Supplementary Information and Figures for The Human Connectome Project's Neuroimaging Approach

Matthew F. Glasser[1], Stephen M. Smith[2], Daniel S. Marcus[1], Jesper Andersson[2], Edward J. Auerbach[3], Timothy E. J. Behrens[2], Timothy S. Coalson[1], Michael P. Harms[4], Mark Jenkinson[2], Steen Moeller[3], Emma C. Robinson[2], Stamatios N. Sotiropoulos[2], Junqian Xu[5], Essa Yacoub[3], Kamil Ugurbil[3], David C. Van Essen[1]

[1]Department of Neuroscience, Washington University Medical School, Saint Louis, MO, USA; [2]FMRIB Centre, Nuffield Department of Clinical Neurosciences, John Radcliffe Hospital, University of Oxford, Oxford, UK; [3]Center for Magnetic Resonance Research (CMRR), University of Minnesota, Minneapolis, MN, USA; [4]Department of Psychiatry, Washington University Medical School, Saint Louis, MO; [5]Translational and Molecular Imaging Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA

The Supplementary Information is divided into three main parts. Part I (Supplementary Topics) includes 18 brief text topics, each expanding on a specific technical or discussion issue in the order it was raised in the main text. Part II (Supplementary Figures) includes 11 figures that illustrate important points raised in the main text or supplementary topics. Supplementary References follow in Part III.

I. Supplementary Topics

1. What do the terms T1-weighted and T2-weighted mean in terms of MRI and myelin contrast?

Students learning about brain imaging (for research or clinical purposes) are exposed early on to the distinction between "T1-weighted (T1w)" and "T2-weighted (T2w)" images. T1w images have an intensity profile of white matter (WM) > grey matter (GM) > cerebrospinal fluid (CSF). In contrast, T2w images have the opposite intensity profile (CSF > GM > WM). Although these terms were originally based on the types of MR contrast in the images from simple gradient echo or spin echo sequences, they have come to refer more to the direction of contrast in MRI images given the increasing complexity of the MR pulse sequences currently in use for 3D high resolution brain imaging (and scanner manufacturers, radiologists, and many brain imagers use the terms in this way). Two particular 3D sequences in current widespread use are 3D MPRAGE (magnetization prepared rapid gradient echo)[1] and 3D SPACE (single slab variable flip angle turbo/fast spin echo[2]. However, none of these terms is completely accurate in describing the MR contrast mechanisms that create each image. The T1w contrast (WM>GM>CSF) in an MPRAGE image is dominated by the T1 relaxation mechanism but also has T2* and proton density (PD) components[3]. On the other hand, the contrast in a SPACE image can be manipulated via the variable flip angle scheme for the refocusing RF pulse train[4] to generate T2w[5], T1w[6], or proton density weighted (PDw) images. The T2w contrast (CSF>GM>WM), which contains both T1 and T2 relaxation mechanisms[5], is the most commonly used contrast for 3D SPACE and was the choice in the HCP's T2w SPACE protocol. Given the complexity surrounding the MR contrast mechanisms in 3D structural acquisitions, we refer to the HCP MPRAGE and SPACE acquisitions simply as T1w and T2w images, consistent with their intensity profiles (albeit with no intention to imply that just a single underlying contrast mechanism is present). We also note that in our experience the best results for T1w and T2w

Nature Neuroscience: doi:10.1038/nn.4361

images are achieved using the 3D MPRAGE and SPACE sequences at 0.8 mm isotropic resolution or better, half the thickness of the thinnest cortex.

T1w and T2w images can be used to compute a T1w/T2w ratio in order to map myelin content in cortical grey matter[7]. As mentioned above, T1, T2, T2*, and PD all contribute to some degree to the contrast in the T1w and T2w images. The lipids in myelin affect T1 contrast[8], and iron, which colocalizes with myelin particularly in cortical grey matter[9], affects T2* contrast[8]. Proton density also appears to contribute to grey/white contrast and may also contribute to intracortical myelin contrast[10]. Thus, many types of MR contrast also contribute to the T1w/T2w ratio and hence to the grey matter myelin maps produced by this ratio. We treat these maps as an in vivo "myelin stain," rather than a quantitative measure of MR properties, and their validation is primarily neuroanatomical, rather than biophysical (i.e. they correlate strongly with known patterns of myelination within the grey matter[7]). Though there has been debate as to whether the T1w/T2w ratio is the best approach for in vivo myelin mapping (as opposed to alternatives such as quantitative T1 images[10-14], quantitative magnetization transfer images, quantitative T2* images[15], T1w/T2*w ratio images[16], myelin water fraction images, etc), we are unaware of an approach that produces as high a contrast-to-noise ratio (CNR) per unit acquisition time and field strength, enabling spatial resolution to be maximized within a fixed scan time (< 13 min combined between the T1w and T2w acquisitions at 3T and 0.8mm isotropic). Also, T1w and T2w images are both very useful for generating accurate surface models and subcortical segmentations[17]. Thus, we recommend that most studies acquire both images (and then the T1w/T2w myelin maps come 'for free'), but studies focusing on quantification may elect instead to use their available scan time to acquire quantitative MR parameter maps. A FLAIR type acquisition (also based on the 3D SPACE sequence with an inversion recovery preparation) could be used in place of the T2w SPACE acquisition if FLAIR contrast is important for other reasons. However, such FLAIR images do not have as much contrast for myelin or as high SNR as the T2w images mentioned before, so the myelin maps they produce are noisier.

2. Why is the HCP's Diffusion Protocol Special?

The HCP's diffusion protocol differs from traditional diffusion imaging in several important ways. It uses many more diffusion encoding directions (270) enabled by an hour-long scanning session and three-fold multi-band acceleration (MB=3), which incurs little SNR penalty. The WU-Minn-Ox HCP dMRI data were acquired on a customized scanner with stronger diffusion encoding gradients that can operate with 100 mT/m maximal gradient strength (G_{max}). This reduces diffusion encoding times (TE) compared to standard gradients of 40 mT/m (Siemens Trio) for a given diffusion weighting magnitude (b value). Consequently, signal loss due to T2 decay during diffusion encoding is significantly less and the resultant gain in SNR is traded for higher spatial resolution. HCP dMRI data achieved 1.25 mm isotropic resolution with sufficient SNR for b-values as high as 3000 s/mm^2 and used b shells of 1000, 2000, and 3000 s/mm^2 [18]. (Note that Siemens subsequently introduced the 3T Prisma, developed in part through their collaboration with the HCP with a maximum gradient strength of 80 mT/m, which enables users in the community to use HCP-style diffusion protocols). The 7T data also have high SNR with an even higher spatial resolution (1.05 mm isotropic), 130 directions at b-values of 1000 and 2000 s/mm^2 [19] (see Supplementary Fig. 2). The MGH/UCLA HCP consortium took a different approach with their even more powerful gradients (300 mT/m G_{max}) by scanning subjects at 1.5mm isotropic resolution and b values up to 10000 s/mm^2 [20]. Both HCP consortia chose to use the original monopolar (i.e., Stejskal-Tanner[21]) diffusion encoding sequence. It has a shorter TE compared to the bipolar (twice-refocused) sequence[22] that has been popular for whole brain diffusion MRI in the last decade because it reduces eddy currents at the expense of an increased TE (and thereby reduced SNR). The HCP could tolerate substantially worse eddy

Nature Neuroscience: doi:10.1038/nn.4361

current effects with the monopolar diffusion encoding at high b-value because eddy current distortions are instead corrected during post-processing using a newly developed approach[23] that goes beyond traditional affine image registration (see Topic #8 below). The HCP also acquired two sets of scans having reversed phase encoding directions. This enables correction of signal that had piled up in one phase encoding direction and spread out in the other without an SNR penalty. Indeed SNR actually increases across the image when both images are averaged, producing a relatively unbiased estimate of the signal intensities throughout the final image[24]). Tractography examples indicative of the data quality are shown in Sotiropoulos et al.[18]. Another example (Supplementary Fig. 3) illustrates the ability to precisely map patterns of cortico-striatal connectivity from human HCP tractography data that closely parallel the tracer-based connectivity patterns reported in the macaque.

3. Simultaneous Multi-slice (Multi-Band) Echo Planer Imaging (EPI)

The high spatial and temporal resolution of the HCP data would not have been possible without the accelerated imaging enabled by Simultaneous Multi-slice (Multi-Band) EPI. This transformative development has become the new standard for EPI imaging, the workhorse sequence for brain fMRI and dMRI. In MR imaging, there is always a tradeoff between resolution and data acquisition times. Higher resolution data simply take longer to acquire. In fMRI, this tension leads to longer TRs (i.e. worse temporal resolution) with increasing spatial resolution. In dMRI, the trade-off is between spatial resolution and angular resolution (in q-space). The approach adopted by the HCP to substantially alleviate this tradeoff is the simultaneous multi-slice (SMS), multi-band (MB) technique that was first employed in brain imaging to reduce the time to acquire very high resolution functional imaging data[25]. In conventional EPI image acquisition, each 2D slice covering the volume of interest is acquired sequentially in time. In SMS/MB, multiple slices are excited and acquired simultaneously using multi-banded pulses. The resultant image represents the sum of the simultaneously excited slices. These overlapping slices can be unaliased using spatial information inherent in the multi-channel array of coils that are used for receiving the signals (see Supplementary Fig. 4). These coils are sensitive to regions of the brain in close proximity to their physical location and lose sensitivity with increasing distance. Each coil yields a linear combination of signals from the different slices (weighted by their sensitivity profiles), and a matrix inversion provides a solution to separate the slices from one another. The imaging time is reduced by the number of simultaneously excited slices (MB factor), which receive the same contrast encoding. This is especially important for EPI-based fMRI and dMRI, as contrast encoding time usually outweighs the EPI readout time. The MB factor cannot be increased too high, however, as it is affected by the number of receive coils in the array, and their size, proximity to the sample, and layout. There is a penalty to be paid in noise enhancement and cross-slice contamination with increasing MB factors, which must be kept to a negligible level, but these parameters can be measured to determine feasible MB acceleration factors. Displacing slices with respect to each other at the image acquisition stage prior to unaliasing significantly improves the maximum usable MB factors[26]. In the HCP's 3T fMRI data acquisition, an MB factor of 8 was used with a 32 channel array coil after extensive testing. This allowed whole brain 2 mm isotropic coverage in 0.72s as opposed to 5.6s with a single band fMRI sequence, significantly improving data quality and SNR efficiency[27]. For HCP diffusion MRI, a MB factor of 3 was employed, making it feasible to obtain high spatial and q-space resolution data. In dMRI, the ultimate limitation in acceleration is the T_1 relaxation of spins, as going faster than approximately $2xT_1$ per volume of interest (whole brain for the HCP) becomes SNR-inefficient. However, when high spatial resolution is employed in dMRI, MB factors are restricted by RF power deposition before reaching this ultimate limit. The dMRI sequence requires more power than the fMRI sequence because, as a spin echo sequence, it requires the use of a 180-degree refocusing pulse, in

Nature Neuroscience: doi:10.1038/nn.4361

addition to the 90-degree excitation pulse also present in a gradient echo fMRI sequence. Further, the MB pulses aggravate the situation because they also increase the power demands. The combination is limited by the safely achievable peak power and total power deposited in the subject. There are solutions to these limitations developed by the HCP[19,28,29], some of which were employed in the 7 Tesla dMRI component of the HCP data acquisition.

4. HCP-Short Protocols

The HCP-Short protocols are similar in key respects to the HCP-Full protocol, but are less than half the total scan duration (~1-2 h total scanning over one or two sessions vs ~4 h over four sessions). HCP-Short is designed for studies in which the subject population may not tolerate intensive multi-session scanning and individual scans of 15 min duration, owing to age and/or brain disorders. The HCP-Short protocols aim to provide sufficient high-quality data to enable HCP-style analyses, including selective removal of artifacts, intersubject registration using areal-feature-based alignment (e.g., MSMAll), and cortical parcellation using the areal classifier. There are multiple variants of the HCP-Short protocols, because some protocols have project-specific aspects (e.g., related to the particular age range or brain disorder under investigation). The initial HCP-Short protocol was developed for the HCP Lifespan Pilot project, a 1-year effort by both HCP consortia to obtain pilot data for the recently launched HCP Lifespan Development and Lifespan Aging projects (see Figure 5 in main text). We describe below adaptations of the Lifespan HCP-Short protocols proposed for the HCP-Aging and HCP-Development Lifespan projects, which will both use the 3T Prisma platform at each of the four consortium data acquisition sites. Note that these protocols are currently undergoing further piloting, with final protocols and parameters to be available at http://protocols.humanconnectome.org/lifespan/.

Structural scans for HCP-Short are similar to that for HCP-Full, but the spatial resolution is slightly lower (0.8 mm isotropic voxels instead of 0.7 mm), and only a single pair of T1w and T2w scans is collected (albeit with a repeated acquisition if the scanning technician deems them to be only “fair” or worse quality). This increases SNR and shortens scan duration slightly (14 min total for T1w and T2w vs 16 min) while preserving the benefits of high spatial resolution for cortical segmentation and maps of myelin and cortical thickness. Additionally, the Lifespan HCP may use ‘volumetric navigators’ (vNavs) to record (and possibly correct for) within-scan head motion[30,31] – see also Topic #4 below.

For fMRI scans, the HCP-Short protocol acquires data at the same spatial and temporal resolution as for HCP-Full (2 mm isotropic; TR = 0.72-0.8 s) using a MB factor of 8. Phase encoding is along the A/P axis for the Prisma (vs the L/R axis for the customized WU Connectom scanner). In addition, HCP-Short acquires a spin-echo AP-PA field map pair.

For rfMRI, HCP-Short scans are shorter in duration (usually 5.0-5.5 min/scan, but 3 min/scan for younger children, vs 15 min/scan for HCP-Full), and in total scan time (30-40 min vs 1 h). The reproducibility of resting-state networks (RSNs) increases steeply with total scan time over the first 30 min of rfMRI acquisitions (with additional improvement up to 80-100 min)[32]. It may be possible to augment information about RSNs using task-fMRI scans (see below), given that functional connectivity is similar when measured during tasks or the resting state[33], and this may add to the useful data for intersubject alignment and individual-subject parcellation.

The task-fMRI aspect of the HCP-Short protocols utilized to date is much briefer (ranging from no task fMRI to ~30 min total, vs 1 h for HCP-Full) and allows for project-specific tasks appropriate for particular age ranges or brain disorders. The HCP Lifespan projects are currently considering options to acquire task data using up to five tasks, each with a notably

Nature Neuroscience: doi:10.1038/nn.4361

short (~2 min) scan duration. More generally, the HCP-style paradigm supports diverse task fMRI protocols to reflect project-specific aims without major adverse impacts on intersubject registration ("MSMAll" surface registration does not use task-fMRI activations) or areal classification (which also does not require task-fMRI activations, cf. Ref. [34]).

For dMRI, HCP-Short acquires 22 min of dMRI data (four 5.5 min scans) at a spatial resolution of 1.5 mm isotropic voxels (vs. 1.25 mm for HCP-Full), two b values (1,500 and 3,000 s/mm^2) and ~75-90 directions/shell. The lower maximal gradient strength for the Prisma (80 mT/m, vs 100 mT/m for the WU customized Connectom scanner) yields a longer TE and slightly lower SNR per image, but HCP-Short uses a simultaneous multi-slice ("multi-band") factor of 4 for dMRI on the Prisma (vs. 3 for main HCP) while keeping within SAR and peak power limits because of additional pulse sequence development that occurred after the main HCP began. Possible refinements are being evaluated that would enable a MB factor of 5 or 6 using the two channel parallel transmit (pTx) capability of the Prisma and the local and global SAR regularized multi-band pTx pulses recently developed at UMinn[29,35,36].

The Lifespan HCP-Short protocol will also likely include a ~5 min Arterial Spin Labeling (ASL) perfusion MRI scan, which provides a quantitative measurement of cerebral blood flow (CBF). Since perfusion is normally coupled to metabolism, ASL is a useful surrogate marker of basal brain function, as well as aiding in the discrimination between neuronally-related vs cardiovascular-related effects seen in BOLD imaging. The Lifespan HCP projects are currently evaluating both background suppressed 3D segmented GRASE pseudo-continuous ASL (pCASL)[37] and 2D multi-band EPI based pCASLa[38].

