDP_dMRI_TBSS_ICVF_Superior_longitudinal_fasciculus_R
193	IDP_dMRI_TBSS_MD_Fornix_cres+Stria_terminalis_L	**270**	IDP_dMRI_TBSS_ICVF_Superior_longitudinal_fasciculus_L
194	IDP_dMRI_TBSS_MD_Superior_longitudinal_fasciculus_R	**271**	IDP_dMRI_TBSS_ICVF_Superior_fronto-occipital_fasciculus_R
195	IDP_dMRI_TBSS_MD_Superior_longitudinal_fasciculus_L	**272**	IDP_dMRI_TBSS_ICVF_Superior_fronto-occipital_fasciculus_L
196	IDP_dMRI_TBSS_MD_Superior_fronto-occipital_fasciculus_R	**273**	IDP_dMRI_TBSS_ICVF_Uncinate_fasciculus_R
197	IDP_dMRI_TBSS_MD_Superior_fronto-occipital_fasciculus_L	**274**	IDP_dMRI_TBSS_ICVF_Uncinate_fasciculus_L
198	IDP_dMRI_TBSS_MD_Uncinate_fasciculus_R	**275**	IDP_dMRI_TBSS_ICVF_Tapetum_R
199	IDP_dMRI_TBSS_MD_Uncinate_fasciculus_L	**276**	IDP_dMRI_TBSS_ICVF_Tapetum_L
200	IDP_dMRI_TBSS_MD_Tapetum_R	**277**	IDP_dMRI_ProbtrackX_ICVF_ar_l
201	IDP_dMRI_TBSS_MD_Tapetum_L	**278**	IDP_dMRI_ProbtrackX_ICVF_ar_r
202	IDP_dMRI_ProbtrackX_MD_ar_l	**279**	IDP_dMRI_ProbtrackX_ICVF_atr_l
203	IDP_dMRI_ProbtrackX_MD_ar_r	**280**	IDP_dMRI_ProbtrackX_ICVF_atr_r
204	IDP_dMRI_ProbtrackX_MD_atr_l	**281**	IDP_dMRI_ProbtrackX_ICVF_cgc_l
205	IDP_dMRI_ProbtrackX_MD_atr_r	**282**	IDP_dMRI_ProbtrackX_ICVF_cgc_r
206	IDP_dMRI_ProbtrackX_MD_cgc_l	**283**	IDP_dMRI_ProbtrackX_ICVF_cgh_l
207	IDP_dMRI_ProbtrackX_MD_cgc_r	**284**	IDP_dMRI_ProbtrackX_ICVF_cgh_r
208	IDP_dMRI_ProbtrackX_MD_cgh_l	**285**	IDP_dMRI_ProbtrackX_ICVF_cst_l
209	IDP_dMRI_ProbtrackX_MD_cgh_r	**286**	IDP_dMRI_ProbtrackX_ICVF_cst_r
210	IDP_dMRI_ProbtrackX_MD_cst_l	**287**	IDP_dMRI_ProbtrackX_ICVF_fma
211	IDP_dMRI_ProbtrackX_MD_cst_r	**288**	IDP_dMRI_ProbtrackX_ICVF_fmi
212	IDP_dMRI_ProbtrackX_MD_fma	**289**	IDP_dMRI_ProbtrackX_ICVF_ifo_l
213	IDP_dMRI_ProbtrackX_MD_fmi	**290**	IDP_dMRI_ProbtrackX_ICVF_ifo_r
214	IDP_dMRI_ProbtrackX_MD_ifo_l	**291**	IDP_dMRI_ProbtrackX_ICVF_ilf_l
215	IDP_dMRI_ProbtrackX_MD_ifo_r	**292**	IDP_dMRI_ProbtrackX_ICVF_ilf_r
216	IDP_dMRI_ProbtrackX_MD_ilf_l	**293**	IDP_dMRI_ProbtrackX_ICVF_mcp
217	IDP_dMRI_ProbtrackX_MD_ilf_r	**294**	IDP_dMRI_ProbtrackX_ICVF_ml_l
218	IDP_dMRI_ProbtrackX_MD_mcp	**295**	IDP_dMRI_ProbtrackX_ICVF_ml_r
219	IDP_dMRI_ProbtrackX_MD_ml_l	**296**	IDP_dMRI_ProbtrackX_ICVF_ptr_l
220	IDP_dMRI_ProbtrackX_MD_ml_r	**297**	IDP_dMRI_ProbtrackX_ICVF_ptr_r
221	IDP_dMRI_ProbtrackX_MD_ptr_l	**298**	IDP_dMRI_ProbtrackX_ICVF_slf_l
222	IDP_dMRI_ProbtrackX_MD_ptr_r	**299**	IDP_dMRI_ProbtrackX_ICVF_slf_r
223	IDP_dMRI_ProbtrackX_MD_slf_l	**300**	IDP_dMRI_ProbtrackX_ICVF_str_l
