Behavioral/Cognitive

Multimodal Imaging Brain Markers in Early Adolescence Are Linked with a Physically Active Lifestyle

Piergiorgio Salvan,1 Thomas Wassenaar,1 Catherine Wheatley,1 Nicholas Beale,2 Michiel Cottaar,1 Daniel Papp,1 Matteo Bastiani,1,3,4 Sean Fitzgibbon,2 Eugene Duff,1 Jesper Andersson,1 Anderson M. Winkler,5,6 Gwenaëlle Douaud,1 Thomas E. Nichols,1,7,8 Stephen Smith,1 Helen Dawes,2 and Heidi Johansen-Berg1

1Wellcome Centre for Integrative Neuroimaging, FMRIB Centre, Nuffield Department of Clinical Neurosciences, University of Oxford, John Radcliffe Hospital, Oxford, OX3 9DU, United Kingdom, 2Centre for Movement, Occupational and Rehabilitation Sciences, Oxford Brookes University, Oxford, OX3 0BP, United Kingdom, 3Sir Peter Mansfield Imaging Centre, School of Medicine, University of Nottingham, Nottingham, NG7 2RD, United Kingdom, 4National Institute for Health Research Biomedical Research Centre, University of Nottingham, Nottingham, NG7 2UH, United Kingdom, 5Emotion and Development Branch, National Institute of Mental Health, National Institutes of Health, Bethesda, MD 20892-9663, Maryland, 6Department of Psychiatry, Yale University School of Medicine, New Haven, CT 06511, Connecticut, 7Oxford Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, Nuffield Department of Population Health, University of Oxford, Oxford, OX3 7LF, United Kingdom, and 8Department of Statistics, University of Warwick, Coventry, CV4 7AL, United Kingdom

The World Health Organization promotes physical exercise and a healthy lifestyle as means to improve youth development. However, relationships between physical lifestyle and human brain development are not fully understood. Here, we asked whether a human brain–physical latent mode of covariation underpins the relationship between physical activity, fitness, and physical health measures with multimodal neuroimaging markers. In 50 12-year old school pupils (26 females), we acquired multimodal whole-brain MRI, characterizing brain structure, microstructure, function, myelin content, and blood perfusion. We also acquired physical variables measuring objective fitness levels, 7 d physical activity, body mass index, heart rate, and blood pressure. Using canonical correlation analysis, we unravel a latent mode of brain–physical covariation, independent of demographics, school, or socioeconomic status. We show that MRI metrics with greater involvement in this mode also showed spatially extended patterns across the brain. Specifically, global patterns of greater gray matter perfusion, volume, cortical surface area, greater white matter extra-neurite density, and resting state networks activity covaried positively with measures reflecting a physically active phenotype (high fit, low sedentary individuals). Showing that a physically active lifestyle is linked with systems-level brain MRI metrics, these results suggest widespread associations relating to several biological processes. These results support the notion of close brain-body relationships and underline the importance of investigating modifiable lifestyle factors not only for physical health but also for brain health early in adolescence.

Key words: adolescence; canonical correlation analysis; fitness; lifestyle; multimodal MRI; physical activity

Significance Statement

An active lifestyle is key for healthy development. In this work, we answer the following question: How do brain neuroimaging markers relate with young adolescents’ level of physical activity, fitness, and physical health? Combining advanced whole-brain multimodal MRI metrics with computational approaches, we show a robust relationship between physically active lifestyles and spatially extended, multimodal brain imaging-derived phenotypes. Suggesting a wider effect on brain neuroimaging metrics than previously thought, this work underlies the importance of studying physical lifestyle, as well as other brain–body relationships in an effort to foster brain health at this crucial stage in development.
Introduction

The World Health Organization encourages early positive lifestyle choices aimed to improve both physical and mental health (World Health Organization, 2010). Physical activity is a powerful and rapid means to improve fitness and physical health throughout the life-span (Cotman, 2002; Hillman et al., 2008). During adolescence, however, levels of physical activity decline (Guthold et al., 2020).

Public health guidelines recommend that school-aged children engage in 60 min of moderate-to-vigorous physical activity daily (Piercy and Troiano, 2018), yet globally only ~22% of boys and 15% of girls achieve that (Guthold et al., 2020). In addition to its importance to physical health, there is growing evidence that a physically active lifestyle during childhood is associated with improved mental and cognitive health through adulthood (Department of Health & Human Services Office of Disease Prevention and Health Promotion, 2000). While there is limited available evidence in adolescents, similar patterns have been reported (Lubans et al., 2016).

A body of work has studied the relationship between single physical measures of activity, fitness, or body mass, and separate MRI metrics of brain structure, microstructure, or function, showing focal neural correlates (for review, see Donnelly et al., 2016; Valkenboorns et al., 2019). However, it is unlikely that a single physical measure fully captures active lifestyles, or that a single MRI metric fully quantifies the condition of the brain. Rather, lifestyles are better characterized by a range of physical measures, and the state of the brain is better quantified by combinations of metrics.

Multimodal MRI can probe different aspects of brain structure and function. While each metric provides an indirect probe of the underlying biology, in combination they provide insights into a range of biological processes (Tardif et al., 2016). Further, these measures can be acquired simultaneously across the whole brain. Many previous brain imaging studies of physical activity and fitness have focused on the hippocampus, where changes in noninvasive imaging measures of tissue volume or perfusion have been argued to relate to processes of neurogenesis and angiogenesis triggered by exercise (van Praag et al., 1999; Pereira et al., 2007; Chaddock et al., 2011; Thomas et al., 2012). However, in addition to such focal changes, more global biological processes might also be triggered by exercise (Tardif et al., 2016). It remains unknown whether whole-brain patterns of multimodal brain metrics are related to cardiorespiratory fitness, physical activity, and physical health.