5. Head Movement During 3D Structural Scans

Head movement during 3D T1w and T2w scans produces currently uncorrectable image blurring and banding and is a major problem for both research and clinical neuroimaging. Even a single translation of a millimeter or so during a scan of 6 – 8 min duration will cause blurring. The main HCP reduced the problem by acquiring two T1w and T2w images. Moreover, if these were not of adequate quality - rated good or excellent based on a systematic manual QC evaluation (http://humanconnectome.org/documentation/S900/HCP_S900_Release_Appendix_IV.pdf, parts I – III, pp. 17 - 28), subjects were asked to complete an additional scanning session where up to two pairs of T1w and T2w images were acquired. All T1w and T2w scans from the same session that were rated good or excellent in overall quality were used in further processing. Though the HCP invested in an external motion tracking camera, this technology was in its very early stages, and problems associated with tracking chip attachment to skin and skin movement independent of the rigid body motion of the head were not solved. As a result this approach was abandoned early in the project. An alterative approach that shows promise involves 'navigator' sequences that acquire low resolution EPI data within the dead time of T1w and T2w sequences. This data can be used to detect subject motion, and, in addition, adjust the imaging gradients to compensate for subject movement, avoiding blurring issues, and reacquire lines in k-space that were acquired during subject movement so as not to use k-space data corrupted by motion in the final Fourier reconstruction. If the robustness of these approaches can be established, future studies may wish to take advantage of such online movement correction[30,31,39] to enable consistently high quality structural image acquisitions.

Nature Neuroscience: doi:10.1038/nn.4361

6. Gradient Distortion

Spatial localization as well as diffusion encoding is achieved through the gradients of the MR system. However, the gradients are not perfectly linear in space, which leads to image distortions as well as errors in diffusion encoding. For dMRI in particular, such gradient nonlinearities can influence the magnitude and direction of diffusion-sensitizing gradients[40] on top of the spatial distortions that they induce. Spatial distortions due to gradient nonlinearities need to be corrected if images acquired on different scanners are to be compared. Also the correct diffusion weighting in magnitude and direction must be used if one is to avoid errors in the estimated diffusion parameters that drive tractography algorithms. In the HCP these corrections were implemented within the HCP's minimal preprocessing pipelines[17].

A second problem exists with respect to calibration of the gradients. For example, if the sequence calls for 40 mT/m, do we actually get 40 mT/m or just 39.2 mT/m? Usually the calibrations are within a small percentage of the targeted value, but they are not exact. This was initially a problem for combining 3T and 7T data on the same subject in the main HCP. These issues were carefully evaluated and calibrated using a grid "phantom" and corrected in postprocessing, by enabling additional degrees of freedom in the boundary-based registration of the HCP's minimal preprocessing pipelines[17,41] (from 6 DOF to 12 DOF).

7. Why Spin Echo Field Maps?

Maps of susceptibility-induced image distortion can be derived from a pair of spin-echo EPI images with reversed phase encoding directions. We call these "Spin Echo Field Map" scans in the HCP-style protocols. They offer several advantages over the more traditional dual-echo, gradient-echo based field maps, where the magnetic field is computed based on the phase difference of images from two differing echo times. Spin echo field maps are faster to acquire (and thus more robust to subject motion), and can be matched exactly to the gradient echo fMRI data in terms of resolution, geometry, and distortion (via matching the echo spacing). As a result, the distortion is identical between the two images, making image registration of the distorted spin echo and gradient echo fMRI images substantially easier. The HCP tested the two distortion correction approaches and found that they performed similarly, leading to the recommendation that the faster imaging approach be used. Importantly, additional information can be obtained from the spin echo images. Gradient echo fMRI images suffer from signal dropout due to dephasing in areas of high magnetic field inhomogeneity that is absent in the spin echo field map images with similar image distortion. This difference can be used to segment these regions of dropout, and these segmentations can serve as useful information to know where the fMRI signal is likely to be unreliable. Additionally, the spin echo and gradient echo images can be combined to estimate the difference in the effects of the transmit B1 field on each of the images plus an estimate of the receive B1 field on the fMRI image[34]. This receive field is more accurate than that generated from T1w and T2w images acquired in a different imaging session (and which experience somewhat different transmit field effects), and thus can be used to bias-correct the fMRI data without the need to acquire additional sequences or use additional preparatory time on the scanner. For some fMRI analyses (e.g., resting-state correlations and task z-statistics), bias correction of fMRI is not needed. However, for other analyses (e.g., beta maps, variance maps), an fMRI intensity bias field can make the data less biologically interpretable across the whole image. For such analyses, accurate intensity bias correction is important[34]. The latest version of the HCP's minimal preprocessing pipelines (v3.15.1 or later) now enables this sort of correction; moreover, as a simple multiplicative bias field, it can be retrospectively applied to prior HCP data as well.

Nature Neuroscience: doi:10.1038/nn.4361

8. Eddy Current Distortion Correction and dMRI Outlier Removal

A major problem for dMRI data arises from eddy currents induced by magnetic field modulations that are distinct for each diffusion direction and are more pronounced in the mono-polar diffusion-encoding scheme chosen by the HCP (for SNR efficiency as mentioned in Topic #2). dMRI data also suffer from signal dropouts caused by head movement and other sources, analogous to those that plague fMRI data. The HCP's 'eddy' software (released in FSL, and not to be confused with the older 'eddy_correct' script based on affine registration) corrects for linear and nonlinear eddy current distortions and also has an option to use model-based outlier detection and replacement to correct corrupted individual slices[24,42,43]. Counter to the typical practice of individually registering each diffusion-weighted volume to a reference, eddy uses all data volumes jointly. A Gaussian-process-based generative data model is used to make predictions about what the 3D dMRI volumes should look like at every point in q-space given knowledge of the different distortions. Inversion of this model enables accurate estimation of the distortion fields, which can then be applied to correct the data. This combination of monopolar acquisition and eddy-based post-processing results in fewer artifacts and less distortion than traditionally achieved, yet with lower TE and therefore higher SNR that the HCP used to achieve higher spatial resolution (see Topic #2).

9. More about the Global Signal and fMRI Data Temporal Cleanup

One idea that has gained substantial currency in the literature is that global signal fluctuations are often induced by direct effects of subject movement[44-46]. Invoking this hypothesis, some argue that it is necessary to regress out the global signal in order to adequately correct for subject movement. However, to our knowledge, no direct biophysical causal mechanism has been proposed to account for subject movement leading to global fluctuations in the fMRI signal (particularly those with the tissue specific and area specific effects shown in Supplementary Figs. 5 and 6). Instead, it is worth considering the kinds of effects seen in fMRI images that are clearly attributable to head movement and which can be removed using ICA+FIX (ICA+FIX consists of detrending, aggressive 24 movement parameter regression, and non-aggressive noise ICA component regression): 1) Largely linear effects of moving the head within an inhomogeneous receive coil. These effects tend to produce "rings" around the perimeter of the brain, reflecting signal intensity increases as the head gets closer to the head coil on one side and decreases as the head gets further from the coil on the other side. Such spatially specific effects are readily identified by ICA+FIX and eliminated from the data. 2) Complex interactions between head motion and the magnetic field inhomogeneities in the fronto-polar, orbito-frontal, anterior temporal, and inferior temporal cortices. These effects are also spatially specific and are readily identified by ICA+FIX and eliminated from the data. 3) "Spin history" effects of movement caused by some portions of the brain being excited sooner than they should and other parts being excited later than they should by the slice selection profiles. In a 'worst case' scenario where the head moved exactly one slice up or down in the slice direction, this would produce a strong banding pattern in the data. Such bands are readily identified by ICA+FIX and eliminated from the data. In summary, no plausible biophysical mechanism has been proposed by which head movement would induce a truly global change in fMRI signal intensity (short of the subject completely exiting the head coil), as opposed to spatially specific changes in fMRI signal intensity that are visible to spatial ICA-based approaches. Further, the global fMRI signals are highly tissue and area specific (Supplementary Figs. 5 and 6), and no plausible biophysical mechanism has been proposed for how subject motion could directly cause such phenomena. Instead, the hypothesis advanced in the main text Box 1 about global fluctuations being partly due to non-neuronal physiological fluctuations and partly to subject-specific neural effects seems much more likely.

Nature Neuroscience: doi:10.1038/nn.4361

Approaches such as ICA+FIX and similar methods[47,48] are attractive for removing the direct effects of movement from fMRI data (along with other spatially specific effects like MRI physics related artifacts or arterial and venous pulsation artifacts). Indeed, ICA+FIX results approach that of the more aggressive censoring ('scrubbing') approach when movement becomes severe, but are not forced to make binary 'scrub/not scrub' choices based on thresholds in framewise displacement (FD) and/or framewise image intensity difference (DVARS) measures that may themselves be quite noisy in high temporal resolution data. Instead ICA+FIX derives a weighted solution directly from the data. That being said, it remains an open question whether scrubbing (for example, based on those frames where ICA+FIX begins to approach the scrubbing solution by removing substantial amounts of unstructured noise variance from the data) would be useful if we put aside for a moment the effects of global noise. A similar question can be asked regarding the standard bandpass filtering that is often using to clean fMRI data, but has not been demonstrated to be useful above and beyond ICA+FIX and appropriate global noise cleanup. Unfortunately, white matter and CSF timeseries regression are unlikely to represent an effective approach to global noise cleanup, as they have relatively little correlation with the global fluctuations as long as they are derived from ROIs free from grey matter contamination (which can arise from spatial smoothing or partial volume effects of 'white matter' voxels encroaching into gray matter, particularly in low resolution datasets), see also Supplementary Fig. 5. Similarly it is not clear that physiological noise modeling approaches, such as that currently implemented in FSL's PMN[47] will deal with all global noise, though they may help. Imperfections in the modeled effects of physiological noise parameters on BOLD fMRI data will lead to incomplete physiological noise clean up. As mentioned in the main text Box 1, a recommendation for the complete approach to fMRI timeseries noise cleanup will have to wait until a selective and effective approach to global noise cleanup becomes available.

Besides ICA+FIX[48,49] and ICA-AROMA[50,51], Multi-Echo-ICA (ME-ICA) is another approach that uses ICA to identify and remove non-BOLD noise components from a functional timeseries[52,53]. In this approach, multiple images are acquired at different echo times (TEs), and each component is scored as to how well it fits a T2* or BOLD-like decay curve. Importantly, although this approach is able to separate BOLD-like from non-BOLD-like components, it does not remove BOLD-like global fluctuations from physiological sources (e.g., the global fluctuations in deoxyhemoglobin from rate and depth of breathing, heart rate, and end tidal pCO2 mentioned in the main text Box 1). Thus, it does not represent a solution to the global signal/global noise problem raised here. Additionally, multi-echo EPI is currently limited in the spatial and temporal resolution that can be achieved (on the order of 4 mm isotropic in published results with a TR of 2.5 s). With the incorporation of multi-band into a multi-echo sequence, the spatial-temporal resolution of this technique should be improved, enabling accurate cortical analysis (i.e., 2.5 mm isotropic or better). At that point, a thorough evaluation of the pros and cons of standard multi-band vs. combined multi-band and multi-echo should be done, in which there are considerations of the tradeoffs in spatial-temporal resolution and temporal SNR, relative to performance in time series de-noising. Currently, we recommend using one of the other ICA-based techniques with standard multi-band fMRI. It is possible that the temporal resolution recommendations of <1 s could be relaxed for multi-echo techniques, though the impact of fewer timepoints on multi-variate methods such as ICA or partial correlation would need to be investigated.

10. Selective Spatio-temporal Smoothing of Unstructured Noise

A major limitation of spatial and temporal smoothing (or filtering) is that they smooth signal and noise indiscriminately. When using smoothing to improve SNR, the data is also being blurred

Nature Neuroscience: doi:10.1038/nn.4361

across areal boundaries or removing high frequency temporal information that is known not to be fully noise[54,55]. For the purposes of estimating dense functional connectivity, an attractive alternative makes use of Principal Components Analysis (PCA) to represent an fMRI 4D timeseries as a series of weighted spatial eigenvectors, ordered according to the size of their eigenvalues. The eigenvalues of the PCA decomposition together form an eigenspectrum (a Wishart distribution in the null case[56]). A white noise timeseries will have a characteristic spectrum of eigenvalues, and structured signal (or noise) present in the data will increase the eigenvalues above and beyond this null eigenspectrum. Even in a dataset containing structured signal (e.g. area-specific resting-state fluctuations), the tail of the eigenspectrum will be largely unstructured noise. A Wishart function can be fit to this tail, making a prediction of the null eigenspectrum for the dataset. This null distribution can be subtracted from the data, forcing most of the eigenvalues representing noise to zero, and the PCA series can then be converted back to a timeseries that has higher SNR (for individual subjects), or correlated as a PCA series with higher SNR and fewer artifacts (for group data). Although reducing the small eigenvalues may remove some very weak signals, the technique avoids the blurring of strong signals that occurs with typical spatial and temporal smoothing. Additionally, this Wishart rolloff approach can be used to remove the effects of spatially autocorrelated noise and ringing from iteratively generated group PCA series[57] (http://www.humanconnectome.org/documentation/mound-and-moat-effect.html). Supplementary Fig. 7 illustrates both effects.

11. Improving Subcortical Alignment Using Diffusion Fiber Orientations

One arena in which the current HCP Pipelines could likely be improved is in the alignment of subcortical white matter, which is based only on T1w images that have relatively little contrast within white matter. Much as areal-feature-based registration improves the alignment of cortical areas over using only folding patterns, using fiber orientation information to drive white matter tract alignment will provide the registration algorithm more information about the identity of the fiber tracts being aligned. Because of their low white matter contrast, T1w images will only approximately align these tracts. Future versions of the HCP Pipelines may integrate such a diffusion fiber orientation-based nonlinear registration, as has previously been demonstrated (http://dti-tk.sourceforge.net/pmwiki/pmwiki.php?n=Main.HomePage)[58,59].

12. Effects of Thresholding on Reproducibility of Neuroimaging Results

Two interrelated issues arise from the traditional use of statistical thresholding for display of neuroimaging results and identification of regions (which are often then converted to 3D standard space coordinates as the sole means of data sharing):

(1) Statistical significance maps (e.g. Z or p maps) are affected by more than the measure of interest (e.g., %BOLD change or relative myelin content). They also reflect the estimate of the error of the measurement. While such maps are statistically interpretable as the likelihood that a finding differs from chance, and thus have a role for establishing statistically significant effects, they are less biologically interpretable than maps of the original effect of interest. For example, in data generated using modern multi-channel head coils, regions near the center of the brain have lower Z-statistic values owing to the higher amount of unstructured noise in those regions after correcting for the sensitivity profile of the head coil. This will decrease Z-statistics even if the %BOLD change in a subcortical central region is the same as in a cortical region. Thus, when searching for biological boundaries in the data, one should arguably use maps of the effect of interest, rather than Z-statistic maps. Further, biologically important boundaries in the data are likely to be identified most reliably by looking for large and rapid changes in the original effect of interest over a small spatial distance. Thus focusing on gradients (measures of spatial

Nature Neuroscience: doi:10.1038/nn.4361

change) in modalities can be an effective way to identify biologically important boundaries in one's data, as was done in the HCP's recent multi-modal parcellation[34]. Using such gradients may aid in comparing across studies to determine the overlap of the results (i.e., do two studies activate or connect the same brain area or not?).

(2) Traditionally, statistical significance maps are thresholded (usually at $p<0.05$ after some method of multiple comparison correction), and the regions that are identified are generally represented by the stereotaxic coordinates of the spatial center of gravity. Often the thresholded statistical maps are all that is provided in the published figures of a study, making it harder to identify the underlying patterns in the data (only the "peak" and "valley" blobs are identified). One problem that could be caused by this approach is missing evidence of an artifactual pattern that is more obvious in the unthresholded data (e.g. after thresholding one might find a few blobs on gyri and sulci, but the unthresholded data might clearly have a gyral vs sulcal bias). Another issue is that the location of the statistical threshold contour can vary much more across studies than the underlying maps of the effect of interest or their gradients (See Supplementary Fig. 8). Relatively uninteresting differences in Z statistical maps can easily turn activations whose significance is right around the threshold value from being above threshold to below threshold and vice-versa as illustrated in the figure. These differences include slight changes in the spatial distribution of the noise, slight errors in the measure of the effect of interest, or even something as simple as adding more subjects or data. When looking for cortical areal borders, the focus should be on where the effect of interest changes markedly across the cortical surface, not where the variance of non-interest changes. Also, in the limit, given enough data in an individual subject or enough subjects, nearly the entire brain may have statistically significant activation and deactivation in a task fMRI study using even a very conservative statistical threshold, but this says little about the biological significance of this whole brain activation and deactivation[60]. Thus, it is advantageous to use gradients in maps of the effect of interest to delineate areal borders or regions of interest, rather than statistical thresholds. Some of the spatial uncertainty present in the traditional neuroimaging analysis paradigm may reflect the use of statistical thresholding to define clusters from which 3D activation peaks, or centers of mass are computed (also, many of the widely used multiple comparison correction and statistical thresholding approaches also have very inflated false positive rates [61]). Thus, it would be helpful if the maps of the effect of interest, their gradients, and statistical significance maps were displayed in publications and uploaded to neuroimaging databases such as BALSA (see main text), OpenfMRI (https://openfmri.org), and NeuroVault (http://neurovault.org).