224	IDP_dMRI_ProbtrackX_MD_slf_r	**301**	IDP_dMRI_ProbtrackX_ICVF_str_r
225	IDP_dMRI_ProbtrackX_MD_str_l	**302**	IDP_dMRI_ProbtrackX_ICVF_unc_l
226	IDP_dMRI_ProbtrackX_MD_str_r	**303**	IDP_dMRI_ProbtrackX_ICVF_unc_r
227	IDP_dMRI_ProbtrackX_MD_unc_l	**304**	IDP_dMRI_TBSS_OD_Middle_cerebellar_peduncle
228	IDP_dMRI_ProbtrackX_MD_unc_r	**305**	IDP_dMRI_TBSS_OD_Pontine_crossing_tract
229	IDP_dMRI_TBSS_ICVF_Middle_cerebellar_peduncle	**306**	IDP_dMRI_TBSS_OD_Genu_of_corpus_callosum
230	IDP_dMRI_TBSS_ICVF_Pontine_crossing_tract	**307**	IDP_dMRI_TBSS_OD_Body_of_corpus_callosum
231	IDP_dMRI_TBSS_ICVF_Genu_of_corpus_callosum	**308**	IDP_dMRI_TBSS_OD_Splenium_of_corpus_callosum

Table S 11 Names of UK Biobank Imaging Derived Phenotypes (IDPs) related to White Matter-continued: 453. Supplement to **Figure 6**.

IDP #	IDP Name	IDP #	IDP Name
309	IDP_dMRI_TBSS_OD_Fornix	**386**	IDP_dMRI_TBSS_ISOVF_Corticospinal_tract_L
310	IDP_dMRI_TBSS_OD_Corticospinal_tract_R	**387**	IDP_dMRI_TBSS_ISOVF_Medial_lemniscus_R
311	IDP_dMRI_TBSS_OD_Corticospinal_tract_L	**388**	IDP_dMRI_TBSS_ISOVF_Medial_lemniscus_L
312	IDP_dMRI_TBSS_OD_Medial_lemniscus_R	**389**	IDP_dMRI_TBSS_ISOVF_Inferior_cerebellar_peduncle_R
313	IDP_dMRI_TBSS_OD_Medial_lemniscus_L	**390**	IDP_dMRI_TBSS_ISOVF_Inferior_cerebellar_peduncle_L
314	IDP_dMRI_TBSS_OD_Inferior_cerebellar_peduncle_R	**391**	IDP_dMRI_TBSS_ISOVF_Superior_cerebellar_peduncle_R
315	IDP_dMRI_TBSS_OD_Inferior_cerebellar_peduncle_L	**392**	IDP_dMRI_TBSS_ISOVF_Superior_cerebellar_peduncle_L
316	IDP_dMRI_TBSS_OD_Superior_cerebellar_peduncle_R	**393**	IDP_dMRI_TBSS_ISOVF_Cerebral_peduncle_R
317	IDP_dMRI_TBSS_OD_Superior_cerebellar_peduncle_L	**394**	IDP_dMRI_TBSS_ISOVF_Cerebral_peduncle_L
318	IDP_dMRI_TBSS_OD_Cerebral_peduncle_R	**395**	IDP_dMRI_TBSS_ISOVF_Anterior_limb_of_internal_capsule_R
319	IDP_dMRI_TBSS_OD_Cerebral_peduncle_L	**396**	IDP_dMRI_TBSS_ISOVF_Anterior_limb_of_internal_capsule_L
320	IDP_dMRI_TBSS_OD_Anterior_limb_of_internal_capsule_R	**397**	IDP_dMRI_TBSS_ISOVF_Posterior_limb_of_internal_capsule_R
321	IDP_dMRI_TBSS_OD_Anterior_limb_of_internal_capsule_L	**398**	IDP_dMRI_TBSS_ISOVF_Posterior_limb_of_internal_capsule_L
322	IDP_dMRI_TBSS_OD_Posterior_limb_of_internal_capsule_R	**399**	IDP_dMRI_TBSS_ISOVF_Retrolenticular_part_of_internal_capsule_R
323	IDP_dMRI_TBSS_OD_Posterior_limb_of_internal_capsule_L	**400**	IDP_dMRI_TBSS_ISOVF_Retrolenticular_part_of_internal_capsule_L
324	IDP_dMRI_TBSS_OD_Retrolenticular_part_of_internal_capsule_R	**401**	IDP_dMRI_TBSS_ISOVF_Anterior_corona_radiata_R