Physical activity influences physical health and contributes to physical fitness, but both activity and fitness may be considered part of an underlying, latent factor. In order to characterize a phenotype of physical lifestyle, measuring whole-day physical activity levels during a normal school week is therefore at least as important as assessing gold-standard measures of cardiorespiratory fitness, such as VO2max measured on an incremental step-test on a cycle ergometer.

In this study, in 50 12-year-old pupils, we acquired multimodal whole-brain MRI metrics to measure resting state networks (RSNs), gray matter (GM) volume and perfusion, cortical surface (area and thickness), white matter (WM) microstructure, and myelin content (R1 and R2*), resulting in a total of 18 different metrics. These metrics are combined into multimodal whole-brain phenotypes whose variation across individuals can be interrogated. We also acquired a rich set of variables depicting physical lifestyle, measuring cardiorespiratory fitness (VO2max and workload), objective physical activity (7 d actigraphy, measuring total week time of brief bursts and long-lasting physical activity), and reported (questionnaire item) and physical health (resting heart rate, blood pressure, and body mass index) (Fig. 1). We hypothesized that, across pupils, intersubject differences in brain phenotypes covaried with differences in physical lifestyle, independent of sex, socioeconomic status, age, pubertal level, and school. A single holistic multivariate analysis allowed us to identify a latent mode of covariation between brain and physical phenotypes, representing a pattern of active physical lifestyle features that significantly covaries with spatially extended patterns of brain metrics.

Materials and Methods

Participants and recruitment

Year 7 pupils from a subset of 10 United Kingdom schools participating in the Fit to Study project Main trial (total 93 schools) were invited to take part in a brain imaging substudy (Wassenaar et al., 2019). The 10 recruitment schools were selected for being conveniently located for travel to Oxford. Researchers visited recruitment schools to pitch the study to pupils and collected pupils’ expressions of interest. For each school, an upper limit was determined in order not to over-recruit from a given school. For schools in which the number of interested pupils exceeded this limit, pupils with lower physical activity scores (based on values already collected through the Main trial) were given priority to select a sample that was the most representative as possible of the entire population. The expressions of interest, however, varied greatly by school. Some schools were therefore topped up with more pupils despite creating an imbalance to pragmatically increase study sample size.

After taking consent and assent in accordance with the University of Oxford ethical guidelines (CUREC reference number: R51313/RE001), 61 pupils were recruited to the brain imaging substudy. Participants attended a testing session in the University of Oxford during which brain imaging, cognitive, and behavioral data were collected.

The analysis required high-quality complete multimodal MRI data. One participant withdrew during the scan session. One complete dataset was lost because of hardware failure. Of the remaining datasets, quality control process identified issues with data quality (e.g., head motion, ringing artifacts, blurring, etc.) in one or more modalities in 9 pupils. Therefore, only 50 pupils (median age: 12 years; 26 females, 52%; Table 1) had high-quality, complete multimodal MRI data that could be taken forward into our final analysis. This substudy population is representative of the larger Fit to Study population in terms of demographics (Table 1). However, the substudy pupils were more active and less likely to qualify for free school meals. These differences can be explained by difficulties in recruitment (i.e., low socioeconomic status households were less keen to travel and more active pupils might have been more interested in participation). All statistical analyses were conducted on this sample of 50 pupils sampled from 10 schools (Table 2). The number of participants per school ranged from 1 to 13. We consistently followed the recruitment process described above, but levels of interest varied...
considerably across schools because of several reasons, including variable interest from pupils, concerns from parents regarding the study or travel distance, and unavailability during summer holidays.

**Behavioral testing**
All pupils underwent a half-day testing session at the Functional Magnetic Resonance Imaging of the Brain (FMRIB) building in Oxford during summer 2017. Over a period of ~5 h, and with multiple breaks, pupils performed, in this order, cognitive testing, multimodal MRI scans, physical health monitoring, and physical activity testing. Actigraphy monitoring aimed at capturing activity levels over a normal school week time was conducted previous to the day of testing at FMRIB.

**Cardiorespiratory fitness**
Objective measures of cardiorespiratory fitness were acquired through an incremental step-test on a cycle ergometer (Lode Excalibur Sport). We then extracted values for maximal oxygen consumption per kilogram (VO2/kg max) (ml/min/kg), and work load maximum (Watts) as primary measures of interest.

**Physical activity**
Objective physical activity was assessed over 5 weekdays and 2 weekend days using the Axivity AX3 wrist-worn accelerometer (Open Lab, Newcastle University) (Ladha et al., 2013). We therefore chose to define a valid wear day as 12 consecutive hours from 08:00 to 20:00 to capture travel to and from school and after-school sports and activities. To account for later weekend waking times, we accepted any consecutive 10 h period between 08:00 and 20:00 on Saturdays and Sundays, and standardized total activity to 12 h. We then aimed to capture both brief bursts and long-lasting activity. We summarized raw accelerometer data from three axes of movement into the signal vector magnitude, or activity “count,” expressed per 60 s epoch and also per 1 s epoch to characterize sustained bouts of activity and also shorter bursts of movement. Axivity’s Open Movement GUI software calculated whether each 60 s epoch was spent in sedentary, light, moderate, or vigorous activity by applying “cut-points” or “count” thresholds corresponding to different activity intensities derived from a validation study with young people aged 8-14 (Phillips et al., 2013). The software identified nonwear time as periods of at least 30 consecutive minutes of zero activity counts. We used a bespoke program, designed to handle large volumes of data, to apply the same cut-points to each 1 s epoch. Participants who had at least three valid weekdays and one valid weekend day were included in the analysis (Troiano et al., 2014). For both brief bursts and long-lasting physical activity, participants’ total minutes of sedentary; moderate, and vigorous activity per day were calculated.