13. Group Average Registration Drift

Group average registration drift is a phenomenon that occurs when registration templates are being generated with non-rigid registration algorithms or when different registration algorithms or templates are used to drive alignment to a common standard space. A particularly striking example of drift is the ~37% increase in brain volume that occurs when one registers HCP study brains to the MNI152 template (see Supplementary Fig. 9), owing to a gradual expansion in the non-rigid components of the registration during the iterative process of template generation used to create the original MNI space[62,63]. Drift can be removed by computing the group average registration effect and concatenating its inverse onto each individual subject's registration. It is both desirable and feasible to avoid such drifts in future brain templates. If such drifts are eliminated[64], very precise cross-study, cross-modal boundary comparisons can be made. This is true even when using different areal features in different studies after a gentle folding-based registration to a common geographic template (see Supplementary Fig. 10), so long as the two studies' samples are drawn from the same population. Importantly, the current

Nature Neuroscience: doi:10.1038/nn.4361

grayordinates space is still based on MNI space for subcortical structures and thus still has this drift embedded in it. A future version of the CIFTI grayordinates space may eliminate this drift.

14. 'Weighted' Approaches to Parcellation

Parcellation of the brain is often carried out by clustering together neighboring voxels or vertices on the basis of the similarity of their timeseries. This typically yields a large number of non-overlapping parcels, with a single contiguous group of voxels in each parcel (or network node). This can be referred to as a "hard parcellation"[65,66]. However the clustering does not have to be "hard", but may instead be "fuzzy" or even probabilistic. One approach to generating a "fuzzy parcellation" involves high-dimensional spatial independent-component analysis (ICA)[67]. Using ICA, each network "node" is described by a spatial map of varying weights. Notably each map may overlap with other nodes' maps and may include more than one set of contiguously neighbouring voxels. Network edges (connections between nodes) are estimated by comparing the fMRI time series associated with the nodes (usually derived using multiple regression). With any of these parcellation approaches, brain connectivity can be represented as a 'parcellated connectome', which can be visualized simply as an N_{nodes} x N_{nodes} network matrix or a graph (explicitly showing nodes and the strongest edges[68]) or using more sophisticated visualization approaches that embed nodes and edges into spatial representations of the brain[69]. Hard parcellations have the advantage of being simpler representations and are therefore a more parsimonious yet accurate representation, if they are a good biological model of the data. They are also easier to interpret as distinct neuroanatomical entities. On the other hand, soft parcellations may be useful where the biology does not fit the model of non-overlapping spatial segregation or there is a premium on maximizing the homogeneity of signals. Multiple regression using a weighted parcellation may be able to produce "purer," more homogeneous timecourses when multiple neural signals are spatially superimposed (e.g. the distinct signals related to the eccentricity and polar angle axes of visuotopic organization in the visual cortex[34]). Further, an ICA-based weighted parcellation by definition will not allow rank-deficient node-timeseries matrices to occur, whereas this condition can be a problem for partial correlation of hard parcellations (i.e. especially if a parcellation includes homotopically related parcels across the two hemispheres).

15. Current Limitations to Generating Human Connectomes Using Non-invasive Approaches

The Human Connectome Project's core goal is to create area-to-area connectivity matrices using diffusion tractography and functional connectivity. While much progress has been made in this direction (in particular in defining the cortical areas to be connected[34]), much work remains to assess and maximize the accuracy of the non-invasive measures of brain connectivity by comparing them to invasive tracers in non-human primates.

Despite its well documented capabilities for localizing major fiber tracts in the white matter (see below), using diffusion tractography to map connections between distant gray matter regions has several limitations and is a field of active research[70-72]. Some of the limitations are fundamental, whereas others are amenable to methodological advances. One fundamental limitation of diffusion tractography is that it is unable to provide accurate directionality information about connections (it is impossible to know if an estimated connection is A → B, A ← B, or A ←→ B) because water diffusion in axons is bidirectional. In contrast, invasive retrograde and anterograde tracers reveal the directionality of connections[71,73]. Thus, directionality information would have to be inferred through some other means (indeed, this prospect was one reason that MEG was included as an imaging modality in the HCP). Another problem is the indirect nature of the measurements: a mapping is necessary from

Nature Neuroscience: doi:10.1038/nn.4361

measurements of diffusion displacements to axons and fiber patterns. Due to the limited resolution of MRI (~mm^3), hundreds of thousands of axons coexist in the same brain imaging voxel[74], giving rise to different complex geometries (crossing, "kissing", branching, fanning, converging). Distinguishing these configurations from the measured diffusion signal is an ill-posed problem in general. For instance some axons in white matter branch by bifurcation at quasi-right angles (e.g.,[75]), making it difficult for tractography to distinguish axonal branching from crossing of independent fiber bundles. These issues can cause errors when tracking either deep in the white matter or close to the cortex.

Current tractography paradigms are also challenged in reaching terminations within gray matter and in accurately tracking in the vicinity of the WM/GM boundary[74,76,77]. There is a pronounced bias in tractography streamlines towards being higher on gyral crowns and much lower in sulci. When streamline density is normalized relative to flat regions (zero surface curvature in sulcal banks), it becomes clear that the bias is actually more of an anti-sulcal bias rather than as strong an increase in gyral connectivity. Such biases are strongest when tractography streamlines return to the cortex from the deep white matter, and so they are most prevalent when seeding bidirectionally within the white matter. Pushing to higher resolution does help (as shown when fusing 3T and 7T HCP dMRI data[78]), but does not fully address this issue (as shown with 250um isotropic resolution in post-mortem macaques[77]). The problems arise because (i) near sulcal fundi, superficial white matter and deep grey matter are dominated by tangentially running "U-Fibers," which direct streamlines around the sulci, rather than allowing them to penetrate the grey matter; and (ii) in 'gyral blades' (white matter lying between sulcal banks), fiber bundles are predominantly oriented towards gyral crowns. It is possible to predict components of a neuroanatomically reasonable gyral bias in the tractography using simple geometric models[74,76,77] that assume that each cubic mm (or square mm of the cortical midthickness surface) should, to first order, send out and receive a comparable number of axons. Thus gyral crowns should have somewhat more axons crossing their white matter surfaces and sulcal fundi somewhat less; however, the bias observed in tractography is frequently much higher[74,76]. This geometric model can serve as a reference upon which tractographic model improvements can be objectively evaluated.

Concerns about the performance of tractography near the grey-white interface and the influence of folding patterns on the locations of gradients in tractographic connectivity led to tractography not being used as a modality in the HCP's multi-modal parcellation of the cerebral cortex[34]. That said, it will be interesting to compare tractographic connectivity gradients to the multi-modal consensus areal borders once large amounts of HCP tractographic data have been generated and shared. Should tractography produce gradients that overlap substantially with the multi-modally defined cortical areal borders, tractography features could aid in predicting the locations of cortical areas in individual subjects using the multi-modal areal classifier[34]. If tractography gradients do not have strong overlap with the multi-modal consensus areal borders, future improvements in tractographic models could be objectively evaluated with respect to these reference areal borders to aim for better agreement.

Collectively, these issues give rise to significant challenges in keeping track of area-to-area connections as they travel from cortex through the deep white matter within fascicles and then fan out again to reach distant regions of cortex. To the extent that axons do not remain clustered together relative to the voxel size of the diffusion scan, these fascicles—which themselves have complex cross-sectional shapes—form bottlenecks where the accuracy of information about the strengths of inter-areal connections will be eroded. If fibers from one area tend to cluster together within the deep white matter and then fan out again when they reach cortex, this will relax the spatial resolution requirements needed to accurately reconstruct area-

Nature Neuroscience: doi:10.1038/nn.4361

to-area pathways. On the other hand, if axons relatively randomly distribute themselves within fascicles, accurate reconstructions of area-to-area pathways would require dramatic improvements over current spatial resolution. Invasive tracer studies in non-human primates compared with tractography may help to resolve which reality predominates within the brain; though the plethora of studies that show agreement between tractography results and other modalities and in various contexts (see Ref.[71] for a review of such evidence) suggest that the geometrical realities may be somewhat on the favorable side.

The above limitations in estimating direct connection strengths using tractography are difficult to fully address. However, better data and better methods can help to reduce errors. Work within the HCP consortium has aimed towards these goals[18,19]. The HCP pushed the spatial resolution of dMRI to reduce partial volume effects, while keeping high SNR and high angular resolution. At the same time, data at multiple angular contrasts (b values) have been sampled, which can further help in accurately mapping of fiber orientations[79,80]. Up to three fiber crossings within the deep white matter, as well as various microstructure metrics,[81-84] can be routinely modeled using the current HCP data releases. New paradigms for mapping connections are also being explored, including incorporation of fiber dispersion[85] in tractography and neighbourhood-wise tracking[86]. Better, neuroanatomically informed models of axonal trajectories near the cortical ribbon incorporated into tractography may help to reduce gyral and sulcal biases. In Ref, [87], the relationship of fiber orientations with cortical features is explored using post-mortem high-resolution histology and used to inform generative models for in-vivo diffusion MRI in an attempt to use mesoscale information to better estimate the paths of tractography at the transition between white and gray matter. As improved tractography models aimed at measuring accurate area-to-area connection strengths emerge, results can be compared directly with invasive tracer studies in the macaque to see whether modifications to tractography models improve the agreement between tracers and tractography[76].

Aside from the question of accurate area-to-area connections, tractography can still provide useful information about the major fascicles themselves, for understanding how large territories of cortex are connected (e.g., refs. [88,89], for understanding how large fascicles have changed through evolution[90,91], for understanding differences in fascicles in health and disease, or for the purposes of surgical planning. Thus, even without yet having a fully accurate and validated area-to-area connectome, investigators are able to see how large ensembles of brain regions are connected to one another and to investigate white matter microstructural changes along these fascicles.

A different set of limitations affect functional connectivity measures. Obtaining connectivity estimates using statistical dependences of the time-varying BOLD signal is also very indirect and has its own caveats. Functional connectivity can be strong between parts of cortex that are known not to be directly connected (e.g., the horizontal meridian representations in V1 across hemispheres[92]). Thus, functional connectivity reflects both direct and indirect connections, likely including common inputs. One approach that attempts to make functional connectivity more direct is partial correlation functional connectivity. This multi-variate approach attempts to identify the unique pairwise influences between cortical areas. Such analyses are very demanding in terms of the number of timepoints and data quality required to produce meaningful data, and may not behave optimally in the context of neuroanatomical parcellations (as opposed to the ICA-based parcellations mentioned in Topic #14). Estimating reliable directionality information from rfMRI timeseries is also problematic, partly because of the temporal smoothing of the BOLD response function, although causal inference methods not based on temporal lag are being developed and evaluated for rfMRI data (e.g., see references and discussions in Ref. [93]). Another conceptual limitation, raised in Ref. [71] involves an area that

Nature Neuroscience: doi:10.1038/nn.4361

participates in many resting state networks and as a result integrates many competing signals. Such an area might actually appear less connected than other areas, and a graph theoretic analysis searching for hubs might overlook it[94].

Significant progress had been made, in part by the HCP consortium, on how to correct for the temporal fluctuations induced by subject movement (main text and Topic #9). That said, as discussed in Box 1 and Topic #9, standard full correlation-based functional connectivity has a global signal/global noise problem (as opposed to partial correlation, which is largely immune) that has been much debated in the literature. This problem will need to be solved before full correlation functional connectivity provides an unbiased measure for assessing differences between groups with differing amounts of global noise (fortunately, global noise, by virtue of its global nature, does not affect the locations of gradients). All these functional connectivity approaches (full correlation, partial correlation, attempts at directional correlation) would benefit from comparison with invasive tracer data from non-human primates, just as tractography does, for validation and to improve their accuracy at representing brain connectivity. Such a framework, just like for tractography, could aid in making decisions about processing and modeling strategies for the generation of more accurate non-invasive human connectomes.

16. The 'Standard' Individual Lab's Study in the HCP-Style Paradigm

The HCP-Style neuroimaging paradigm includes a set of standard, concrete recommendations that are applicable to many if not most human neuroimaging studies (i.e., what is desirable for investigators to do in the absence of a good reason to do something different):

Acquisition: A standard study should acquire 3D T1w and T2w structural images on a 3T scanner with a multi-channel head coil (ideally 32 or more receive channels) with 0.8 mm isotropic resolution or better (half the minimum thickness of the cerebral cortex), which can be acquired in ~13 minutes. (3T is currently the most mature platform that produces higher SNR than 1.5T, and Siemens Prisma scanners are farthest along with the necessary technology and stability). Functional MRI data should be acquired using task and/or resting state paradigms according to the purpose of the study. The functional MRI data are acquired whole-brain with isotropic spatial resolution of 2.5 mm or better (less than the 2.6mm mean thickness of the cerebral cortex) and a TR of 1 second or less. As much fMRI data is acquired as is feasible (ideally 1 hour or more, and at least 30 min is strongly recommended[32]). A phase-reversed pair of spin echo images is acquired with the same geometry as the fMRI data in about 1 minute (for mapping the b0 distortion field and the fMRI receive intensity bias field). If diffusion imaging is a part of the study, it should be acquired using a monopolar diffusion encoding scheme with whole brain coverage and with as high spatial resolution (i.e. less than 2 mm isotropic), as many directions (more than 60 for high angular resolution), and as high a b-value (ideally more than b=1000 s/mm^2, and multiple shells are also generally helpful), as is feasible for the scanner that is used. At a minimum, phase reversed b0s should be acquired and as many unique directions as possible (rather than repeating directions for averaging). Ideally phase-reversed averages are acquired if there is time for hundreds of directions to be acquired as in the main HCP. Both the fMRI and diffusion MRI data must be acquired using a multi-band sequence to maximize efficiency and to reach the targets for spatial resolution, temporal resolution, and angular resolution.

Preprocessing: The HCP's minimal preprocessing pipelines, or alternatives with demonstrably superior capabilities, should be used to remove all image distortions, remove the spatial effects of subject motion, align data across modalities, and generate individual subject cortical surfaces, subcortical segmentations, and myelin and thickness maps. Structured temporal

Nature Neuroscience: doi:10.1038/nn.4361

noise is removed from fMRI data using ICA+FIX or a similar approach and it is removed from DWI data using the outlier modeling available in FSL's 'eddy' or a similar approach. Data are registered on the surface using MSMAll or a similar areal-feature-based approach to maximize alignment of areas across subjects and, if the study is creating its own areal-feature-based registration template, group registration drift relative to a gentle folding-based geographic registration is removed. Volume data are registered using a nonlinear volume registration approach (perhaps in the future based on fiber orientation information if diffusion data has been acquired, but currently based on the T1w image). CIFTI files are created as the final outputs of these registrations in the 2mm standard grayordinates space. Spatial and temporal smoothing are avoided as a rule (unless specific study goals require it, e.g. matching the spatial resolution between fMRI and a grid of electrodes in ECOG), and unstructured noise is reduced using parcellation or an approach like the Wishart rolloff method in Topic #10. Parcellation is done using a neuroanatomical parcellation that reflects the structural and functional organization of the brain (e.g. the HCP's multi-modal parcellation) that was generated based on accurately aligned data of hundreds of subjects. If greater accuracy is desired, individual subject parcellations can also be created using the previously trained cortical areal classifier (so far this has been tested on 1 hour of resting state fMRI data, but in principle other kinds of fMRI data should also work, though the minimum amount of data needed has not yet been established).