325	IDP_dMRI_TBSS_OD_Retrolenticular_part_of_internal_capsule_L	**402**	IDP_dMRI_TBSS_ISOVF_Anterior_corona_radiata_L
326	IDP_dMRI_TBSS_OD_Anterior_corona_radiata_R	**403**	IDP_dMRI_TBSS_ISOVF_Superior_corona_radiata_R
327	IDP_dMRI_TBSS_OD_Anterior_corona_radiata_L	**404**	IDP_dMRI_TBSS_ISOVF_Superior_corona_radiata_L
328	IDP_dMRI_TBSS_OD_Superior_corona_radiata_R	**405**	IDP_dMRI_TBSS_ISOVF_Posterior_corona_radiata_R
329	IDP_dMRI_TBSS_OD_Superior_corona_radiata_L	**406**	IDP_dMRI_TBSS_ISOVF_Posterior_corona_radiata_L
330	IDP_dMRI_TBSS_OD_Posterior_corona_radiata_R	**407**	IDP_dMRI_TBSS_ISOVF_Posterior_thalamic_radiation_R
331	IDP_dMRI_TBSS_OD_Posterior_corona_radiata_L	**408**	IDP_dMRI_TBSS_ISOVF_Posterior_thalamic_radiation_L
332	IDP_dMRI_TBSS_OD_Posterior_thalamic_radiation_R	**409**	IDP_dMRI_TBSS_ISOVF_Sagittal_stratum_R
333	IDP_dMRI_TBSS_OD_Posterior_thalamic_radiation_L	**410**	IDP_dMRI_TBSS_ISOVF_Sagittal_stratum_L
334	IDP_dMRI_TBSS_OD_Sagittal_stratum_R	**411**	IDP_dMRI_TBSS_ISOVF_External_capsule_R
335	IDP_dMRI_TBSS_OD_Sagittal_stratum_L	**412**	IDP_dMRI_TBSS_ISOVF_External_capsule_L
336	IDP_dMRI_TBSS_OD_External_capsule_R	**413**	IDP_dMRI_TBSS_ISOVF_Cingulum_cingulate_gyrus_R
337	IDP_dMRI_TBSS_OD_External_capsule_L	**414**	IDP_dMRI_TBSS_ISOVF_Cingulum_cingulate_gyrus_L
338	IDP_dMRI_TBSS_OD_Cingulum_cingulate_gyrus_R	**415**	IDP_dMRI_TBSS_ISOVF_Cingulum_hippocampus_R
339	IDP_dMRI_TBSS_OD_Cingulum_cingulate_gyrus_L	**416**	IDP_dMRI_TBSS_ISOVF_Cingulum_hippocampus_L
340	IDP_dMRI_TBSS_OD_Cingulum_hippocampus_R	**417**	IDP_dMRI_TBSS_ISOVF_Fornix_cres+Stria_terminalis_R
341	IDP_dMRI_TBSS_OD_Cingulum_hippocampus_L	**418**	IDP_dMRI_TBSS_ISOVF_Fornix_cres+Stria_terminalis_L
342	IDP_dMRI_TBSS_OD_Fornix_cres+Stria_terminalis_R	**419**	IDP_dMRI_TBSS_ISOVF_Superior_longitudinal_fasciculus_R
343	IDP_dMRI_TBSS_OD_Fornix_cres+Stria_terminalis_L	**420**	IDP_dMRI_TBSS_ISOVF_Superior_longitudinal_fasciculus_L
344	IDP_dMRI_TBSS_OD_Superior_longitudinal_fasciculus_R	**421**	IDP_dMRI_TBSS_ISOVF_Superior_fronto-occipital_fasciculus_R
345	IDP_dMRI_TBSS_OD_Superior_longitudinal_fasciculus_L	**422**	IDP_dMRI_TBSS_ISOVF_Superior_fronto-occipital_fasciculus_L
346	IDP_dMRI_TBSS_OD_Superior_fronto-occipital_fasciculus_R	**423**	IDP_dMRI_TBSS_ISOVF_Uncinate_fasciculus_R
347	IDP_dMRI_TBSS_OD_Superior_fronto-occipital_fasciculus_L	**424**	IDP_dMRI_TBSS_ISOVF_Uncinate_fasciculus_L
348	IDP_dMRI_TBSS_OD_Uncinate_fasciculus_R	**425**	IDP_dMRI_TBSS_ISOVF_Tapetum_R