**Physical health**
Physical health was assessed on the day of testing at rest (prior cardiorespiratory testing) by measuring heart rate and systolic and diastolic blood pressure. Compared with publicly available age-matched normative values (Flynn JT et al., 2018), blood pressure (5th-95th percentiles) was found within healthy values (normative values for 12 years old pupils: systolic: 102-131 with average of 113; diastolic: 61-83 with average of 75; study sample, systolic: median = 106, 5th-95th percentiles: 89-123; diastolic: median = 73, fifth-95th percentiles: 58-85).

**Negative behaviors not considered in the analysis**
As part of the study, we also obtained ethics to ask pupils information about negative behaviors, such as smoking, drinking alcohol, or drug use. However, none of the pupils reported having used any of these substances.

**MR imaging**
All MRI scans were conducted during summer 2017 at the Oxford Center for FMRIB using a 3T Siemens Magnetom Prisma scanner with a 32-channel head coil.

The MRI protocol included the following:
1. **T1 weighted (T1w)** three-dimensional rapid gradient echo sequence (3D MPRA GE): TR = 1900 ms; TE = 3.97 ms; flip angle = 8°; FOV = 192 mm; voxel size: 1×1×1 mm. Sequence duration: 5 min 31 s.
2. **Diffusion-weighted MRI (DW-MRI)**: multishell, multiband EPI sequence; b values = 0, 1250, 2500 s/mm², with, respectively 11, 60, 60 diffusion-weighted directions; TR = 2483 ms; TE = 78.20 ms; FOV = 214 mm; voxel size: 1.75×1.75×1.75 mm; multiband acceleration factor = 4. Sequence duration: 5 min 40 s. In addition, 4 b = 0 s/mm² images were acquired with reversed phase encoding, for the purpose of EPI distortion correction. Sequence duration: 32 s.
3. **Quantitative FLASH-MRI** (Weiskopf et al., 2013): two 3D multiecho FLASH datasets, one predominantly proton-density weighted (PDw, flip angle = 6 deg), and one predominantly T1w (flip angle = 21 deg); FOV = 256 mm; voxel size: 1×1×1 mm; TR = 25 ms; first TE = 2.34 ms; eight equally space echoes, echo spacing = 2.3; GRAPPA acceleration factor = 2 in both phase-encoded directions, with 40 reference lines in each direction. Duration for each FLASH sequence: 5 min 11 s. Two single-echo, low-resolution (4 mm isotropic) FLASH scans were acquired before each high-resolution scan; identical FOV; TR = 4 ms; TE = 2 ms; one was acquired receiving on the 32-channel receive head coil, the other receiving on the body coil. To correct for the effect of RF inhomogeneities, the local RF field was mapped using a 2D DAM method with a FLASH readout.
4. **Pseudo-continuous arterial spin labeling with background saturation** (Okell et al., 2013): six imaging blocks, each with different post-labeling delays 0.25, 0.5, 0.75, 1, 1.25, and 1.5 s. Arterial blood was magnetically tagged using a labeling duration of 1.4 s. Other imaging parameters were as follows: single-shot EPI; TR = 4100 ms;
In order to provide a more comfortable experience, during all structural scans, a wildlife documentary was shown. A fixation cross was instead shown during rs-FMRI and ASL in order not to bias cognitive processing to certain areas/networks during assessment of resting brain activity.

**MRI preprocessing**

MRI data were processed primarily using FSL software (Jenkinson et al., 2012) and FreeSurfer (Dale et al., 1999).

**Gradient distortion correction (GDC).** GDC was applied within image analysis pipelines using tools developed by FreeSurfer and HCP groups (https://github.com/Washington-University/Pipelines), using the Siemens scanner-specific table of gradient nonlinearities.

**Structural.** Brain extraction was performed in native space after GDC unwarping using FSL BET (Smith, 2002). Tissue-type segmentation was estimated based on FSL FAST (Zhang et al., 2001), providing hard segmentation as well as partial-volume images for each tissue type. This tool was also used to provide a fully bias-field-corrected version of brain extracted structural brain images. Subcortical structures were modeled using FSL FIRST (Patenaude et al., 2011).

**Cortical surface reconstruction.** Subject-specific cortical surface reconstruction and cortical parcellation were estimated based on the GDC, brain-extracted T1 image, using the command recon-all from FreeSurfer (Dale et al., 1999).

**Registration.** Rigid registrations between multimodal MRI native spaces were estimated through FSL FLIRT with boundary-based cost function (Jenkinson et al., 2002; Greve and Fischl, 2009). Nonlinear warps to MNI152 standard-space T1 template were estimated through FSL FNIRT. This set of nonlinear warps is then carried over to all MRI modalities, such as in the case of rs-FMRI.

**EPI distortion correction.** B0 fieldmap processing was estimated through FSL Topup (Andersson et al., 2003) based on AP-PA image pairs from DWI-MRI protocol.