Analysis: Analyses of fine spatial patterns are carried out on the dense (grayordinate-wise) CIFTI files, however statistical analyses of whole brain, area-level patterns are carried out on parcellated CIFTI files to benefit from gains in sensitivity and power. Statistical comparisons across the whole brain are thus carried out on parcellated data, whereas finer comparisons might be carried out only within predefined brain areas. Obviously there are a very large number of analyses that could arise from HCP-Style studies, including those that wish to go beyond areal-feature-based registration to establish non-spatial correspondences across subjects such as those in hyperalignment[95]. The critical features of an HCP-Style analysis relate to the neuroanatomical fidelity of the spatial localization at the areal level and the use of neuroanatomy to describe the results.

Data Sharing: Results are visualized on cortical surfaces for the cerebral cortex and in volume slices for subcortical nuclei. If thresholded statistical maps are generated, unthresholded versions of the effect of interest are also made available. When appropriate, figures of brain images are created directly in Connectome Workbench or another neuroimaging software with equivalent features for figure creation and saved as "scenes" which link the data to the figure along with any annotations of the figure. These scenes are linked within the resulting publication using unique identifying URLs and the scenes are uploaded to a neuroimaging database so that they and their data are available for download by other investigators. This will allow users access to the brain maps of published figures to facilitate direct cross-study comparisons. Any customized analysis code is uploaded to a code sharing service (e.g., GitHub) and also shared with readers.

17. Can't I Just Map between the Group Average Surface and Group Average Volume MNI Space?

A frequent question on the HCP-Users mailing list concerns how to compare group average or atlas data in MNI space to data within the standard CIFTI grayordinates space (91282 grayordinates, 2mm spacing). Although it is possible to compare group average or atlas data in the standard grayordinate space to an individual subject's physical volume space (by using the MSMAll registered individual surfaces to form a cortical ribbon in the subject's physical volume space and projecting the data into this ribbon), there exists no method to bring data from a

Nature Neuroscience: doi:10.1038/nn.4361

group average surface into group average MNI volume space that maintains the spatial localization accuracy of the HCP-style neuroimaging paradigm. Insofar as a particular cortical region has a consistent folding pattern across subjects and these folds have a consistent relationship with cortical areas, the decrease in accuracy is smaller, but this condition holds for only a minority of cortical regions (e.g., the central sulcus, the calcarine sulcus, and the insula). Even in these regions, there is a substantial amount of degradation in spatial localization precision in the volume (e.g., V1 in Fischl et al[96]), and inaccuracy in mapping to a group average cortical surface (e.g., area 2 in Van Essen et al[97]). This makes neuroanatomically valid comparisons between group average data in MNI volume space and group average data on the surface inherently challenging[7].

The problem manifests itself most strikingly when comparing data mapped to group average cortical surfaces, where matching between vertices is derived from areal feature-based registration, to data mapped to average structural images registered in the volume using nonlinear volume registration (in this case FNIRT, as in Glasser et al 2013[17]). Supplementary Fig. 12 shows the issue with a comparison between group average and individual data. In most cortical regions, group average surfaces and curvature maps are much blurrier than individual subject surfaces and curvature maps. The only exceptions are regions where there is strong consistency in folding patterns across subjects and between cortical areas and folds[96], such as the central sulcus, the calcarine sulcus, and the insula. This has an important impact on the group average surface contour, which shrinks from geometric effects of the averaging. This shrinkage can be approximately corrected for on the surface when performing computations on group average surfaces (e.g., computing gradients or resampling data) by also making use of the group average of the vertex areas across the individual subjects' midthickness surfaces in their native physical volume spaces. This shrinkage makes mapping to or from group average surfaces into group average volume space (e.g., MNI standard space) highly problematic, as the surfaces will shrink and miss large portions of grey matter. To compare group average data from two studies while maintaining the neuroanatomical fidelity of the HCP-style paradigm, it is necessary to process both studies using the HCP-style neuroimaging analysis approach. This means cortical atlases should be surface-based and any averaging across subjects should occur on the surface after areal features have been aligned. Another approach, generating a surface from a group average volume template and registering it to a surface template using folding patterns, may avoid the shrinkage issue[98,99]; however, surfaces previously produced using this approach are of poorer quality than individual subject surfaces owing to the blurring of partial voluming effects on the volume template from inadequate cross-subject cortical alignment. There is also no guarantee that cortical areas will be aligned across templates, and the volume-based cortical data of interest will still be poorly aligned.

18. How can I compare surface-based data generated on other platforms to the HCP_MMP1.0 parcellation and other group average surface data?

Many investigators have carried out surface-based analyses using atlases and platforms other than fs_LR[97] and Connectome Workbench, but may be interested in comparing their results to HCP group average data, particularly the HCP_MMP1.0 cortical parcellation. This requires a mapping between the surfaces (individual or group average) on which the investigator's analysis has been carried out and the HCP group average surface mesh. The HCP_MMP1.0 parcellation and most other HCP group average datasets are represented on the '32k_fs_LR' mesh (32,492 vertices per hemisphere, with a 2 mm average vertex spacing on the midthickness surface). This mesh has good geographic correspondence between the left and right hemisphere surfaces as a result of landmark-constrained interhemispheric registration between the left and right hemispheres (fs_L and fs_R) of FreeSurfer's fsaverage atlas

Nature Neuroscience: doi:10.1038/nn.4361

surface[97]. For datasets analyzed in FreeSurfer, data from the fsaverage atlas surfaces can be mapped to the 32k_fs_LR mesh by resampling from the registered fsaverage atlas spherical surface to the 32k_fs_LR atlas surface (see https://wiki.humanconnectome.org/display/PublicData/HCP+Users+FAQ; entry #9 – "*How do I map data between FreeSurfer (native or fsaverage mesh) and HCP's fs_LR mesh?"*)[97]. Conversely, data on the 32k_fs_LR mesh can be mapped onto FreeSurfer surfaces by reversing the order of files called in the resampling command. In general, we encourage investigators to map data onto the 32k_fs_LR mesh (the correspondence between left and right hemispheres is useful in a variety of contexts), but recognize there may be project-specific reasons for mapping in the reverse direction. For investigators using other surface-based analysis platforms (e.g., SUMA, BrainVoyager, or BrainVisa), analogous inter-atlas transformations have yet to be generated but such methods will be posted to the HCP-users FAQ if and when they are implemented.

Nature Neuroscience: doi:10.1038/nn.4361

III. Supplementary References

1 Mugler, J. P., 3rd & Brookeman, J. R. Three-dimensional magnetization-prepared rapid gradient-echo imaging (3D MP RAGE). *Magnetic resonance in medicine* **15**, 152-157 (1990).

2 Mugler, J. P., 3rd *et al.* Optimized single-slab three-dimensional spin-echo MR imaging of the brain. *Radiology* **216**, 891-899, doi:10.1148/radiology.216.3.r00au46891 (2000).

3 Van de Moortele, P. F. *et al.* T1 weighted brain images at 7 Tesla unbiased for Proton Density, T2* contrast and RF coil receive B1 sensitivity with simultaneous vessel visualization. *NeuroImage* **46**, 432-446, doi:10.1016/j.neuroimage.2009.02.009 (2009).

4 Mugler, J. P., 3rd. Optimized three-dimensional fast-spin-echo MRI. *Journal of magnetic resonance imaging : JMRI* **39**, 745-767, doi:10.1002/jmri.24542 (2014).

5 Busse, R. F., Hariharan, H., Vu, A. & Brittain, J. H. Fast spin echo sequences with very long echo trains: design of variable refocusing flip angle schedules and generation of clinical T2 contrast. *Magnetic resonance in medicine* **55**, 1030-1037, doi:10.1002/mrm.20863 (2006).

6 Park, J., Mugler, J. P., 3rd, Horger, W. & Kiefer, B. Optimized T1-weighted contrast for single-slab 3D turbo spin-echo imaging with long echo trains: application to whole-brain imaging. *Magnetic resonance in medicine* **58**, 982-992, doi:10.1002/mrm.21386 (2007).

7 Glasser, M. F. & Van Essen, D. C. Mapping human cortical areas in vivo based on myelin content as revealed by T1- and T2-weighted MRI. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **31**, 11597-11616, doi:10.1523/JNEUROSCI.2180-11.2011 (2011).

8 Stuber, C. *et al.* Myelin and iron concentration in the human brain: a quantitative study of MRI contrast. *NeuroImage* **93 Pt 1**, 95-106, doi:10.1016/j.neuroimage.2014.02.026 (2014).

9 Fukunaga, M. *et al.* Layer-specific variation of iron content in cerebral cortex as a source of MRI contrast. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 3834-3839, doi:10.1073/pnas.0911177107 (2010).

10 Weiskopf, N. *et al.* Quantitative multi-parameter mapping of R1, PD(*), MT, and R2(*) at 3T: a multi-center validation. *Frontiers in neuroscience* **7**, 95, doi:10.3389/fnins.2013.00095 (2013).

11 Dick, F. *et al.* In vivo functional and myeloarchitectonic mapping of human primary auditory areas. *The Journal of Neuroscience* **32**, 16095-16105 (2012).

12 Geyer, S., Weiss, M., Reimann, K., Lohmann, G. & Turner, R. Microstructural parcellation of the human cerebral cortex–from Brodmann's post-mortem map to in vivo mapping with high-field magnetic resonance imaging. *Frontiers in human neuroscience* **5**, 19 (2011).

13 Lutti, A., Dick, F., Sereno, M. I. & Weiskopf, N. Using high-resolution quantitative mapping of R1 as an index of cortical myelination. *NeuroImage* **93**, 176-188 (2014).

14 Sereno, M. I., Lutti, A., Weiskopf, N. & Dick, F. Mapping the human cortical surface by combining quantitative T1 with retinotopy. *Cerebral Cortex* **23**, 2261-2268 (2013).

15 Cohen-Adad, J. *et al.* T 2* mapping and B 0 orientation-dependence at 7T reveal cyto-and myeloarchitecture organization of the human cortex. *NeuroImage* **60**, 1006-1014 (2012).
16 De Martino, F. *et al.* High-resolution mapping of myeloarchitecture in vivo: localization of auditory areas in the human brain. *Cerebral Cortex* **25**, 3394-3405 (2015).
17 Glasser, M. F. *et al.* The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage* **80**, 105-124, doi:10.1016/j.neuroimage.2013.04.127 (2013).
18 Sotiropoulos, S. N. *et al.* Advances in diffusion MRI acquisition and processing in the Human Connectome Project. *NeuroImage* **80**, 125-143, doi:10.1016/j.neuroimage.2013.05.057 (2013).
19 Vu, A. T. *et al.* High resolution whole brain diffusion imaging at 7T for the Human Connectome Project. *NeuroImage* **122**, 318-331, doi:10.1016/j.neuroimage.2015.08.004 (2015).
20 Setsompop, K. *et al.* Pushing the limits of in vivo diffusion MRI for the Human Connectome Project. *NeuroImage* **80**, 220-233, doi:10.1016/j.neuroimage.2013.05.078 (2013).
21 Stejskal, E. O. & Tanner, J. E. Diffusion measurements: Spin echoes in the presence of a time-dependent field gradient. *J. Chem. Phys.* **42**, 288 (1965).
22 Reese, T. G., Heid, O., Weisskoff, R. M. & Wedeen, V. J. Reduction of eddy-current-induced distortion in diffusion MRI using a twice-refocused spin echo. *Magnetic resonance in medicine* **49**, 177-182, doi:10.1002/mrm.10308 (2003).
23 Andersson, J. L. & Sotiropoulos, S. N. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *NeuroImage* **125**, 1063-1078, doi:10.1016/j.neuroimage.2015.10.019 (2016).
24 Andersson, J. L. & Skare, S. A model-based method for retrospective correction of geometric distortions in diffusion-weighted EPI. *NeuroImage* **16**, 177-199, doi:10.1006/nimg.2001.1039 (2002).
25 Moeller, S. *et al.* Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI. *Magnetic resonance in medicine* **63**, 1144-1153, doi:10.1002/mrm.22361 (2010).
26 Setsompop, K. *et al.* Improving diffusion MRI using simultaneous multi-slice echo planar imaging. *NeuroImage* **63**, 569-580, doi:10.1016/j.neuroimage.2012.06.033 (2012).
27 Smith, S. M. *et al.* Resting-state fMRI in the Human Connectome Project. *NeuroImage* **80**, 144-168, doi:10.1016/j.neuroimage.2013.05.039 (2013).
28 Auerbach, E. J., Xu, J., Yacoub, E., Moeller, S. & Ugurbil, K. Multiband accelerated spin-echo echo planar imaging with reduced peak RF power using time-shifted RF pulses. *Magnetic resonance in medicine* **69**, 1261-1267, doi:10.1002/mrm.24719 (2013).
29 Ugurbil, K. *et al.* Pushing spatial and temporal resolution for functional and diffusion MRI in the Human Connectome Project. *NeuroImage* **80**, 80-104, doi:10.1016/j.neuroimage.2013.05.012 (2013).
30 Tisdall, M. D. *et al.* Volumetric navigators for prospective motion correction and selective reacquisition in neuroanatomical MRI. *Magnetic resonance in medicine* **68**, 389-399, doi:10.1002/mrm.23228 (2012).