349	IDP_dMRI_TBSS_OD_Uncinate_fasciculus_L	**426**	IDP_dMRI_TBSS_ISOVF_Tapetum_L
350	IDP_dMRI_TBSS_OD_Tapetum_R	**427**	IDP_dMRI_ProbtrackX_ISOVF_ar_l
351	IDP_dMRI_TBSS_OD_Tapetum_L	**428**	IDP_dMRI_ProbtrackX_ISOVF_ar_r
352	IDP_dMRI_ProbtrackX_OD_ar_l	**429**	IDP_dMRI_ProbtrackX_ISOVF_atr_l
353	IDP_dMRI_ProbtrackX_OD_ar_r	**430**	IDP_dMRI_ProbtrackX_ISOVF_atr_r
354	IDP_dMRI_ProbtrackX_OD_atr_l	**431**	IDP_dMRI_ProbtrackX_ISOVF_cgc_l
355	IDP_dMRI_ProbtrackX_OD_atr_r	**432**	IDP_dMRI_ProbtrackX_ISOVF_cgc_r
356	IDP_dMRI_ProbtrackX_OD_cgc_l	**433**	IDP_dMRI_ProbtrackX_ISOVF_cgh_l
357	IDP_dMRI_ProbtrackX_OD_cgc_r	**434**	IDP_dMRI_ProbtrackX_ISOVF_cgh_r
358	IDP_dMRI_ProbtrackX_OD_cgh_l	**435**	IDP_dMRI_ProbtrackX_ISOVF_cst_l
359	IDP_dMRI_ProbtrackX_OD_cgh_r	**436**	IDP_dMRI_ProbtrackX_ISOVF_cst_r
360	IDP_dMRI_ProbtrackX_OD_cst_l	**437**	IDP_dMRI_ProbtrackX_ISOVF_fma
361	IDP_dMRI_ProbtrackX_OD_cst_r	**438**	IDP_dMRI_ProbtrackX_ISOVF_fmi
362	IDP_dMRI_ProbtrackX_OD_fma	**439**	IDP_dMRI_ProbtrackX_ISOVF_ifo_l
363	IDP_dMRI_ProbtrackX_OD_fmi	**440**	IDP_dMRI_ProbtrackX_ISOVF_ifo_r
364	IDP_dMRI_ProbtrackX_OD_ifo_l	**441**	IDP_dMRI_ProbtrackX_ISOVF_ilf_l
365	IDP_dMRI_ProbtrackX_OD_ifo_r	**442**	IDP_dMRI_ProbtrackX_ISOVF_ilf_r
366	IDP_dMRI_ProbtrackX_OD_ilf_l	**443**	IDP_dMRI_ProbtrackX_ISOVF_mcp
367	IDP_dMRI_ProbtrackX_OD_ilf_r	**444**	IDP_dMRI_ProbtrackX_ISOVF_ml_l
368	IDP_dMRI_ProbtrackX_OD_mcp	**445**	IDP_dMRI_ProbtrackX_ISOVF_ml_r
369	IDP_dMRI_ProbtrackX_OD_ml_l	**446**	IDP_dMRI_ProbtrackX_ISOVF_ptr_l
370	IDP_dMRI_ProbtrackX_OD_ml_r	**447**	IDP_dMRI_ProbtrackX_ISOVF_ptr_r
371	IDP_dMRI_ProbtrackX_OD_ptr_l	**448**	IDP_dMRI_ProbtrackX_ISOVF_slf_l
372	IDP_dMRI_ProbtrackX_OD_ptr_r	**449**	IDP_dMRI_ProbtrackX_ISOVF_slf_r
373	IDP_dMRI_ProbtrackX_OD_slf_l	**450**	IDP_dMRI_ProbtrackX_ISOVF_str_l
374	IDP_dMRI_ProbtrackX_OD_slf_r	**451**	IDP_dMRI_ProbtrackX_ISOVF_str_r
375	IDP_dMRI_ProbtrackX_OD_str_l	**452**	IDP_dMRI_ProbtrackX_ISOVF_unc_l
376	IDP_dMRI_ProbtrackX_OD_str_r	**453**	IDP_dMRI_ProbtrackX_ISOVF_unc_r
377	IDP_dMRI_ProbtrackX_OD_unc_l		
378	IDP_dMRI_ProbtrackX_OD_unc_r		
379	IDP_dMRI_TBSS_ISOVF_Middle_cerebellar_peduncle		
380	IDP_dMRI_TBSS_ISOVF_Pontine_crossing_tract		
381	IDP_dMRI_TBSS_ISOVF_Genu_of_corpus_callosum		
382	IDP_dMRI_TBSS_ISOVF_Body_of_corpus_callosum		
383	IDP_dMRI_TBSS_ISOVF_Splenium_of_corpus_callosum		
384	IDP_dMRI_TBSS_ISOVF_Fornix		
385	IDP_dMRI_TBSS_ISOVF_Corticospinal_tract_R		