**Functional.** rs-FMRI data were preprocessed using a custom pipeline previously validated on developmental datasets (Baxter et al., 2019; Fitzgibbon et al., 2020). rs-FMRI data were corrected for interval and intravolume subject head motion and EPI distortions (Andersson and Sotiropoulos, 2015); high pass temporal filtering and GDC unwarping were also applied. Registration to structural was improved by an extra rigid registration step aided by a single-band EPI image. Structured artifacts were removed by FSL ICA+FIX processing (Beckmann and Smith, 2004; Griffanti et al., 2013; Salimi-Khorshidi et al., 2014). The FSL FIX classifier was specifically trained for these data and provided the following scores in leave-one subject-out accuracy: true positive ratio (TPR) = 98.8%; true negative ratio (TNR) = 95.3%; weighted ratio (TPR + TNR)/4) = 97.9%. Independent components (separately identified for each individual) were then regressed into the rs-FMRI signal to obtain denoised (clean) rs-FMRI signal, thus minimizing the effect of head-motion, physiological, and MRI-related artifact at the individual subject level. FSL MELODIC was then used to estimate 50 group-average independent components. We then calculated median absolute (ridge) partial correlation (with a regularization value of 0.1) and amplitude for each of the 25 independent components identified as RSNs.

**Diffusion.** DWI-MRI data were first corrected for eddy currents, EPI distortions, and interval and intravolume subject head motion, with outlier-slice replacement, using FSL Eddy (Andersson and Sotiropoulos, 2015). GDP unwarping was then applied (Miller et al., 2016). Diffusion tensor imaging (DTI) fitting was conducted with FSL DTIFIT using a kurtosis model (Behrens et al., 2007). Neurite Orientation Dispersion and Density Imaging (NODDI) modeling was estimated using FSL cuDIMOT based on the Bingham-NODDI model (Tariq et al., 2016). In order to resolve crossing-fiber configurations, multishell voxelwise diffusion was modeled using FSL BedpostX (Behrens et al., 2012). Probabilistic tractography was then conducted with FSL ProbtrackX (Behrens et al., 2007), and 29 major WM bundles were reconstructed as implemented in FSL AutoPtx (de Groot et al., 2013).

**Myelin and iron maps.** Quantitative MRI data were processed to produce the quantitative maps of myelination (1/T1) and iron level (1/T2⁎), using the Voxel-Based Quantification toolbox (Callaghan et al., 2014) in Statistical Parametric Mapping (http://www.fil.ion.ucl.ac.uk/spm/). Although R1 (1/T1, longitudinal relaxation rate) and R2⁎ (1/T2⁎, effective transverse relaxation rate) are not direct quantitative maps of myelination or iron (as other biological factors can also affect them), these quantitative maps have a high degree of sensitivity to myelination and iron (Weiskopf et al., 2013; Callaghan et al., 2014; Lutti et al., 2014).

**Perfusion.** Perfusion images were processed using FSL BASIL (Chappell et al., 2009). Images were first corrected with fieldmap and GDC unwarping; then, to obtain maps of cerebral blood flow and arrival time in absolute units, a calibration step was implemented based on cerebrospinal fluid values.

**Image-derived phenotypes (IDPs)**

Each MRI parameter was summarized in a series of IDPs: anatomy-specific average values that span three sets of ROIs. For cortical and subcortical regions, we used the Desikan-Killiany Atlas (84 parcels, 68 cortical, and 16 subcortical) from the individual FreeSurfer parcellation (Fischl et al., 2002). This parcellation was then warped into each (relevant) modality in order not to interpolate MRI-map values. For ASL, we used this parcellation while opting for a conservative approach to minimize coverage issues in frontal and temporal pole ROIs (voxels size being too large for these thin cortical ribbons). We removed bilaterally the frontal and temporal poles ROIs, thus resulting in 80 ROIs for ASL perfusion and 80 ROIs for ASL arrival time (instead of 84 and 84 ROIs). For the white matter, we used the 29 white matter bundles from the AutoPtx reconstruction; first averaged at group level; optimally thresholded; and then warped back to native spaces. For functional activity, 25 group-level RSNs were identified.

A total of 859 IDPs were then fed into statistical analysis. Functional IDPs (RSNs and ASL IDPs) represented 25% of all IDPs, while WM and GM IDPs represented, respectively, 30% and 45% of all IDPs.

**Cognitive testing and reported mental health and general health measures**

All measures acquired during the testing are reported in detail in Wassenaar et al. (2019).

**Cognitive skills**

Here we considered a summary measure for three tasks of interest: the relational memory task (correct valid answers, %) (Chaddock et al., 2010); task switching (switch cost, ms) (Hillman et al., 2014); and object-location task (identification errors 8 s delay; Table 3) (Pertsov et al., 2012).
Table 4. MRI modalities and MRI metrics used to define IDPs of brain function and structure