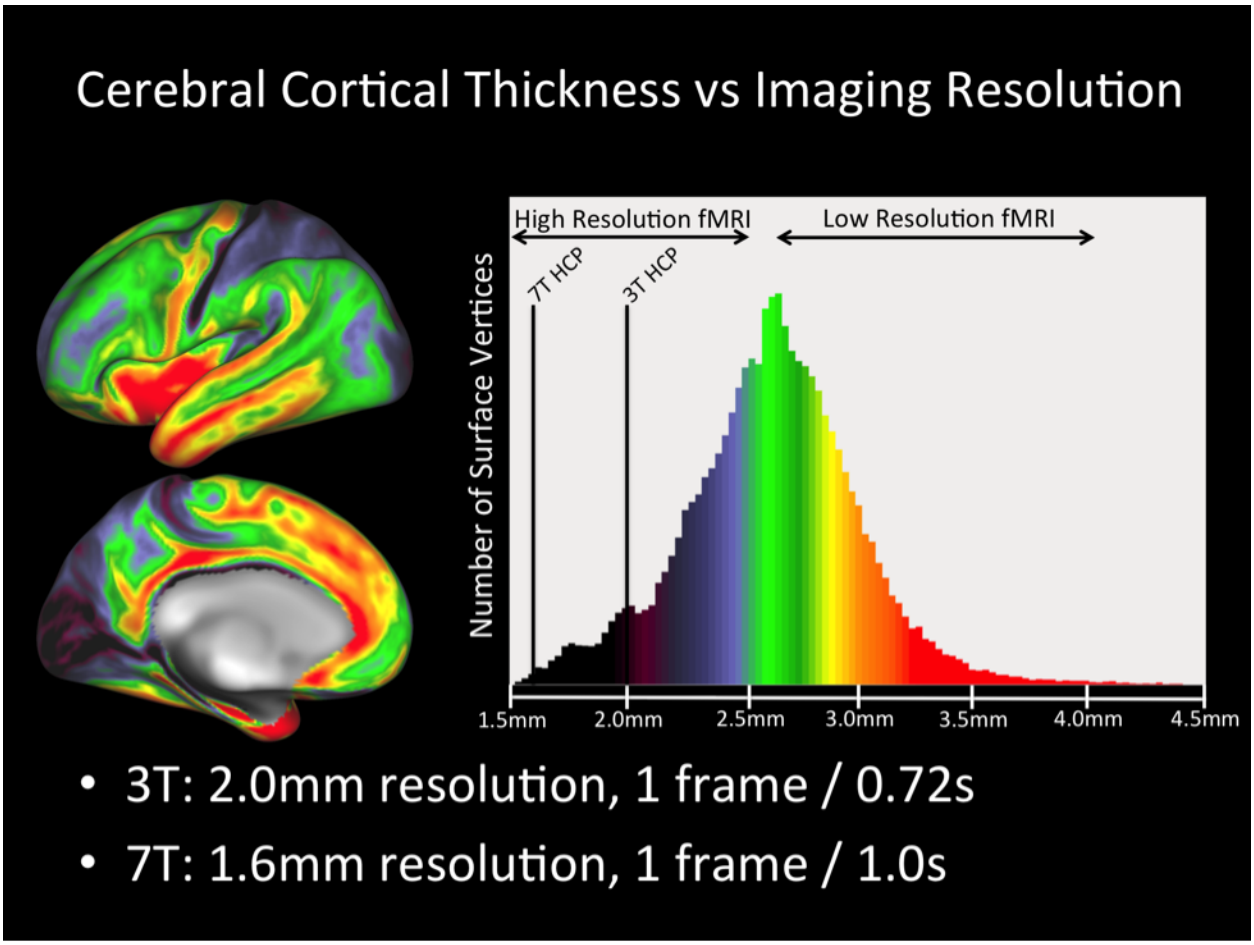
31 Tisdall, M. D. *et al.* Prospective motion correction with volumetric navigators (vNavs) reduces the bias and variance in brain morphometry induced by subject motion. *NeuroImage* **127**, 11-22, doi:10.1016/j.neuroimage.2015.11.054 (2016).
32 Laumann, T. O. *et al.* Functional System and Areal Organization of a Highly Sampled Individual Human Brain. *Neuron* **87**, 657-670, doi:10.1016/j.neuron.2015.06.037 (2015).
33 Cole, M. W., Bassett, D. S., Power, J. D., Braver, T. S. & Petersen, S. E. Intrinsic and task-evoked network architectures of the human brain. *Neuron* **83**, 238-251, doi:10.1016/j.neuron.2014.05.014 (2014).
34 Glasser, M. F. *et al.* A multi-modal parcellation of human cerebral cortex. *Nature*, doi:doi 10.1038/nature18933 (2016).
35 Wu, X. *et al.* Simultaneous multislice multiband parallel radiofrequency excitation with independent slice-specific transmit B1 homogenization. *Magnetic resonance in medicine* **70**, 630-638, doi:10.1002/mrm.24828 (2013).
36 Wu, X., Schmitter, S., Ugurbil, K. & Van de Moortele, P.-F. Slab-wise parallel transmit multiband RF pulse design for simultaneous multislice imaging with volumetric coverage. *Proc Int Soc Mag Reson Med* **22**, 4333 (2014).
37 Jann, K. *et al.* Functional connectivity in BOLD and CBF data: similarity and reliability of resting brain networks. *NeuroImage* **106**, 111-122, doi:10.1016/j.neuroimage.2014.11.028 (2015).
38 Li, X. *et al.* Theoretical and experimental evaluation of multi-band EPI for high-resolution whole brain pCASL Imaging. *NeuroImage* **106**, 170-181, doi:10.1016/j.neuroimage.2014.10.029 (2015).
39 White, N. *et al.* PROMO: Real-time prospective motion correction in MRI using image-based tracking. *Magnetic resonance in medicine* **63**, 91-105, doi:10.1002/mrm.22176 (2010).
40 Bammer, R. *et al.* Analysis and generalized correction of the effect of spatial gradient field distortions in diffusion-weighted imaging. *Magnetic resonance in medicine* **50**, 560-569, doi:10.1002/mrm.10545 (2003).
41 Greve, D. N. & Fischl, B. Accurate and robust brain image alignment using boundary-based registration. *NeuroImage* **48**, 63-72, doi:10.1016/j.neuroimage.2009.06.060 (2009).
42 Andersson, J. L. & Sotiropoulos, S. N. Non-parametric representation and prediction of single- and multi-shell diffusion-weighted MRI data using Gaussian processes. *NeuroImage* **122**, 166-176, doi:10.1016/j.neuroimage.2015.07.067 (2015).
43 Andersson, J. L. R., Graham, M. S., Zsoldos, E. & Sotiropoulos, S. N. Incorporating outlier detection and replacement into a non-parameteric framework for movement and distortion correction of diffusion MR images. *Neuroimage (in review)* (2016).
44 Power, J. D. *et al.* Methods to detect, characterize, and remove motion artifact in resting state fMRI. *NeuroImage* **84**, 320-341, doi:10.1016/j.neuroimage.2013.08.048 (2014).
45 Power, J. D., Schlaggar, B. L. & Petersen, S. E. Recent progress and outstanding issues in motion correction in resting state fMRI. *NeuroImage* **105**, 536-551, doi:10.1016/j.neuroimage.2014.10.044 (2015).

46 Satterthwaite, T. D. *et al.* Impact of in-scanner head motion on multiple measures of functional connectivity: relevance for studies of neurodevelopment in youth. *NeuroImage* **60**, 623-632, doi:10.1016/j.neuroimage.2011.12.063 (2012).
47 Brooks, J. C. *et al.* Physiological noise modelling for spinal functional magnetic resonance imaging studies. *NeuroImage* **39**, 680-692, doi:10.1016/j.neuroimage.2007.09.018 (2008).
48 Salimi-Khorshidi, G. *et al.* Automatic denoising of functional MRI data: combining independent component analysis and hierarchical fusion of classifiers. *NeuroImage* **90**, 449-468, doi:10.1016/j.neuroimage.2013.11.046 (2014).
49 Griffanti, L. *et al.* ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. *NeuroImage* **95**, 232-247, doi:10.1016/j.neuroimage.2014.03.034 (2014).
50 Pruim, R. H., Mennes, M., Buitelaar, J. K. & Beckmann, C. F. Evaluation of ICA-AROMA and alternative strategies for motion artifact removal in resting state fMRI. *NeuroImage* **112**, 278-287, doi:10.1016/j.neuroimage.2015.02.063 (2015).
51 Pruim, R. H. *et al.* ICA-AROMA: A robust ICA-based strategy for removing motion artifacts from fMRI data. *NeuroImage* **112**, 267-277, doi:10.1016/j.neuroimage.2015.02.064 (2015).
52 Kundu, P. *et al.* Integrated strategy for improving functional connectivity mapping using multiecho fMRI. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 16187-16192, doi:10.1073/pnas.1301725110 (2013).
53 Kundu, P., Inati, S. J., Evans, J. W., Luh, W. M. & Bandettini, P. A. Differentiating BOLD and non-BOLD signals in fMRI time series using multi-echo EPI. *NeuroImage* **60**, 1759-1770, doi:10.1016/j.neuroimage.2011.12.028 (2012).
54 Chen, J. E. & Glover, G. H. BOLD fractional contribution to resting-state functional connectivity above 0.1 Hz. *NeuroImage* **107**, 207-218, doi:10.1016/j.neuroimage.2014.12.012 (2015).
55 Niazy, R. K., Xie, J., Miller, K., Beckmann, C. F. & Smith, S. M. Spectral characteristics of resting state networks. *Progress in brain research* **193**, 259-276, doi:10.1016/B978-0-444-53839-0.00017-X (2011).
56 Wishart, J. The generalizeed product moment distribution in samples from a normal multivariate population. *Biometrika* **20A**, 32-52 (1928).
57 Smith, S. M., Hyvarinen, A., Varoquaux, G., Miller, K. L. & Beckmann, C. F. Group-PCA for very large fMRI datasets. *NeuroImage* **101**, 738-749, doi:10.1016/j.neuroimage.2014.07.051 (2014).
58 Zhang, H. *et al.* High-dimensional spatial normalization of diffusion tensor images improves the detection of white matter differences: an example study using amyotrophic lateral sclerosis. *IEEE transactions on medical imaging* **26**, 1585-1597, doi:10.1109/TMI.2007.906784 (2007).
59 Zhang, H., Yushkevich, P. A., Alexander, D. C. & Gee, J. C. Deformable registration of diffusion tensor MR images with explicit orientation optimization. *Medical image analysis* **10**, 764-785, doi:10.1016/j.media.2006.06.004 (2006).
60 Gonzalez-Castillo, J. *et al.* Whole-brain, time-locked activation with simple tasks revealed using massive averaging and model-free analysis. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 5487-5492, doi:10.1073/pnas.1121049109 (2012).

61 Eklund, A., Nichols, T. E. & Knutsson, H. Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *Proceedings of the National Academy of Sciences*, 201602413 (2016).
62 Evans, A. C. *et al.* 3D statistical neuroanatomical models from 305 MRI volumes. *Proc IEEE-Nuclear Science Symp and Med Imag Conf* (1993).
63 Mazziotta, J. *et al.* A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **356**, 1293-1322, doi:10.1098/rstb.2001.0915 (2001).
64 Abdollahi, R. O. *et al.* Correspondences between retinotopic areas and myelin maps in human visual cortex. *NeuroImage* **99**, 509-524, doi:10.1016/j.neuroimage.2014.06.042 (2014).
65 Blumensath, T. *et al.* Spatially constrained hierarchical parcellation of the brain with resting-state fMRI. *NeuroImage* **76**, 313-324, doi:10.1016/j.neuroimage.2013.03.024 (2013).
66 de Reus, M. A. & van den Heuvel, M. P. The parcellation-based connectome: limitations and extensions. *NeuroImage* **80**, 397-404, doi:10.1016/j.neuroimage.2013.03.053 (2013).
67 Beckmann, C. F. Modelling with independent components. *NeuroImage* **62**, 891-901, doi:10.1016/j.neuroimage.2012.02.020 (2012).
68 Smith, S. M. *et al.* Functional connectomics from resting-state fMRI. *Trends in cognitive sciences* **17**, 666-682 (2013).
69 Margulies, D. S., Böttger, J., Watanabe, A. & Gorgolewski, K. J. Visualizing the human connectome. *NeuroImage* **80**, 445-461 (2013).
70 Jbabdi, S. & Johansen-Berg, H. Tractography: where do we go from here? *Brain connectivity* **1**, 169-183, doi:10.1089/brain.2011.0033 (2011).
71 Jbabdi, S., Sotiropoulos, S. N., Haber, S. N., Van Essen, D. C. & Behrens, T. E. Measuring macroscopic brain connections in vivo. *Nature neuroscience* **18**, 1546-1555, doi:10.1038/nn.4134 (2015).
72 Tournier, J. D., Mori, S. & Leemans, A. Diffusion tensor imaging and beyond. *Magnetic resonance in medicine* **65**, 1532-1556, doi:10.1002/mrm.22924 (2011).
73 Markov, N. T. *et al.* A weighted and directed interareal connectivity matrix for macaque cerebral cortex. *Cereb Cortex* **24**, 17-36, doi:10.1093/cercor/bhs270 (2014).
74 Van Essen, D. *et al.* Mapping connections in humans and non-human primates: Aspirations and challenges for diffusion. *Diffusion MRI, 2nd Ed, pp. 337-358, New York; Elsevier* (2014).
75 Economo, M. N. *et al.* A platform for brain-wide imaging and reconstruction of individual neurons. *eLife* **5**, e10566, doi:10.7554/eLife.10566 (2016).
76 Donahue, C. *et al.* Using diffusion tractography to predict cortical connection strength and distance: A quantitative comparison with tracers in the monkey. *Neuroimage. 2016. pii: S1053-8119(16)30035-0. doi: 10.1016/j.neuroimage.2016.04.002. [Epub ahead of print]* (2016).
77 Reveley, C. *et al.* Superficial white matter fiber systems impede detection of long-range cortical connections in diffusion MR tractography. *Proceedings of the National*

Academy of Sciences of the United States of America **112**, E2820-2828, doi:10.1073/pnas.1418198112 (2015).
78 Sotiropoulos, S. N. *et al.* Fusion in diffusion MRI for improved fibre orientation estimation: An application to the 3T and 7T data of the human connectome project. *NeuroImage* **134**, 396-409, doi:10.1016/j.neuroimage.2016.04.014 (2016).
79 Jbabdi, S., Sotiropoulos, S. N., Savio, A. M., Grana, M. & Behrens, T. E. Model-based analysis of multishell diffusion MR data for tractography: how to get over fitting problems. *Magnetic resonance in medicine* **68**, 1846-1855, doi:10.1002/mrm.24204 (2012).
80 Jeurissen, B., Tournier, J. D., Dhollander, T., Connelly, A. & Sijbers, J. Multi-tissue constrained spherical deconvolution for improved analysis of multi-shell diffusion MRI data. *NeuroImage* **103**, 411-426, doi:10.1016/j.neuroimage.2014.07.061 (2014).
81 Basser, P. J., Mattiello, J. & LeBihan, D. MR diffusion tensor spectroscopy and imaging. *Biophysical journal* **66**, 259-267, doi:10.1016/S0006-3495(94)80775-1 (1994).
82 Fieremans, E., Jensen, J. H. & Helpern, J. A. White matter characterization with diffusional kurtosis imaging. *NeuroImage* **58**, 177-188, doi:10.1016/j.neuroimage.2011.06.006 (2011).
83 Jensen, J. H., Helpern, J. A., Ramani, A., Lu, H. & Kaczynski, K. Diffusional kurtosis imaging: the quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *Magnetic resonance in medicine* **53**, 1432-1440, doi:10.1002/mrm.20508 (2005).
84 Zhang, H., Schneider, T., Wheeler-Kingshott, C. A. & Alexander, D. C. NODDI: practical in vivo neurite orientation dispersion and density imaging of the human brain. *NeuroImage* **61**, 1000-1016, doi:10.1016/j.neuroimage.2012.03.072 (2012).
85 Sotiropoulos, S. N., Behrens, T. E. & Jbabdi, S. Ball and rackets: Inferring fiber fanning from diffusion-weighted MRI. *NeuroImage* **60**, 1412-1425, doi:10.1016/j.neuroimage.2012.01.056 (2012).
86 Bastiani, M. *et al.* Improved tractography by modelling sub-voxel fibre patterns using asymmetric fibre orientation distributions. *ISMRM Proceedings, Singapore, p. 10* (2016).
87 Cottaar, M. *et al.* A generative model of white matter axonal orientations near the cortex. *ISMRM Proceedings, Toronoto, p. 351* (2015).
88 Glasser, M. F. & Rilling, J. K. DTI tractography of the human brain's language pathways. *Cereb Cortex* **18**, 2471-2482, doi:10.1093/cercor/bhn011 (2008).
89 Ramayya, A. G., Glasser, M. F. & Rilling, J. K. A DTI investigation of neural substrates supporting tool use. *Cereb Cortex* **20**, 507-516, doi:10.1093/cercor/bhp141 (2010).
90 Rilling, J. K., Glasser, M. F., Jbabdi, S., Andersson, J. & Preuss, T. M. Continuity, divergence, and the evolution of brain language pathways. *Frontiers in evolutionary neuroscience* **3**, 11, doi:10.3389/fnevo.2011.00011 (2011).
91 Rilling, J. K. *et al.* The evolution of the arcuate fasciculus revealed with comparative DTI. *Nature neuroscience* **11**, 426-428, doi:10.1038/nn2072 (2008).
92 Vincent, J. L. *et al.* Intrinsic functional architecture in the anaesthetized monkey brain. *Nature* **447**, 83-86, doi:10.1038/nature05758 (2007).