1460 Table S 12 Names of UK Biobank Non-Imaging Derived Phenotypes (nIDPs) related to Cardiovascular Health: 77. Supplement to **Figure 6**.

nIDP #	nIDP Name	nIDP #	nIDP Name
1	Pulse rate, automated reading (0.0),	**39**	Stroke volume during PWA (2.1),
2	Pulse rate, automated reading (0.1),	**40**	Mean arterial pressure during PWA (2.0),
3	Pulse rate, automated reading (2.0),	**41**	Mean arterial pressure during PWA (2.1),
4	Pulse rate, automated reading (2.1),	**42**	Ventricular rate (2.0),
5	Diastolic blood pressure, automated reading (0.0),	**43**	P duration (2.0),
6	Diastolic blood pressure, automated reading (0.1),	**44**	QRS duration (2.0),
7	Diastolic blood pressure, automated reading (2.0),	**45**	Systolic brachial blood pressure (2.0),
8	Diastolic blood pressure, automated reading (2.1),	**46**	Diastolic brachial blood pressure (2.0),
9	Systolic blood pressure, automated reading (0.0),	**47**	Cardiac index during PWA (2.0),
10	Systolic blood pressure, automated reading (0.1),	**48**	Cardiac index during PWA (2.1),
11	Systolic blood pressure, automated reading (2.0),	**49**	Pulse wave Arterial Stiffness index (0.0),
12	Systolic blood pressure, automated reading (2.1),	**50**	Pulse wave Arterial Stiffness index (2.0),
13	Pulse rate (0.0),	**51**	PQ interval (2.0),
14	Pulse rate (2.0),	**52**	QT interval (2.0),
15	Pulse wave reflection index (0.0),	**53**	QTC interval (2.0),
16	Pulse wave reflection index (2.0),	**54**	RR interval (2.0),
17	Pulse wave peak to peak time (0.0),	**55**	PP interval (2.0),
18	Pulse wave peak to peak time (2.0),	**56**	P axis (2.0),
19	Heart rate during PWA (2.0),	**57**	R axis (2.0),
20	Heart rate during PWA (2.1),	**58**	T axis (2.0),
21	Peripheral pulse pressure during PWA (2.0),	**59**	QRS num (2.0),
22	Central systolic blood pressure during PWA (2.0),	**60**	LV ejection fraction (2.0),
23	Central pulse pressure during PWA (2.0),	**61**	LV end diastolic volume (2.0),
24	Number of beats in waveform average for PWA (2.0),	**62**	LV end systolic volume (2.0),
25	Number of beats in waveform average for PWA (2.1),	**63**	LV stroke volume (2.0),
26	Central augmentation pressure during PWA (2.0),	**64**	Cardiac output (2.0),
27	Central augmentation pressure during PWA (2.1),	**65**	Average heart rate (2.0),
28	Augmentation index for PWA (2.0),	**66**	Minimum carotid IMT (intima-medial thickness) at 120 degrees (2.0),
29	Augmentation index for PWA (2.1),	**67**	Mean carotid IMT (intima-medial thickness) at 120 degrees (2.0),
30	Cardiac output during PWA (2.0),	**68**	Maximum carotid IMT (intima-medial thickness) at 120 degrees (2.0),