<table>
<thead>
<tr>
<th>MRI modality</th>
<th>Metric</th>
<th>ROI type</th>
<th>No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs-fMRI</td>
<td>Functional connectivity</td>
<td>RSN</td>
<td>25</td>
</tr>
<tr>
<td>rs-fMRI</td>
<td>Amplitude</td>
<td>RSN</td>
<td>25</td>
</tr>
<tr>
<td>ASL</td>
<td>Perfusion</td>
<td>GM</td>
<td>80</td>
</tr>
<tr>
<td>ASL</td>
<td>Arrival time</td>
<td>GM</td>
<td>80</td>
</tr>
<tr>
<td>T1w-MRI</td>
<td>Volume (VBM-like)</td>
<td>GM</td>
<td>84</td>
</tr>
<tr>
<td>q-MRI</td>
<td>R1</td>
<td>GM</td>
<td>84</td>
</tr>
<tr>
<td>q-MRI</td>
<td>R2*</td>
<td>GM</td>
<td>84</td>
</tr>
<tr>
<td>T1w-MRI</td>
<td>Surface area</td>
<td>Cortex</td>
<td>68</td>
</tr>
<tr>
<td>T1w-MRI</td>
<td>Surface thickness</td>
<td>Cortex</td>
<td>68</td>
</tr>
<tr>
<td>DW-MRI</td>
<td>DTI - FA</td>
<td>WM</td>
<td>29</td>
</tr>
<tr>
<td>DW-MRI</td>
<td>DTI - MD</td>
<td>WM</td>
<td>29</td>
</tr>
<tr>
<td>DW-MRI</td>
<td>DTI - Kurtosis</td>
<td>WM</td>
<td>29</td>
</tr>
<tr>
<td>q-MRI</td>
<td>R1</td>
<td>WM</td>
<td>29</td>
</tr>
<tr>
<td>q-MRI</td>
<td>R2*</td>
<td>WM</td>
<td>29</td>
</tr>
<tr>
<td>DW-MRI</td>
<td>NODDI f-iso</td>
<td>WM</td>
<td>29</td>
</tr>
<tr>
<td>DW-MRI</td>
<td>NODDI f-isnta</td>
<td>WM</td>
<td>29</td>
</tr>
<tr>
<td>DW-MRI</td>
<td>NODDI DA</td>
<td>WM</td>
<td>29</td>
</tr>
<tr>
<td>DW-MRI</td>
<td>NODDI ODN</td>
<td>WM</td>
<td>29</td>
</tr>
</tbody>
</table>

Five MRI sequences were used to quantify 18 different MRI metrics. Specific sets of ROIs were then used for each MRI metric to extract whole-brain MM IDPs quantifying brain structure, microstructure, function, myelin content, and blood perfusion.

Mental health
Mental health was assessed with the Strengths and Difficulties Questionnaire (Goodman, 1997).

Questionnaire on general health
From the Behavior Health in School-aged Children questionnaire (World Health Organization, 2016), we used the positive health items (self-rated health, life satisfaction, multiple health complaints) to measure reported general health.

Experimental design and statistical analyses
This is a cross-sectional study with a sample size of N = 50 subjects. Because of the limited sample size compared with the number of variables of interest, we strove to reduce input data and nuisances variables dimensions as much as possible. Standardization of variables before decomposition methods (principal component analysis [PCA] and canonical correlation analysis [CCA]) was applied to avoid variables with disproportionately greater variance driving the decomposition. All statistical analyses were conducted in MATLAB 2018.

Confounds
Before all statistical analyses, a series of relevant confounds was chosen: age; sex; pubertal developmental level (assessed through the Pubertal Development Rating Scale) (Petersen et al., 1988), a self-report measure of physical development for youths under the age of 16; socioeconomic status (assessed through the United Kingdom Index of Multiple Deprivation); and head size/scaling factor (computed through FSL SIENAX).

On these nuisance variables, we perform a dimensionality reduction through means of PCA (Nz = 2) accounting for 60% of total variance. These confounds were then regressed out of all IDPs and behavioral variables and the residuals standardized.

Dimensionality reduction of IDPs and physical variables
In order to avoid an overdetermined, rank-deficient CCA solution, and to limit the chances of overfitting, a dimensionality reduction step was performed to both IDPs and physical variables. Using the same approach previously applied by Smith et al. (2015), IDPs were reduced into 10 PCAs (Nx = 10; variance explained = 53%), whereas physical variables were reduced into 5 PCAs (Ny = 5; variance explained = 79%).

CCA
We sought to characterize a mode of brain–physical covariation across pupils: a data-driven latent factor linking a linear combination of neuroimaging metrics (Table 4) with a linear combination of physical measures (Table 5). To this end, we used CCA, an approach that has successfully been applied in recent studies and that, compared with pairwise association testing, has shown greater sensitivity for complex biological processes and greater explained variance (Smith et al., 2015; Miller et al., 2016).

CCA is a symmetric, cross-decomposition method that characterizes covariation modes between a pair of two-dimensional datasets. This is achieved by finding two sets of free parameters (or canonical coefficients, i.e., one set of coefficient vectors per set of brain metrics and one set of coefficient vectors per set of physical metrics) that maximize the correlation of the projections of the two datasets into the identified latent space (or canonical variates or subject scores). In other words, the variation in mode strength between subjects is maximally correlated. Here, this was computed using MATLAB ‘canoncorr’ function.

Unbiased statistical inference through block-aware permutation testing
Deconfounding, as required to ensure that the CCA is not driven by nuisance factors, induces a dependency among the rows of the data submitted to CCA. While this dependency is weak and diminishes with increasing sample size, it represents a violation of the exchangeability assumption required by permutation, which can inflate permutation significance. To account for this deconfounding-induced dependency that violates exchangeability, we use a method that, without changing the canonical correlations, reduces the data from N observations to N-Nz observations that are exchangeable and, thus, can be subjected to a permutation test (Theil, 1965; Winkler et al., 2020). We randomly chose 1000 sets of Nz rows for removal, conducting 1000 permutations for each set.

Permutations were performed among subjects within school, respecting dependencies given by the hierarchical structure of the data (Winkler et al., 2015). For each of the 1000 repetitions, a p value was computed based on this null distribution for the first CCA mode. Across repetitions, a distribution of statistical significance values was built and the final statistical significance level was computed as its average value. The results of this analysis are shown in Figure 2b.