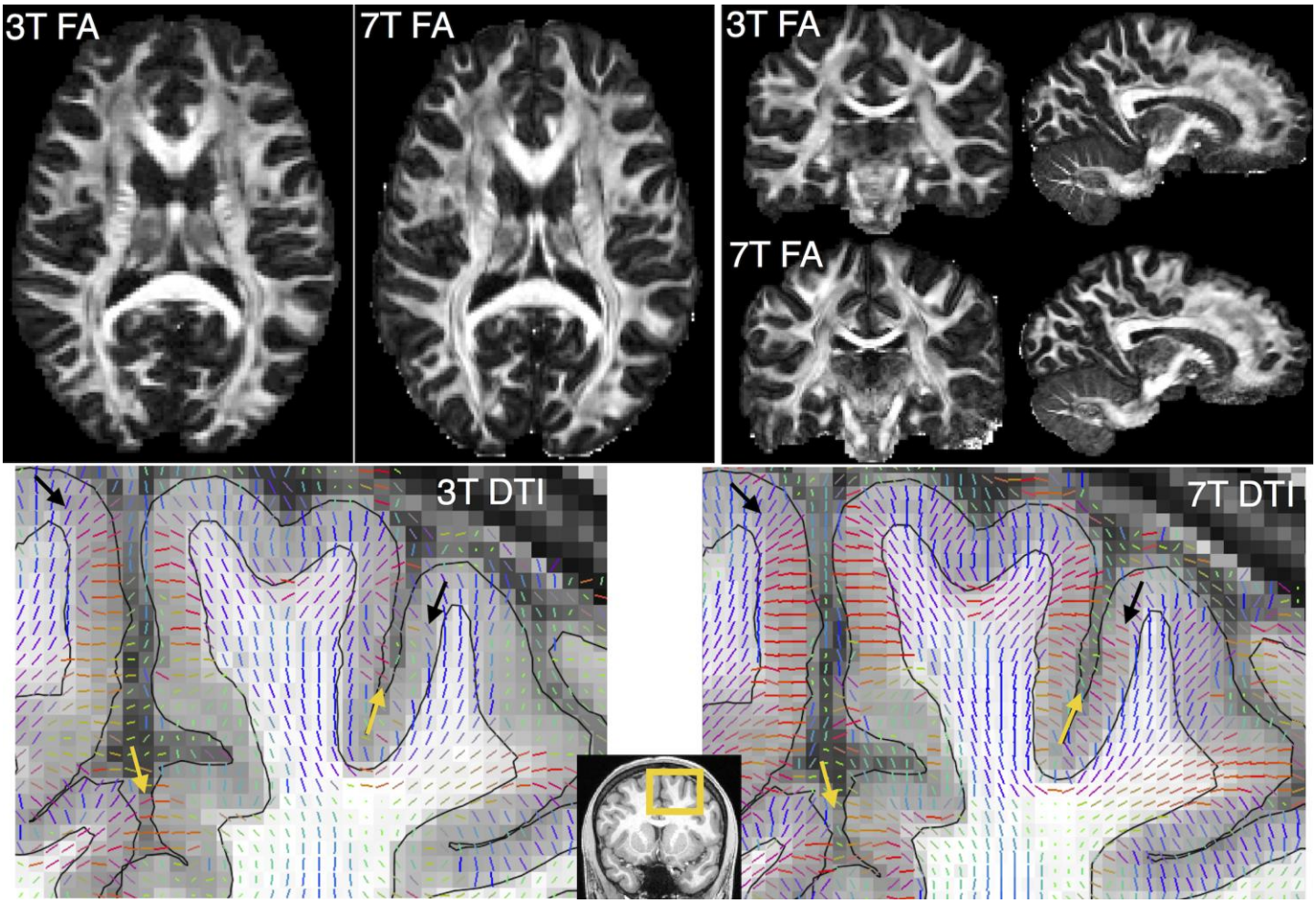
93 Smith, S. M. The future of FMRI connectivity. *NeuroImage* **62**, 1257-1266, doi:10.1016/j.neuroimage.2012.01.022 (2012).
94 Power, J. D., Schlaggar, B. L., Lessov-Schlaggar, C. N. & Petersen, S. E. Evidence for hubs in human functional brain networks. *Neuron* **79**, 798-813 (2013).
95 Haxby, J. V. *et al.* A common, high-dimensional model of the representational space in human ventral temporal cortex. *Neuron* **72**, 404-416, doi:10.1016/j.neuron.2011.08.026 (2011).
96 Fischl, B. *et al.* Cortical folding patterns and predicting cytoarchitecture. *Cereb Cortex* **18**, 1973-1980, doi:10.1093/cercor/bhm225 (2008).
97 Van Essen, D. C., Glasser, M. F., Dierker, D. L., Harwell, J. & Coalson, T. Parcellations and hemispheric asymmetries of human cerebral cortex analyzed on surface-based atlases. *Cereb Cortex* **22**, 2241-2262, doi:10.1093/cercor/bhr291 (2012).
98 Buckner, R. L., Krienen, F. M., Castellanos, A., Diaz, J. C. & Yeo, B. T. The organization of the human cerebellum estimated by intrinsic functional connectivity. *Journal of neurophysiology* **106**, 2322-2345, doi:10.1152/jn.00339.2011 (2011).
99 Yeo, B. T. *et al.* The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of neurophysiology* **106**, 1125-1165, doi:10.1152/jn.00338.2011 (2011).
100 Haber, S. N. & Behrens, T. E. The neural network underlying incentive-based learning: implications for interpreting circuit disruptions in psychiatric disorders. *Neuron* **83**, 1019-1039, doi:10.1016/j.neuron.2014.08.031 (2014).
101 Xu, J. *et al.* Evaluation of slice accelerations using multiband echo planar imaging at 3 T. *NeuroImage* **83**, 991-1001, doi:10.1016/j.neuroimage.2013.07.055 (2013).
102 Robinson, E. C. *et al.* MSM: A new flxible framework for multimodal surface matching. *NeuroImage* **100**, 414-426 (2014).



Supplementary Figure 1

Average cortical thickness map of 210 HCP subjects at each left hemisphere vertex and the associated colorized histogram.

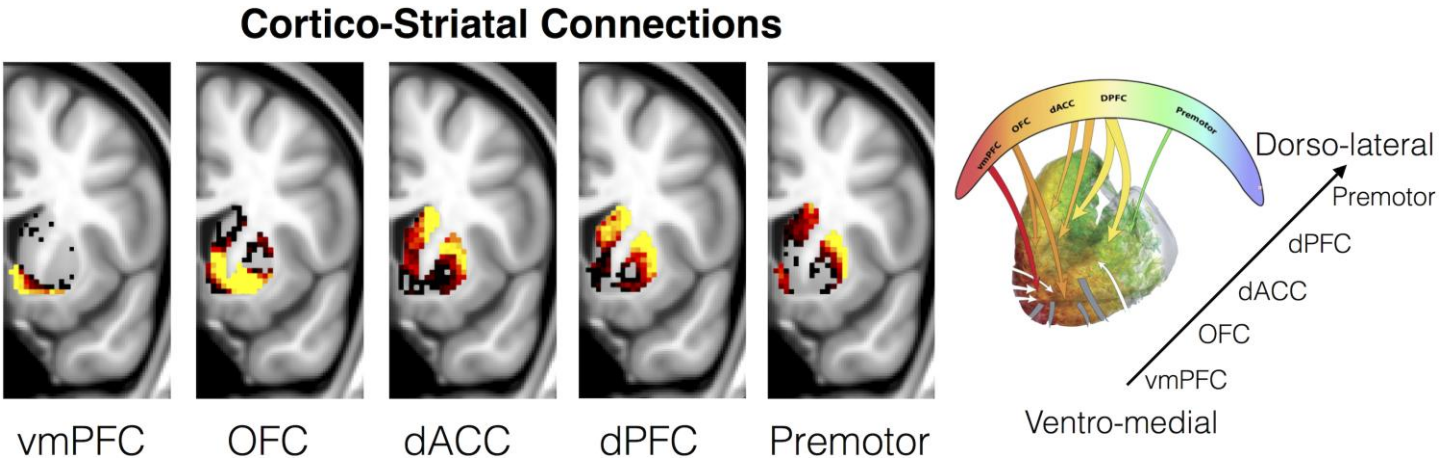
The mean cortical thickness is around 2.6 mm, and this roughly divides fMRI data into high resolution (less than the mean cortical thickness) and low resolution (more than the mean cortical thickness). The HCP's 3T and 7T chosen resolutions are also plotted. For the data in this figure and other Supplementary Figures, subject recruitment procedures and informed consent forms, including consent to share de-identified data, were approved by the Washington University institutional review board.



Supplementary Figure 2

dMRI data for 3T and 7T scans of the same subject (HCP Subject 158035).

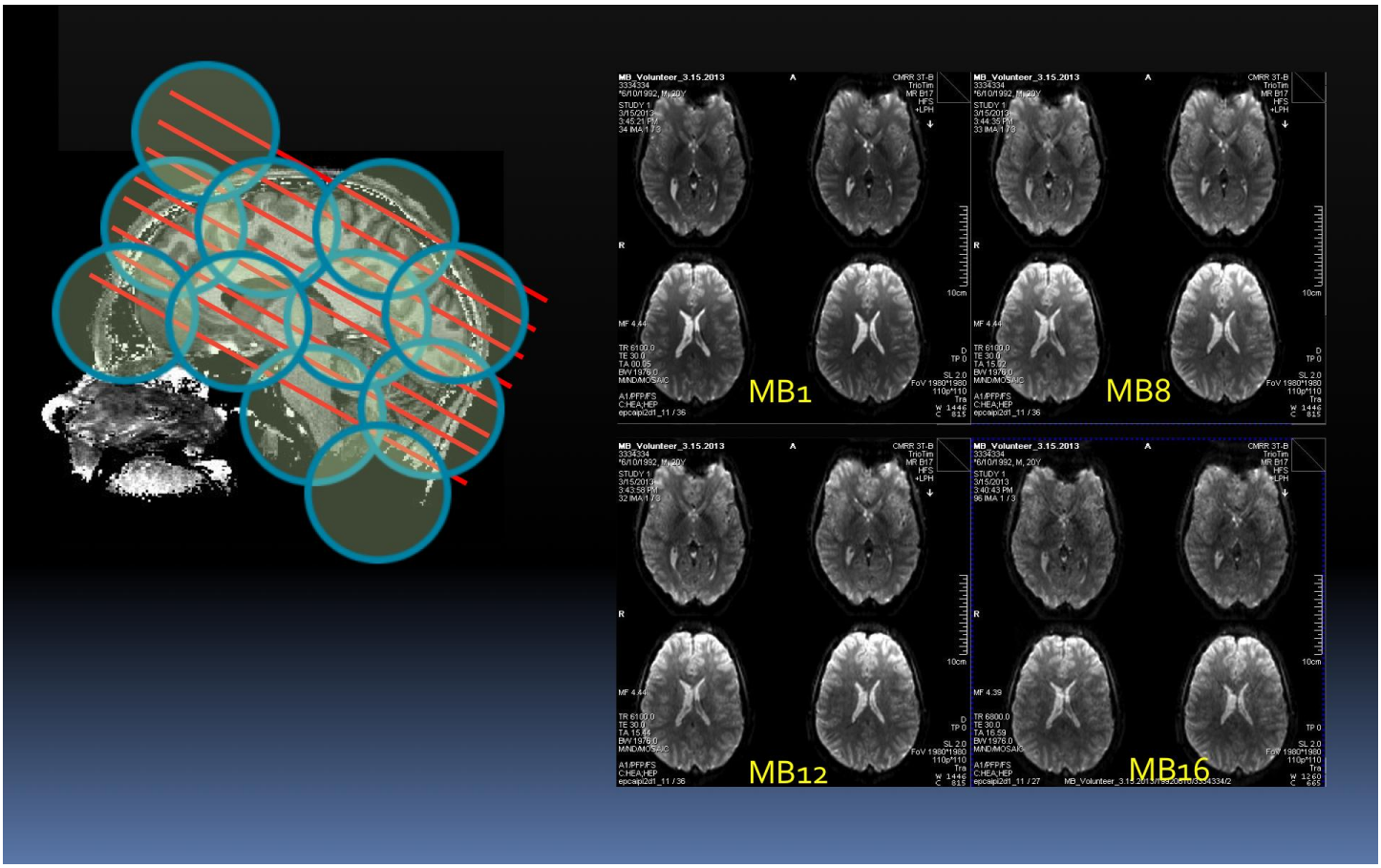
Top row: Fractional anisotropy (FA) maps (axial, coronal and sagittal views). Notice that B1 inhomogeneities at 7T lead to poor SNR and noisy FA estimates in the inferior temporal regions (evident in the coronal views), but efforts have been taken to minimize them[19]. Bottom row: DTI principal fiber orientations (coronal zoomed view of the area delineated by the yellow box). The orientations are RGB color-coded (Red: Left–Right, Green: Anterior–Posterior, Blue: Superior–Inferior) and superimposed on the structural T1w image. The pial surface and the WM/GM boundary surface are also shown. Reproduced, with permission, from Ref. 78.



Supplementary Figure 3

Patterns of cortico-striatal connectivity revealed by tractography.

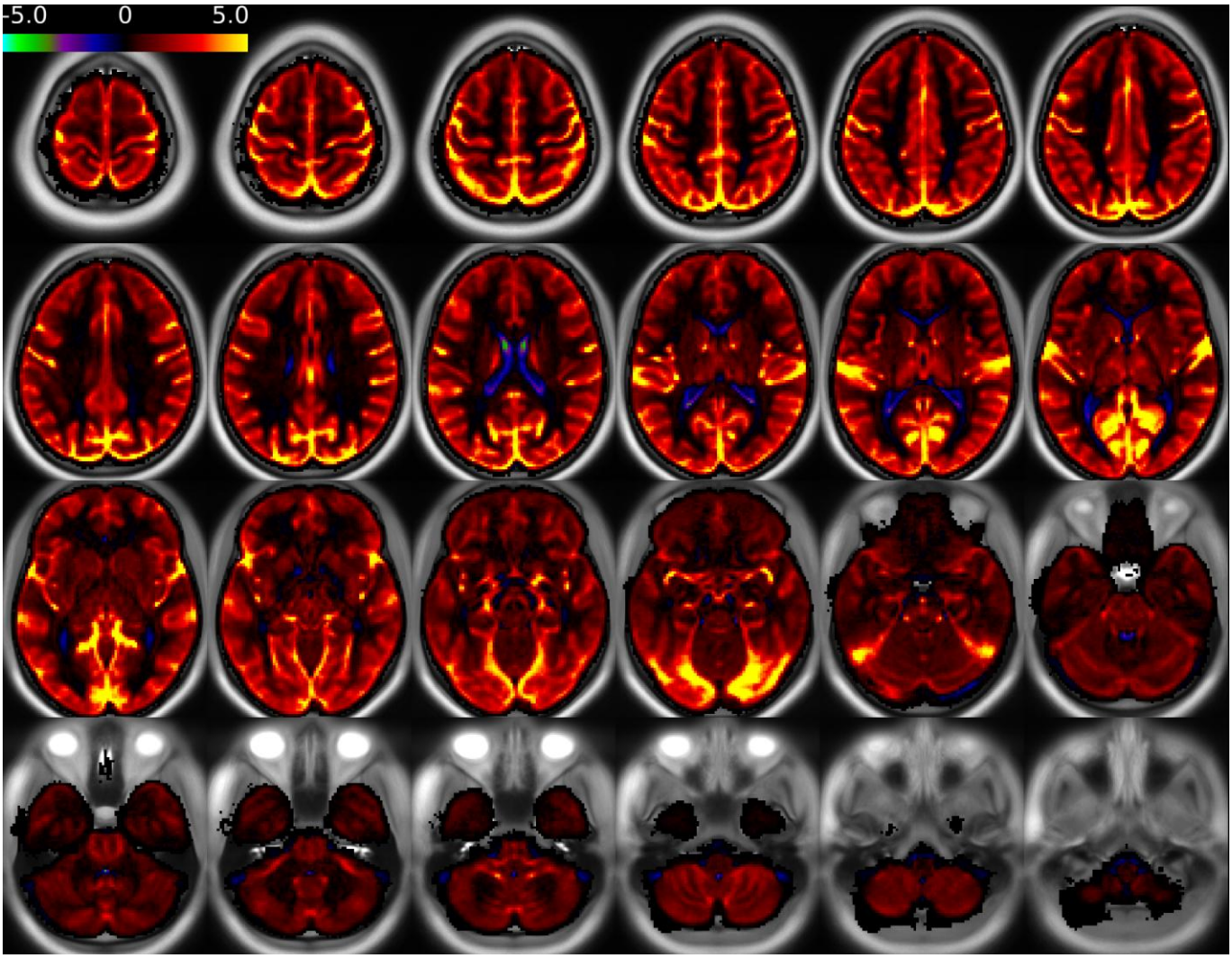
Seed locations were in different cortical regions, including vmPFC (ventromedial prefrontal cortex), OFC (orbito-frontal cortex), dACC (dorsal anterior cingulate cortex), dPFC (dorsal prefrontal cortex) and Premotor cortex. Path probabilities (yellow: high, red: intermediate, black: low) are obtained using probabilistic tractography (FSL's probtracx2, http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FDT/UserGuide) and the Matrix 3 (bidirectional white matter voxel to gray-matter terminations) algorithm to compute dense connectomes. N=150 HCP subjects were analyzed and averaged. Note the strong similarity to patterns of tracer-based connectivity reported in the macaque monkey, shown in the sketch on the right (reproduced with permission from Ref. 100).



Supplementary Figure 4

Multi-band imaging schematic and exemplar results.