31	Cardiac output during PWA (2.1),	**69**	Minimum carotid IMT (intima-medial thickness) at 150 degrees (2.0),
32	End systolic pressure during PWA (2.0),	**70**	Mean carotid IMT (intima-medial thickness) at 150 degrees (2.0),
33	End systolic pressure during PWA (2.1),	**71**	Maximum carotid IMT (intima-medial thickness) at 150 degrees (2.0),
34	End systolic pressure index during PWA (2.0),	**72**	Minimum carotid IMT (intima-medial thickness) at 210 degrees (2.0),
35	End systolic pressure index during PWA (2.1),	**73**	Mean carotid IMT (intima-medial thickness) at 210 degrees (2.0),
36	Total peripheral resistance during PWA (2.0),	**74**	Maximum carotid IMT (intima-medial thickness) at 210 degrees (2.0),
37	Total peripheral resistance during PWA (2.1),	**75**	Minimum carotid IMT (intima-medial thickness) at 240 degrees (2.0),
38	Stroke volume during PWA (2.0),	**76**	Mean carotid IMT (intima-medial thickness) at 240 degrees (2.0),
		77	Maximum carotid IMT (intima-medial thickness) at 240 degrees (2.0)

1461

1462

1463 Table S 13 Names of UK Biobank Non-Imaging Derived Phenotypes (nIDPs) related to Cognition: 68. Supplement to **Figure 6**.

nIDP #	nIDP Name
1	Number of incorrect matches in round (0.1)
2	Number of incorrect matches in round (0.2)
3	Number of incorrect matches in round (2.1)
4	Number of incorrect matches in round (2.2)
5	Number of incorrect matches in round (2.3)
6	Number of times snap-button pressed (0.0)
7	Number of times snap-button pressed (0.1)
8	Number of times snap-button pressed (0.2)
9	Number of times snap-button pressed (0.3)
10	Number of times snap-button pressed (0.4)
11	Number of times snap-button pressed (0.10)
12	Number of times snap-button pressed (0.11)
13	Number of times snap-button pressed (2.0)
14	Number of times snap-button pressed (2.1)
15	Number of times snap-button pressed (2.2)
16	Number of times snap-button pressed (2.3)
17	Number of times snap-button pressed (2.4)
18	Number of times snap-button pressed (2.5)
19	Number of times snap-button pressed (2.7)
20	Number of times snap-button pressed (2.10)
21	Number of times snap-button pressed (2.11)
22	Prospective memory result (0.0)
23	Prospective memory result (2.0)
24	FI8 : chained arithmetic (2.0)
25	FI9 : concept interpolation (2.0)
26	Duration to complete numeric path (trail #1) (2.0)
27	Mean time to correctly identify matches (2.0)
28	Number of incorrect matches in round (0.0)
29	Time elapsed (2.0)
30	Time elapsed (2.1)
31	Time elapsed (2.2)