Unbiased estimation of effect size through leave-one-school-out cross-validation
In order to derive an unbiased estimate of the CCA correlation strength that took into account the hierarchical structure in the data, we implemented a leave-one-school-out cross-validation (CV) approach. In all but one school, we performed all the above steps (except permutation testing), learning all the coefficients of the standardization steps and of the linear transformations. On the left-out school, we then applied those transformations and predicted left-out pupils’ scores in the CCA mode. We repeated this procedure for all folds (here schools). CV performance
For each number of PCs, the whole CV pipeline was performed: CV varying the number of PCs (from \( N_x = 6 \) to 14).

Supplementary analysis for robustness of identified covariation when varying the number of principal components (PCs) on the IDPs was then quantified as the Pearson’s \( \rho \) correlation coefficient and mean squared error (MSE) calculated between predicted brain and physical canonical covariates (or predicted canonical variates). The results of this analysis are shown in Figure 2a.

Supplementary analysis for robustness of identified covariation when varying the number of PCs number (\( N_x = 10 \)). The results of this analysis are shown in Figure 2c–f.

Characterization of brain and physical phenotypes

We then aimed to characterize the CCA phenotypes: the set of brain measures and the set of physical measures symmetrically linked by the CCA covariance mode. To do this (formally, to characterize the CCA crossed loadings, hereafter referred to as loadings), we follow the procedure described by Smith et al. (2015). On the whole sample, CCA brain loadings were calculated as the pairwise Pearson’s partial correlation between CCA physical variate (or subject scores) and the original datasets of brain IDPs, while controlling for the full set of nuisance variables: CCA brain loadings = partial correlation (brain IDPs, CCA physical variate, nuisance variables). The results of this process are shown in Figures 4–6. CCA physical loadings were calculated with the following the same process: CCA physical loadings = partial correlation (physical variables, CCA brain variate, nuisance variables). The results of this process (only for structural IDPs) are shown in Figure 3. CCA loadings are therefore bound between 1 and −1.

For functional measures, each IDP is represented by a whole-brain RSN. To aid interpretation, for each RSN, its CCA brain loading was multiplied by the group RSN map. The results of this process are shown in Figure 5a, b. Then, to derive a summary representation, we concatenated all RSN maps in a 4D file and computed standardized mean across RSNs, separately for both functional connectivity and amplitude. The results of this process are shown in Figure 5c, d.

We then aimed to characterize the average involvement for each type of MRI value. Across IDPs of a MRI metric, we computed the average across CCA brain loadings. This provided a ranked list of MRI parameters representing the average form of the CCA metric with pupils’ physical scores. The results of this process are shown in Figure 7.

Joint-inferences with univariate measures of cognitive skills, mental health, and general health

The tests for association between the identified CCA mode and the multiple variables measuring the domains of cognitive skills, mental health, and general health were conducted (separately for each domain) using multiple linear regression with nonparametric combination (NPC) implemented in FSL PALM (Winkler et al., 2016). NPC works by combining test statistics or \( p \) values of separate (even if not independent) analyses into a single, joint statistic, the significance of which is assessed through synchronized permutations for each of the separate tests. Here we asked whether each CCA covariate (brain covariate while adjusting for physical covariate, and vice versa) was associated with any domain of interest, and the NPC was tested via Fisher statistic with 1000 block-aware permutations while adjusting for nuisance variables in reduced space. For each domain, NPC Fisher significance values were corrected for multiple comparison
testing across CCA covariates (brain and physical) via family-wise error correction (FWE-corr).

**Results**

**Brain–physical mode of covariation across pupils**

Using CCA, we tested the hypothesis that, across pupils, intersubject differences in multimodal whole-brain IDPs covaried significantly with differences in physical lifestyle variables, independent of nuisance variables. We found one significant mode of brain–physical covariation across pupils, linking differences in brain IDPs with individual differences in physical lifestyle (Fig. 2a, CCA: $\rho = 0.34$, MSE = 1.38, using leave-one school-out CV; Fig. 2b, $p = 0.0130$, significance assessed on 1000 repetitions, each with 1000 block-aware permutations; results remained the same if adjusted for the full set of nuisance variables, $p = 0.0312$, significance assessed on 1000 repetitions). We also show that varying the number of PCs on the IDPs (from $N_x = 6$ PCs to 14 PCs) consistently produces the same results (Fig. 2c–f), showing robustness of the identified relationship across a range of PCs on the IDPs. This mode represents a pattern of brain IDPs that covaries with a pattern of physical variables. We next interrogated this physical phenotype and brain phenotype separately, to determine the patterns that underlie this mode.

**Physical phenotype of covariation**

For each physical variable, we calculated the loadings of the physical phenotype relating to the CCA mode (or CCA physical loadings) representing the relationship between each physical variable and the CCA brain variate (or subjects’ brain scores) (Fig. 3). We found that pupils who scored higher in the brain–physical mode of covariation were those with higher cardiovascular fitness; those with lower body mass index; those who spent more time doing long-lasting (both moderate and vigorous) physical activity during a normal school week and spent less time being sedentary.

**Brain phenotypes of covariation**

In order to interpret brain phenotypes of physical covariation, we calculated the canonical loadings for each IDP of brain structure, microstructure, and function. These loadings (or CCA brain loadings) represent the relationship between each brain IDP and the CCA physical variate (or subjects’ physical scores) (to explore the spatial patterns of all structural, microstructural, and functional IDPs, see, respectively, Fig. 4 and Fig. 5a,b; for violin plot of CCA brain loadings for all IDPs, see Fig. 6).