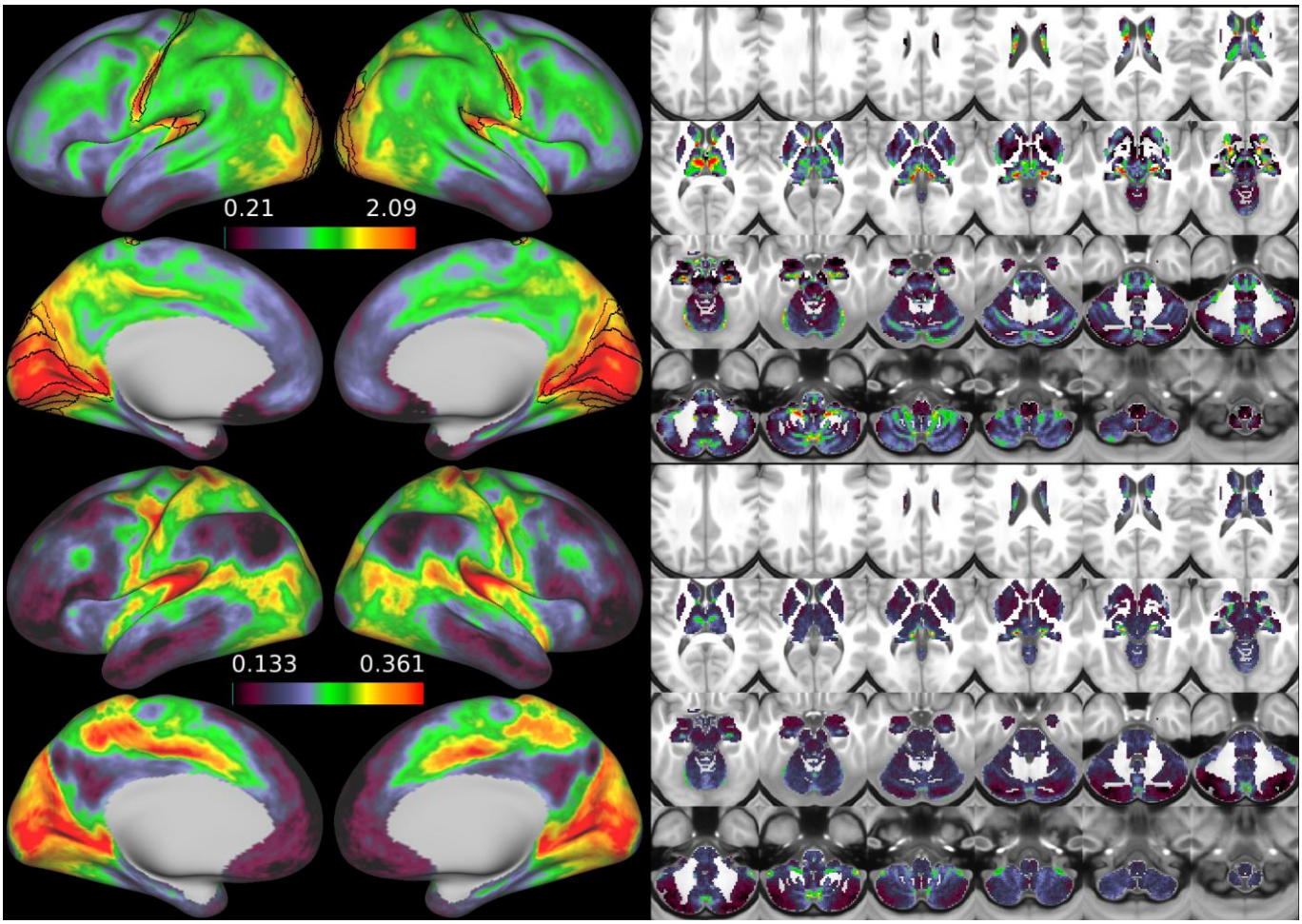
Left: A schematic representation of the array coil elements over a human head image and a multi-band (MB) excitation (8 slices, in red). Each coil detects a linear combination of signals from each slice weighted by the sensitivity profile of that coil. Right. Four slices from a whole brain data set with standard acquisition (upper left, MB1) and with SMS/MB acquisition with MB=8, MB=12, and MB=16, all obtained using a 32-channel coil on a 3T scanner. Visual inspection of the M=16 image reveals discernible artifacts (e.g., ghosting), whereas the MB=12 and MB=8 images appear much cleaner. Quantitative estimates of cross slice contamination and reconstruction noise can be made[101]. The HCP used MB=8 for fMRI and MB=3 for dMRI (now MB=4 or more for HCP Short).



Supplementary Figure 5

Beta map of the mean fMRI timeseries.

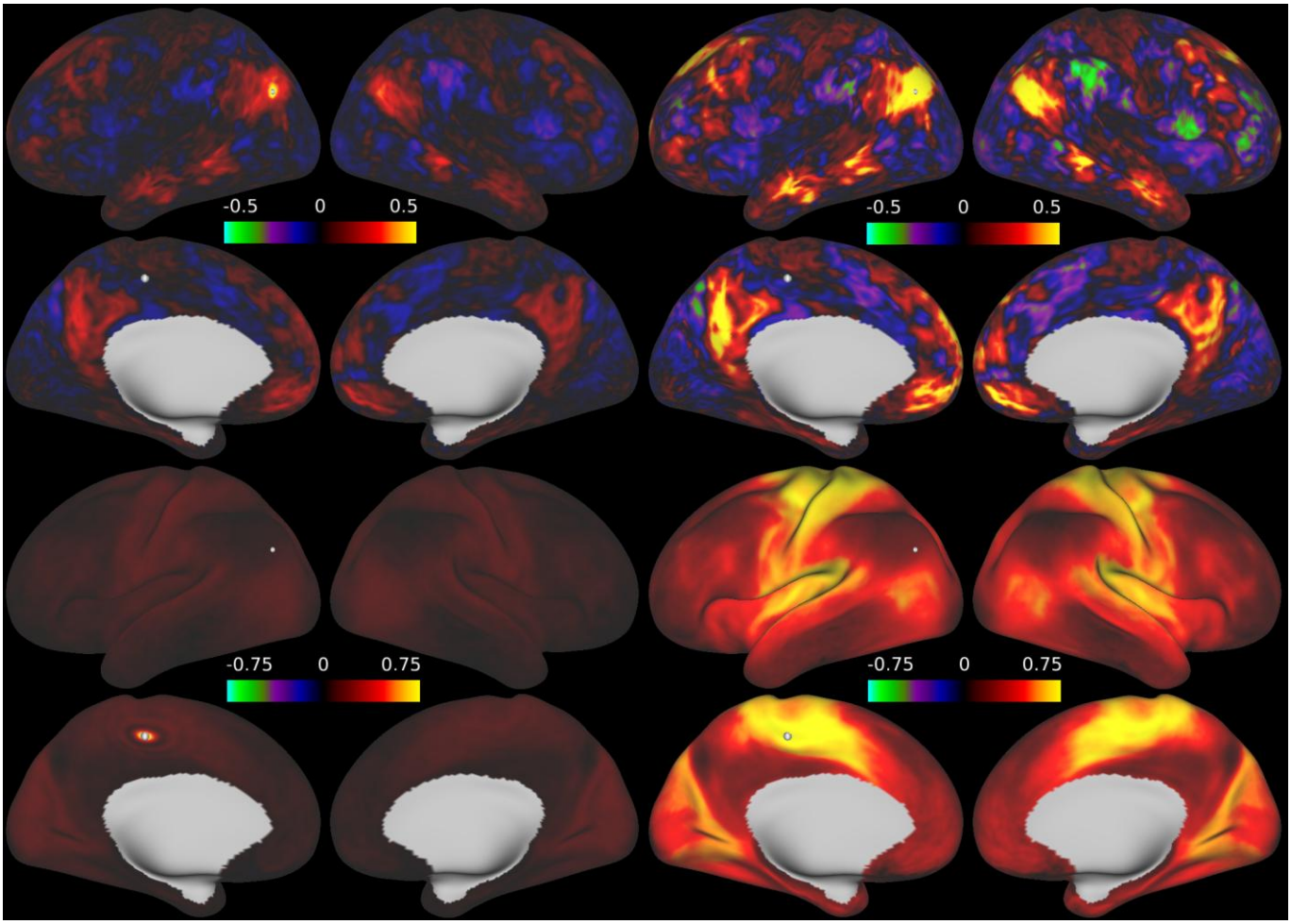
The timecourses were averaged over the whole brain (including gray matter, white matter, and CSF), regressed into the data of each subject, and then averaged across (n = 210 HCP subjects). One particularly striking characteristic of this map is how tissue-specific the global signal is (after ICA+FIX data cleanup), being generally most positive in grey matter, close to zero in white matter, and slightly negative in the ventricles. The tissue specificity of this signal argues against a non-physiologic, non-BOLD contrast-based cause of the signal (e.g., direct biophysical effects of subject head motion). This map by itself does not tell us to what extent the global signal is physiological noise vs. neural signal. Although the data are averaged in the volume across subjects, they still appear relatively sharp because they are not smoothed. Data at http://balsa.wustl.edu/0L1m.



Supplementary Figure 6

Visualization of the mean grey signal.

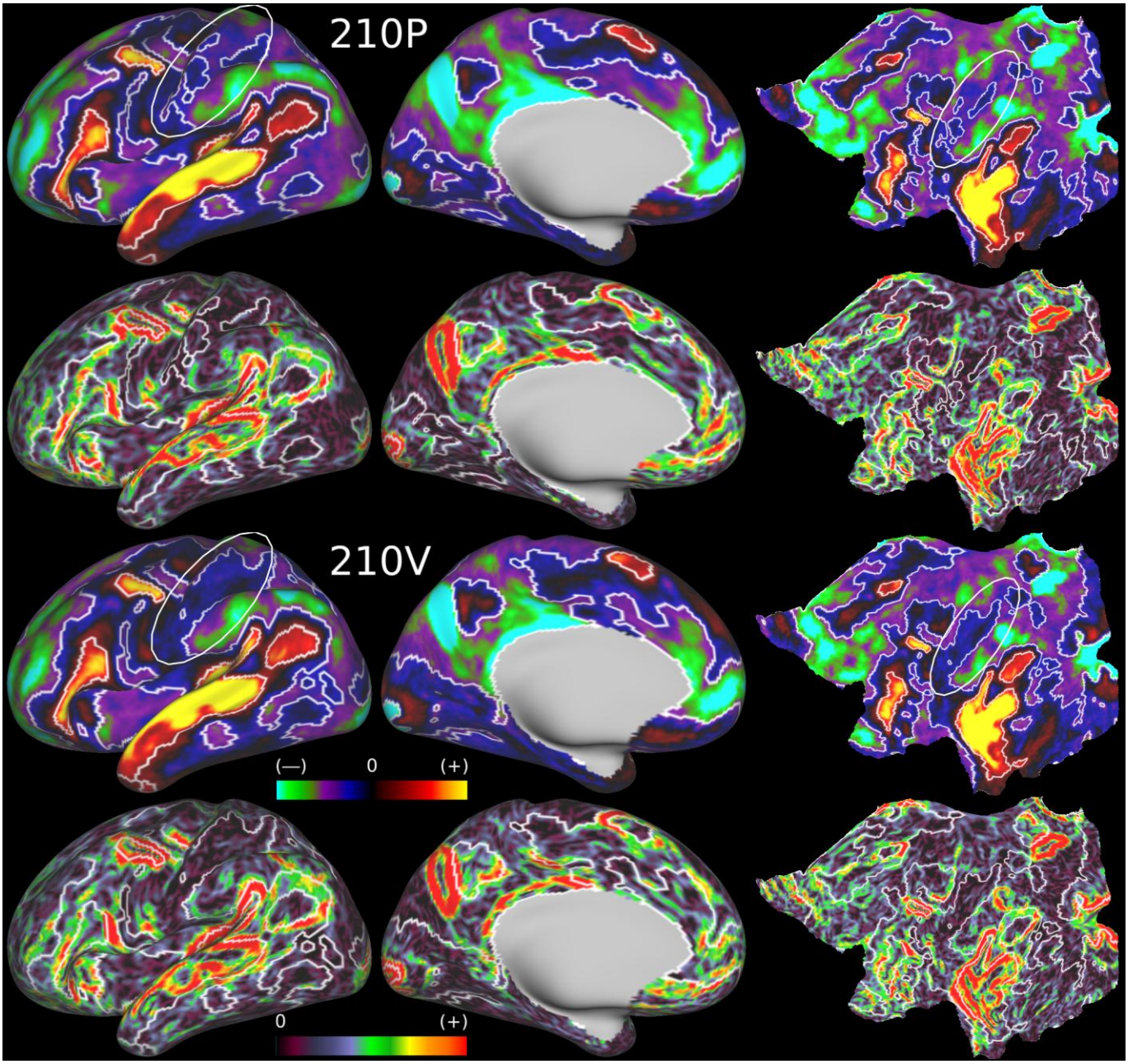
The mean grey timeseries beta map (top rows) and the ratio of the variance of the mean grey timeseries to the total BOLD variance (i.e., variance classified as signal by ICA+FIX; bottom rows). The absolute magnitude of the mean grey signal is highest in sensory regions including visual cortex, early somatosensory cortex (particularly of the face), early auditory cortex, and several thalamic nuclei including the LGN/MGN. Visual, somatosensory, auditory, and likely vestibular cortical areas are highlighted with black outlines. Data based on averaging across 210 HCP subjects, aligned using MSMAll for the cortical surface. When the global signal is strong in individual subjects, it closely matches the group average pattern, however when it is weaker, it may or may not match the group pattern. In these cases, the global signal often looks like one or another widespread RSN (e.g., the task positive or task negative (default mode) network). This is further evidence that removing the global signal as a preprocessing step may distort resting state functional connectivity and hence that we need a more selective way to clean global noise out of the data. The bottom row is a relative measure of how much the fMRI timeseries will be altered by regressing out the mean grey timeseries (on average across subjects), as it is a measure of the proportion of the total BOLD signal represented by the mean grey timeseries. Regressing out the mean grey signal will also tend to cause resting state gradient boundaries to move somewhat in the regions where this map has sharp gradients[34]. Data at http://balsa.wustl.edu/2VnN.



Supplementary Figure 7

Effects of the Wishart rolloff on dense functional connectivity maps of both an individual subject and group data (210 HCP subjects; MSMAll surface registration).

Top rows show an individual subject, before (column 1) and after (column 2) Wishart rolloff for a seed location in lateral parietal cortex (white dot in upper left panels). The correlation increases dramatically as unstructured spatio-temporal noise is reduced, however the map is not substantially "smoothed" as it would be with typical smoothing algorithms. Bottom rows show a group dataset before and after Wishart rolloff for a seed location in the posterior cingulate sulcus (white dot in lower left panels). The dataset has been created using the MIGP[57] algorithm to generate a group PCA series (d=4500) that represents the group concatenated timeseries. Because of the hard cutoff at PCA component number 4500, there is a 'ringing" pattern resulting from spatial autocorrelation in the spatio-temporal noise that is represented by the PCA components with the lowest eigenvalues. If a Gaussian filter had been applied, this pattern of "local connectivity" would be a blob instead of rings. The Wishart rolloff eliminates these rings and again dramatically increases the SNR of the data. Data at http://balsa.wustl.edu/rrpl.