32	Time elapsed (2.3)
33	Time elapsed (2.4)
34	Time elapsed (2.5)
35	Time elapsed (2.6)
36	Time elapsed (2.7)
37	Time elapsed (2.8)
38	Digits entered correctly (2.0)
39	Digits entered correctly (2.1)
40	Digits entered correctly (2.2)
41	Digits entered correctly (2.3)
42	Digits entered correctly (2.4)
43	Digits entered correctly (2.5)
44	Digits entered correctly (2.6)
45	Digits entered correctly (2.7)
46	Total errors traversing numeric path (trail #1) (2.0)
47	Mean time to correctly identify matches (0.0)
48	Fluid intelligence score (0.0)
49	Fluid intelligence score (2.0)
50	Duration to complete numeric path (trail #1) (0.0)
51	Duration to complete alphanumeric path (trail #2) (0.0)
52	Digits entered correctly (2.8)
53	Maximum digits remembered correctly (2.0)
54	Attempted fluid intelligence (FI) test. (2.0)
55	FI1 : numeric addition test (0.0)
56	FI1 : numeric addition test (2.0)
57	FI3 : word interpolation (0.0)
58	FI3 : word interpolation (2.0)
59	FI4 : positional arithmetic (0.0)
60	FI4 : positional arithmetic (2.0)
61	FI5 : family relationship calculation (0.0)
62	FI5 : family relationship calculation (2.0)
63	FI6 : conditional arithmetic (0.0)
64	FI6 : conditional arithmetic (2.0)
65	FI7 : synonym (0.0)
66	FI7 : synonym (2.0)
67	Duration to complete alphanumeric path (trail #2) (2.0)
68	Number of puzzles correctly solved (2.0)

1464

1465

1466 Table S 14 Names of Human Connectome Project Non-Imaging Derived Phenotypes (nIDPs) used in CCA. Supplement to **Figure 6**.

nIDP #	nIDP Name	nIDP #	nIDP Name
1	PicVocab_Unadj	**80**	ASR_Witd_Pct
2	PicVocab_AgeAdj	**81**	IWRD_TOT
3	PMAT24_A_CR	**82**	PainInterf_Tscore
4	DDisc_AUC_200	**83**	MMSE_Score
5	THC	**84**	SSAGA_Alc_12_Frq_Drk
6	LifeSatisf_Unadj	**85**	Odor_Unadj
7	ListSort_AgeAdj	**86**	SSAGA_Alc_D4_Ab_Sx
8	ReadEng_Unadj	**87**	SSAGA_Mj_Use
9	SCPT_SPEC	**88**	ASR_Aggr_Raw
10	ReadEng_AgeAdj	**89**	SSAGA_Mj_Ab_Dep
11	ListSort_Unadj	**90**	DSM_Somp_Raw
12	DDisc_AUC_40K	**91**	FearSomat_Unadj
13	Avg_Weekday_Any_Tobacco_7days	**92**	SSAGA_Alc_12_Drinks_Per_Day
14	Num_Days_Used_Any_Tobacco_7days	**93**	Mars_Log_Score
15	Total_Any_Tobacco_7days	**94**	SelfEff_Unadj
16	PicSeq_AgeAdj	**95**	SCPT_SEN
17	FamHist_Fath_DrgAlc	**96**	NEOFAC_N
18	PicSeq_Unadj	**97**	SSAGA_Agoraphobia
19	Avg_Weekday_Cigarettes_7days	**98**	ASR_Intn_T
20	Avg_Weekend_Any_Tobacco_7days	**99**	AngHostil_Unadj
21	Total_Cigarettes_7days	**100**	Num_Days_Drank_7days
22	Dexterity_AgeAdj	**101**	SSAGA_Times_Used_Cocaine