We observed that some MRI metrics presented a global and homogeneous involvement in the mode of covariation across ROIs. In order to quantify this tendency, for each MRI metric, we computed the average CCA brain loadings across all ROIs (Fig. 7). We found that the strongest CCA brain loadings were found for GM perfusion (and arrival time) as well as cortical surface area, GM volume, and a number of WM diffusion metrics. MRI metrics with the greatest average CCA brain loadings tended to be characterized by spatially extended and homogeneous involvement across the whole brain. Together, these results show that pupils with greater physical scores were those who also showed global patterns of higher blood perfusion (and lower arrival time, i.e., faster perfusion) in the GM, greater GM volume, greater cortical surface area, greater neurite dispersion anisotropy across WM tracts, as well as greater extraneurite fraction (equivalent to lower intraneurite fraction), and lower neurite orientation dispersion.

We also observed that, although the average CCA brain loadings for RSNs functional connectivity and BOLD amplitude was close to zero, there was great variance across RSNs (Fig. 5). Because RSNs are not binary masks but are instead characterized by spatial distributions, to summarize their pattern of involvement in the mode of covariation, for each voxel we computed the standardized mean CCA brain loadings across RSNs (Fig. 5c, d). The resulting maps for functional connectivity (Fig. 5c) and amplitude (Fig. 5d) showed both similarities and differences in their patterns of involvement in the mode of covariation. In Figure 5c, RSN functional connectivity shows greater positive involvement bilaterally in the parietal cortices, supplementary motor cortex, putamen, and right primary motor cortex, whereas it shows greater negative involvement broadly in the occipital cortices. The peak of positive involvement was localized in the right parietal cortex, whereas the negative involvement was localized in the occipital cortex. In Figure 5d, RSN BOLD amplitude shows greater positive involvement in the anterior cingulate gyrus (dACC), superior frontal gyrus, parietal cortices, right inferior frontal gyrus, whereas it shows greater negative involvement broadly in the occipital cortices, and left primary somatosensory cortex. The peak of positive involvement was localized in the dACC, whereas the negative involvement was localized in the occipital cortex. These maps show a common pattern of greater positive involvement bilaterally in the parietal cortices, and a common pattern of negative involvement in the occipital cortices.

**Relationship with measures of cognition, mental health, and general health**

We then tested the hypothesis that the identified CCA mode of brain–physical covariation was significantly associated with measures of (1) cognitive skills, (2) mental health, and (3) general reported health. We used a multiple linear regression to test the association between the CCA variates (brain covariate while adjusting for physical covariate, and vice versa) and the outcome measures. Testing an NPC joint-inference for each domain, we found no statistically significant association with cognitive skills (respectively, for brain covariate and physical covariate, NPC Fisher FWE-corr $p = 0.2600, 0.2360$) or general reported health (respectively, for brain covariate and physical covariate, FWE-corr $p = 0.7240, 0.4910$). We found a trend toward an association between individual differences in the brain covariate and differences in mental health (NPC Fisher FWE-corr $p = 0.0640$), whereas no significant association was found for the physical covariate (NPC Fisher FWE-corr $p = 0.1830$).
In this work, we show that, in 12-year-old pupils, physical activity, fitness, and physical health are linked with global patterns of brain structure, microstructure, and function. In this relationship, whole-brain, homogeneous patterns of multimodal brain phenotypes are linked with a specific, latent pattern of physical measures that capture a physically active lifestyle (high fit, high active, low sedentary individuals). This finding hints at the involvement of multiple underlying biological processes and suggests that physical activity, fitness, and health are associated with specific brain structural and functional patterns. The figure demonstrates the relationship between the identified phenotype of active lifestyle and various MRI metrics for different ROIs. Hot colors represent a positive relationship with the physical phenotype, while cold colors represent a negative relationship. Structural maps are ranked from top to bottom (left column to right) in accordance to average CCA brain loadings (as shown in Fig. 7).

Discussion
In this work, we show that, in 12-year-old pupils, physical activity, fitness, and physical health are linked with global patterns of brain structure, microstructure, and function. In this relationship, whole-brain, homogeneous patterns of multimodal brain phenotypes are linked with a specific, latent pattern of physical measures that capture a physically active lifestyle (high fit, high active, low sedentary individuals). This finding hints at the involvement of multiple underlying biological processes and suggests that physical activity, fitness, and health are associated with specific brain structural and functional patterns. The figure demonstrates the relationship between the identified phenotype of active lifestyle and various MRI metrics for different ROIs. Hot colors represent a positive relationship with the physical phenotype, while cold colors represent a negative relationship. Structural maps are ranked from top to bottom (left column to right) in accordance to average CCA brain loadings (as shown in Fig. 7).
health and aerobic exercise might have a wider effect on brain processes than previously thought.

We applied a holistic approach to provide novel insight into the importance of different aspects of a physically active lifestyle in relation to brain structure and function. While high cardiovascular fitness and physical activity are positively linked with the identified brain phenotypes, sedentary activity and body mass index are negatively related. Furthermore, we showed that long-lasting physical activity, either moderate or vigorous, is more important to this relationship than brief bursts of activity, suggesting that regular moderate-to-vigorous physical activity might be a better driver to promote brain changes. Together, these findings situate pupils along a latent axis according to their physical phenotype: pupils with high cardiorespiratory fitness
and performance and with high weekly levels of physical activity, contrast with pupils spending most time in sedentary or low-energy behaviors.

The novelty of this work is the finding of multimodal global brain phenotypes linked with a physically active lifestyle. Although prior work has studied the relationship between single brain phenotypes linked with a physically active lifestyle.