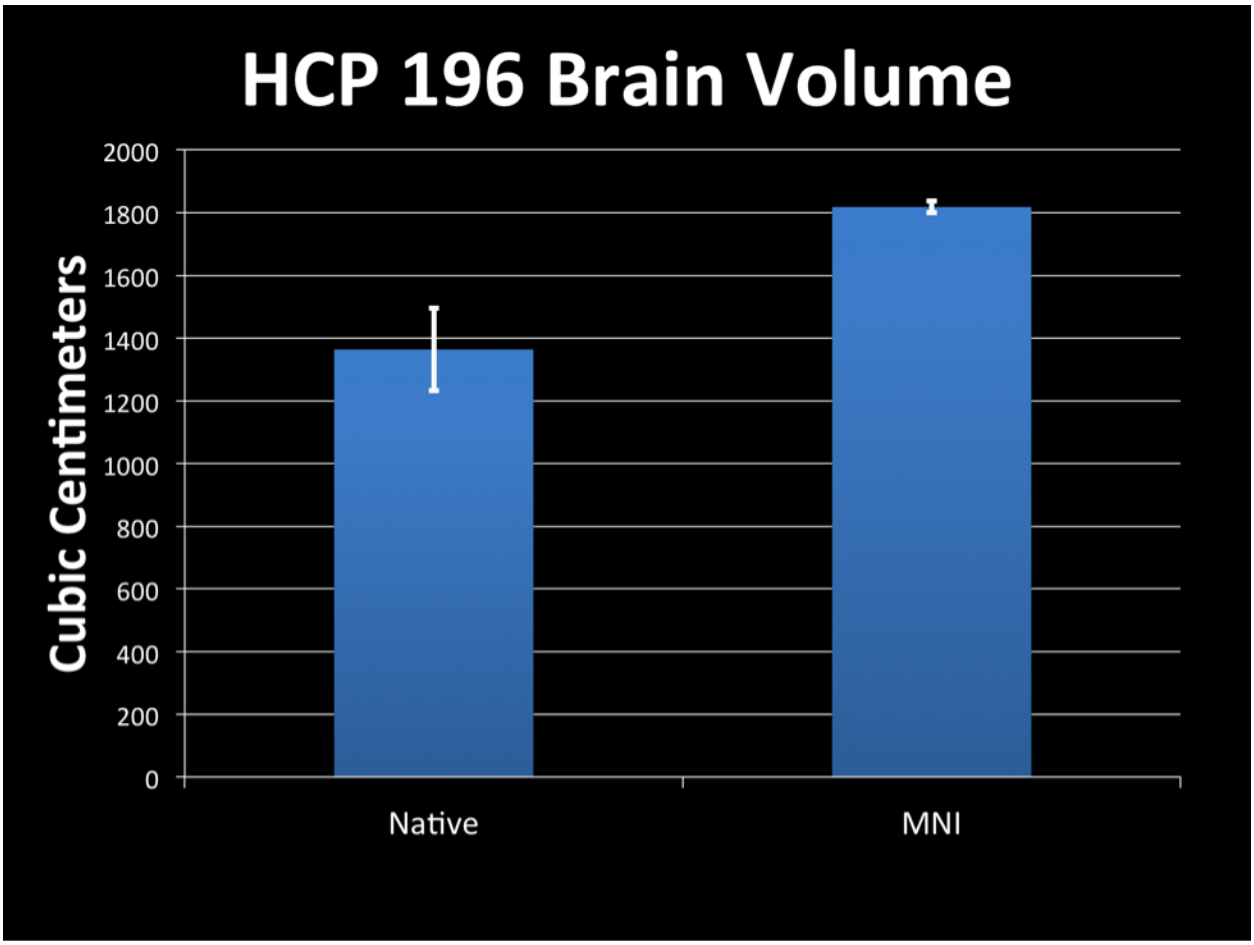


Supplementary Figure 8

The HCP language task (story vs baseline) beta maps and their spatial gradients.

Beta maps (rows 1 and 3) and gradient maps (rows 2 and 4) are from two independent groups of 210 HCP subjects, "210P" (rows 1 and 2) and "210V" (rows 3 and 4). Because of the large number of high quality HCP subjects, the beta maps are very similar across the two groups, and the strong gradients in the beta maps are also very similar. Also shown are white contours of a Bonferroni corrected significance threshold across all 91282 grayordinates (z+/- ~5). Two things are apparent: 1) Because of the large amount of high quality data, most of the brain is either significantly activated or deactivated (an issue that has been discussed before, see Ref. 60). Thus the statistical threshold is not particularly biologically meaningful (a point about statistical thresholds that generally applies). At the same time, the statistical threshold is also not strongly reproducible, in spite of the large amount of high quality data (highlighted ellipses show large differences in the area of activation classified as "significant" that are not particularly impressive when viewing the unthresholded beta maps). In contrast, the strong gradients in the beta maps are much more reproducible, are likely also

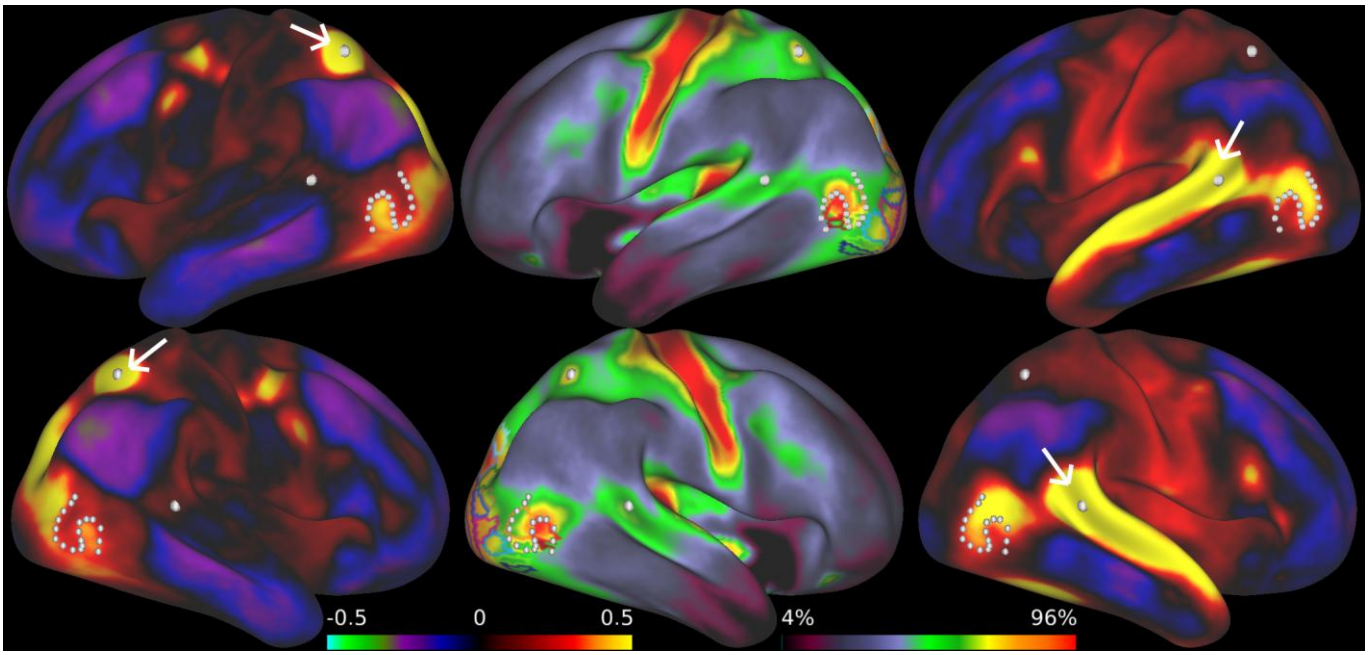
more biologically meaningful, and hence provide a better substrate for defining regions of activation or comparing across studies. Data at http://balsa.wustl.edu/PrmK.



Supplementary Figure 9

Effects on average brain volume of registering 196 HCP brains to MNI space.

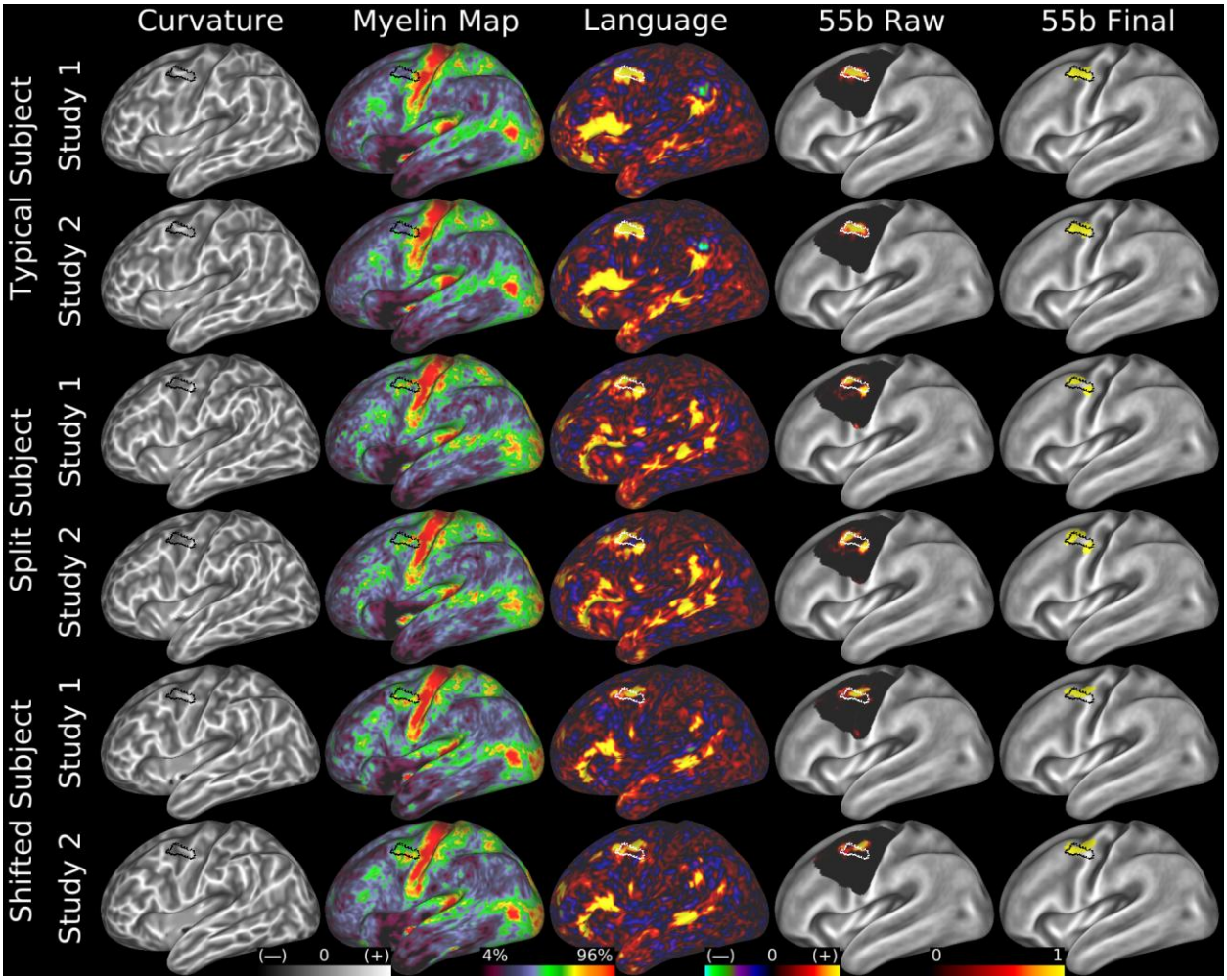
The total brain volume (minimal preprocessing pipelines' whole brain mask) measured in the subject's native physical space is around ~1350 cc; however, after registration to MNI space it is ~1800 cc, though the variability in brain volumes goes down as indicated by the narrower standard deviation error bars. This reduction in variability is the intended effect of registration, but the increase in brain volume is group average registration drift that was built into MNI standard space during the non-rigid registrations of the template generation process[62-64].



Supplementary Figure 10

A comparison between the HCP data and published retinotopic parcellation data.

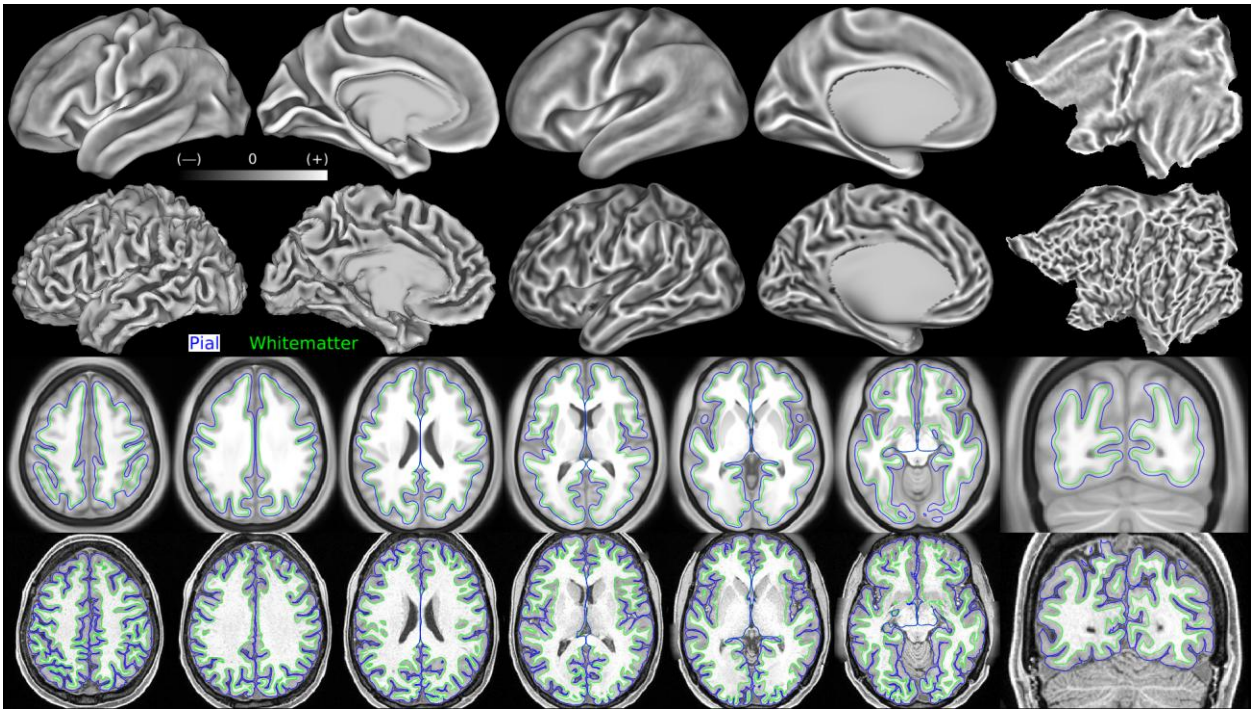
Data from Ref. 64 and from 120 HCP subjects (from Q1-2) were registered using MSM areal-feature-based registration[102] and group average registration drift was removed from both. Because of this, a contour in functional connectivity in the MT+ region distinguishing strong connectivity to the heavily myelinated IPS hotspot (LIPv, column 1) and to the STS (column 3) lines up with the border between MT+V4t (orange and yellow) and MST+FST (red and maroon, middle column). This illustrates the kind of precise cross-modal, cross-study boundary comparisons possible using the HCP-Style paradigm. Data at http://balsa.wustl.edu/x2Lz.



Supplementary Figure 11

Classification of area 55b in individual subjects by the areal classifier.

The typical location of area 55b is shown in black or white outlines on the inflated left hemisphere surface. The top two rows show a subject having an area 55b in the typical location found in the population. Rows 1 and 2 are entirely separate 'test/retest' runs of this subject through the full HCP MRI acquisition and analysis protocol. Column 1 is the subject's individual curvature map, column 2 is the subject's myelin map, column 3 is a d=40 RSN map that shows strong connectivity between 55b and other areas of the language network, column 4 is the raw probabilistic output of the areal classifier, column 5 is the final output of the areal classifier. Rows 3 and 4 show a different subject whose 55b is atypically split (heavy myelination running through the population average location of 55b and a concomitant lack of resting state connectivity). In both runs through the protocol, this subject shows a split 55b that is accurately detected by the areal classifier, showing that the classifier can accurately classify atypical subjects and that these atypical patterns are stable across time. Rows 5 and 6 show a third subject whose 55b is atypically shifted relative to its neighbors. Again the pattern is stable across time and the areal classifier is able to accurately detect the area. Reprinted from Ref. 34. Data at http://balsa.wustl.edu/WPPn.



Supplementary Figure 12

Effects of averaging surface coordinates and folding maps after areal feature-based registration (MSMAll).

The first row shows the group average midthickness surface (left), and the group average curvature map also displayed on inflated and flat surfaces (center and right). The group consisted of 210 subjects. The second row shows an individual subject's midthickness surface (left), and the individual's curvature map displayed also on inflated and flat surfaces (center and right). Note how much less detailed the group average surfaces and curvature maps are in most regions of cortex. However, in regions with consistent folding patterns across subjects and consistent relationships between cortical areas and folds, the group average patterns remain sharp (e.g. the central and calcarine sulci and the insula). The third row shows the group average T1w volume (after FNIRT nonlinear volume registration to MNI152 space of each subject in the group) together with the group average white (green) and pial (blue) surface contours. The fourth row shows the same individual subject's T1w volume together with the individual's white and pial surfaces (after aligning both the T1w volume and the surfaces to the group average in MNI space using FNIRT nonlinear volume registration). Despite the high precision of the white and pial surfaces in following the grey matter ribbon in the individual, the group average surfaces do not follow the group average volume particularly well, except in the regions where there are consistent folding patterns across subjects and consistent relationships between cortical areas and folds (as mentioned above). These issues also occur with folding-based surface registration (not shown), though they are less severe, because for folding-based registration the dominant factor is inconsistency in folding patterns across subjects (as no attempt is made to directly align cortical areas). The midthickness surfaces are the average of the white and pial surfaces (this average is performed on each individual, and the group midthickness surface is the average of the individual midthickness surfaces). Data at http://balsa.wustl.edu/7qP3.