23	Avg_Weekend_Cigarettes_7days	**102**	Loneliness_Unadj
24	Dexterity_Unadj	**103**	ASR_Intn_Raw
25	Times_Used_Any_Tobacco_Today	**104**	SSAGA_Alc_Hvy_Drinks_Per_Day
26	PSQI_Score	**105**	MeanPurp_Unadj
27	AngAggr_Unadj	**106**	DSM_Avoid_Pct
28	Taste_AgeAdj	**107**	NEOFAC_E
29	ASR_Rule_Raw	**108**	Total_Beer_Wine_Cooler_7days
30	Taste_Unadj	**109**	DSM_Avoid_Raw
31	ASR_Thot_Raw	**110**	Avg_Weekday_Wine_7days
32	EVA_Denom	**111**	Flanker_AgeAdj
33	SSAGA_TB_Still_Smoking	**112**	ASR_Anxd_Pct
34	FamHist_Fath_None	**113**	Avg_Weekend_Beer_Wine_Cooler_7days
35	ASR_Thot_Pct	**114**	SSAGA_Alc_D4_Ab_Dx
36	PercStress_Unadj	**115**	Total_Drinks_7days
37	ProcSpeed_AgeAdj	**116**	SSAGA_Alc_Hvy_Max_Drinks
38	ASR_Rule_Pct	**117**	FearAffect_Unadj
39	ProcSpeed_Unadj	**118**	Total_Wine_7days
40	DSM_Antis_Raw	**119**	Avg_Weekday_Drinks_7days
41	ER40_CR	**120**	ER40SAD
42	NEOFAC_A	**121**	Flanker_Unadj
43	ASR_Crit_Raw	**122**	ER40FEAR
44	VSPLOT_TC	**123**	Avg_Weekday_Beer_Wine_Cooler_7days
45	NEOFAC_O	**124**	SSAGA_Times_Used_Illicits
46	ER40ANG	**125**	Avg_Weekend_Drinks_7days
47	VSPLOT_OFF	**126**	SSAGA_Alc_D4_Dp_Sx
48	SSAGA_Times_Used_Stimulants	**127**	NEOFAC_C
49	ASR_Soma_Pct	**128**	Total_Hard_Liquor_7days
50	SSAGA_Mj_Times_Used	**129**	Correction
51	DSM_Antis_Pct	**130**	SSAGA_Alc_Hvy_Frq_5plus
52	CardSort_AgeAdj	**131**	DSM_Adh_Pct
53	ASR_Extn_Raw	**132**	ASR_Attn_Pct
54	ASR_Oth_Raw	**133**	VSPLOT_CRTE
55	ASR_Totp_T	**134**	SSAGA_Depressive_Ep
56	ASR_Extn_T	**135**	AngAffect_Unadj
57	ASR_Totp_Raw	**136**	SSAGA_PanicDisorder
58	EmotSupp_Unadj	**137**	Avg_Weekend_Hard_Liquor_7days
59	DSM_Anxi_Pct	**138**	FamHist_Moth_Dep
60	PercReject_Unadj	**139**	ASR_Anxd_Raw
61	ER40NOE	**140**	SSAGA_Times_Used_Opiates
62	DSM_Anxi_Raw	**141**	SSAGA_Times_Used_Sedatives
63	ASR_TAO_Sum	**142**	SSAGA_Alc_Hvy_Frq
64	SSAGA_TB_Smoking_History	**143**	SSAGA_Alc_12_Frq_5plus
65	CardSort_Unadj	**144**	Friendship_Unadj
66	PosAffect_Unadj	**145**	SSAGA_Depressive_Sx
67	SSAGA_ChildhoodConduct	**146**	ASR_Attn_Raw
68	Odor_AgeAdj	**147**	ASR_Intr_Raw
69	ASR_Witd_Raw	**148**	SSAGA_Alc_12_Frq
70	SSAGA_Alc_Hvy_Frq_Drk	**149**	FamHist_Fath_Dep
71	ASR_Soma_Raw	**150**	InstruSupp_Unadj
72	DSM_Depr_Pct	**151**	ASR_Intr_Pct
73	ASR_Aggr_Pct	**152**	SSAGA_Times_Used_Hallucinogens
74	SSAGA_Alc_12_Max_Drinks	**153**	Avg_Weekend_Wine_7days
75	DSM_Depr_Raw	**154**	FamHist_Moth_None
76	Mars_Final	**155**	Sadness_Unadj
77	PercHostil_Unadj	**156**	DSM_Hype_Raw
78	DSM_Somp_Pct	**157**	DSM_Adh_Raw
79	SSAGA_Alc_Age_1st_Use	**158**	DSM_Inat_Raw

1467