Previous literature has explicitly focused on studying the effects of aerobic exercise on the hippocampus (Cotman, 2002; Pereira et al., 2007; Chaddock-Heyman et al., 2016; Thomas et al., 2016). Cardiorespiratory fitness is indeed known to promote hippocampal neurogenesis and angiogenesis that, in turn, determines macroscale changes that are also visible via noninvasive neuroimaging (van Praag et al., 1999). Here, we extend the current knowledge beyond a uniquely hippocampal pattern, highlighting the global nature of greater volume and faster perfusion across the whole-brain GM. In other words, variation in hippocampal structure alone does not underlie the brain–physical relationship characterized here. Rather, we observed homogeneous loadings across GM areas. The strongest contributions to our brain phenotype came from perfusion measures. These robust associations found with perfusion metrics are in line with a body of literature showing positive effects of physically active lifestyle on vascular health (Department of Health and Human Services Office of Disease Prevention and Health Promotion, 2000; Vaynman et al., 2004), as well as animal studies linking physical exercise to angiogenesis (Kleim et al., 1996; Rhyu et al., 2010).

Further key insights derive from spatially extended patterns of WM covariation. Although the myelin–sensitive metrics (Quantitative-MRI) in the current study made little contribution to the mode of variation, higher scores on the physical phenotype were associated with lower intraneurite density and kurtosis and, to a lesser extent, with lower neurite orientation dispersion and with greater dispersion anisotropy. It is relevant that the DW-MRI protocol used in this study would be sensitive to diffusion properties within large glial cells, such as astrocytes and oligodendrocytes. Our gradient strength is indeed known to promote hippocampal neurogenesis and angiogenesis that, in turn, determines macroscale changes that are also visible via noninvasive neuroimaging (van Praag et al., 1999). Here, we extend the current knowledge beyond a uniquely hippocampal pattern, highlighting the global nature of greater volume and faster perfusion across the whole-brain GM. In other words, variation in hippocampal structure alone does not underlie the brain–physical relationship characterized here. Rather, we observed homogeneous loadings across GM areas. The strongest contributions to our brain phenotype came from perfusion measures. These robust associations found with perfusion metrics are in line with a body of literature showing positive effects of physically active lifestyle on vascular health (Department of Health and Human Services Office of Disease Prevention and Health Promotion, 2000; Vaynman et al., 2004), as well as animal studies linking physical exercise to angiogenesis (Kleim et al., 1996; Rhyu et al., 2010).
association here observed between physical lifestyle scores and WM extraneurite fraction (by construction 1 minus intraneurite density, and specifically, the hindered space outside the neurites prescribed through anisotropic diffusion). Crucially, there exists key histologic evidence from animal studies in support of an association here observed between physical lifestyle scores and macroglia cell density across multiple WM tracts, perhaps reflecting a role in providing enhanced metabolic support for neurons. This hypothesis should be tested using imaging alongside more direct measures from ex vivo studies, or using alternative techniques with greater specificity, such as detecting MRS-visible metabolites with greater sensitivity to astrocytes (Brand et al., 1993).

We also report two patterns of RSNs involvement in the mode of brain–physical covariation. We found that a physically active lifestyle was linked with greater connectivity in the parietal cortices and with lower connectivity in the occipital cortices, showing, respectively, increased and decreased BOLD coupling with all RSNs in more active participants. The same phenotype of a physically active lifestyle was also positively related with greater amplitude in local BOLD fluctuations in the dACC and in the parietal cortices, and with lower amplitude in the occipital cortices. Studying both RSNs amplitude (BOLD variance) and functional connectivity (BOLD covariance) can be important to understand possible sources of change and the related neural processes (Garrett et al., 2010; Duff et al., 2018). While greater activity both in functional connectivity and in BOLD amplitude may suggest greater coactivation between the parietal cortices and multiple RSNs across the whole brain, greater BOLD amplitude with no increase in functional connectivity, as observed in the dACC, may suggest greater local activity that results in a decoupling of the dACC from the rest of brain activity. Greater dACC activity during a cognitive control task was previously associated with higher fitness levels in preadolescent children, with greater dACC activity in the high fitness group positively related to accuracy in task performance (Voss et al., 2011). In this study, however, we found no significant association between pupil’s scores in the mode of brain–physical covariation and differences in cognitive skills. Only a trend for an association with mental health was found, thus not allowing us to infer on the cognitive or mental health relevance of this brain pattern.

Overall, our findings lend support to the growing body of evidence demonstrating a close relationship between the body and the brain. Although the relatively small sample size given the number of variables of interest and the possible cluster effect of schools may represent a limitation of this study, here we used thorough statistical procedures (i.e., block-aware permutation testing and leave-one-cluster-out cross-validation) to explicitly deal with this factor, thus producing robust and unbiased statistics. Larger samples might provide power to detect multiple modes of covariation. Also, the results here reported are correlational; therefore, caution is required in interpreting directionality. It is possible that this relationship also represents the other direction of association, and indeed brain structure, microstructure, and function are key drivers of behavioral choices. Nevertheless, our findings suggest that a complex physical phenotype that is influenced by physiology, and lifestyle choices, might have widespread effects on biological processes influencing brain phenotypes. Future studies may test whether improving physical health and fitness through means of activity interventions promotes diffuse neuroplasticity.

In conclusion, this work provides novel insight into the comprehensive relationship between physically active lifestyle and brain structure and physiology in early adolescence. These findings have broad implications for future research, suggesting novel avenues to study the effect of modifiable lifestyle factors as part of wider brain-body relationships. Understanding how physical pathways may foster healthy human brain development can help us to develop better intervention studies aimed at informing public health and education policies.

References


