NeuroImage 144 (2017) 262–269

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

The Cambridge Centre for Ageing and Neuroscience (Cam-CAN) data repository: Structural and functional MRI, MEG, and cognitive data from a cross-sectional adult lifespan sample

Jason R. Taylor [a,*], Nitin Williams [b], Rhodri Cusack [c], Tibor Auer [b], Meredith A. Shafto [d], Marie Dixon [d], Lorraine K. Tyler [d], Cam-CAN [e], Richard N. Henson [b]

[a] School of Psychological Sciences, The University of Manchester, Zochonis Building, Brunswick Street, Manchester M13 9PL, UK
[b] MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 7EF, UK
[c] Brain and Mind Institute, University of Western Ontario, London, ON, Canada
[d] Department of Psychology, University of Cambridge, Cambridge CB2 3EB, UK
[e] Cambridge Centre for Ageing and Neuroscience (Cam-CAN), University of Cambridge and MRC Cognition and Brain Sciences Unit, Cambridge, UK

ARTICLE INFO

Article history:
Accepted 8 September 2015
Available online 12 September 2015

Keywords:
Data repository
Brain imaging
Magnetic resonance imaging
Magnetoencephalography
Cognition
Ageing

ABSTRACT

This paper describes the data repository for the Cambridge Centre for Ageing and Neuroscience (Cam-CAN) initial study cohort. The Cam-CAN Stage 2 repository contains multi-modal (MRI, MEG, and cognitive-behavioural) data from a large (approximately N = 700), cross-sectional adult lifespan (18–87 years old) population-based sample. The study is designed to characterise age-related changes in cognition and brain structure and function, and to uncover the neurocognitive mechanisms that support healthy cognitive ageing. The database contains raw and preprocessed structural MRI, functional MRI (active tasks and resting state), and MEG data (active tasks and resting state), as well as derived scores from cognitive behavioural experiments spanning five broad domains (attention, emotion, action, language, and memory), and demographic and neuropsychological data. The dataset thus provides a depth of neurocognitive phenotyping that is currently unparalleled, enabling integrative analyses of age-related changes in brain structure, brain function, and cognition, and providing a testbed for novel analyses of multi-modal neuroimaging data.

1. Cam-CAN project

1.1. Overview

The Cambridge Centre for Ageing and Neuroscience (Cam-CAN) Stage 2 cohort study is a large-scale (approx. N = 700), multi-modal (MRI, MEG, and behavioural), cross-sectional, population-based adult lifespan (18–87 years old) investigation of the neural underpinnings of successful cognitive ageing. The project is an interdisciplinary collaboration involving researchers with expertise in cognitive psychology, cognitive neuroscience, psychiatry, engineering, and public health. The full Cam-CAN study consists of three stages, described briefly below (for full protocol, see Shafto et al., 2014). The focus of the present paper is the Stage 2 dataset, which includes raw and preprocessed MRI, MEG, and cognitive-behavioural data, along with demographic data and other cross-referenced measures collected from the Stage 2 cohort in Stage 1. Collection of Stage 3 data has recently completed, and these data will be added to the database in due course.

A key focus of the Cam-CAN project is integrative analysis across domains of cognition and measures of neural structure, function, and connectivity, with the goal of understanding how neurocognitive systems adapt in order to overcome age-related changes. For example, Tsvetanov et al. (2015) combined T2*-weighted resting-state fMRI data with resting-state MEG data to show that ageing affects the vascular response independently of neural activity; Geerligs et al. (in press) showed how age-related differences in functional connectivity change with cognitive state; and Kievit et al. (2014) combined fractional-

Abbreviations: ACE-R, Addenbrooke's Cognitive Exam; BOLD, blood-oxygenation level-dependent; Cam-CAN, Cambridge Centre for Ageing and Neuroscience; DARTEL, diffeomorphic anatomical registration through exponentiated lie algebra; DWI, diffusion-weighted imaging; DTI, diffusion tensor imaging; DKI, diffusion kurtosis imaging; ERF, event-related field; fMRI, [functional] magnetic resonance imaging; GM, grey matter; MNI, Montreal Neurological Institute; MT, magnetisation transfer; MEG, magnetoencephalography; MMSE, Mini Mental Status Exam; ROI, region of interest; VBM, voxel-based morphometry; WMS-III UK, Wechsler Memory Scale Third UK Edition; WM, white matter.

* Corresponding author. Fax: +44 161 306 0400.
E-mail addresses: jason.taylor@manchester.ac.uk (J.R. Taylor), nitin.williams@mrc-cbu.cam.ac.uk (N. Williams), rhodri@cusacklab.org (R. Cusack), tibor.auer@mrc-cbu.cam.ac.uk (T. Auer), mshafto@csl.psychol.cam.ac.uk (M.A. Shafto), admin@cam-can.com (M. Dixon), lktyler@csl.psychol.cam.ac.uk (L.K. Tyler), rik.henson@mrc-cbu.cam.ac.uk (R.N. Henson).
URL: http://www.cam-can.com (Cam-CAN).

http://dx.doi.org/10.1016/j.neuroimage.2015.09.018
1053-8119/

anisotropy measures from the diffusion-weighted (DWI) MRI data with volumetric measures from the T1-weighted MRI data to show that white matter (WM) and grey matter (GM) in frontal cortex make independent contributions to age-related declines in fluid intelligence and multitasking.

Several features of the Cam-CAN Stage 2 dataset together make it unique. First, the sample is derived from a larger, population-based sample (approximate N = 3000) recruited from the general population via Primary Care Trust lists, which can be related to national data (Shafto et al., 2014), and which allows quantification of bias in the Stage 2 sample of people willing and able to undergo neuroimaging. Second, the distribution of ages was selected to be roughly uniform, allowing sufficient statistical power to test for differences within as well as across age groups. Third, Stage 2 of the study involved collecting a broad range of behavioural measures from 14 experiments spanning five main cognitive domains (attention, emotion, action, language, and memory), which can also be related to considerable amounts of demographic, health and lifestyle data obtained in Stage 1. Finally, Stage 2 includes a wide range of neuroimaging measures: high-resolution (1 mm^3) T1- and T2-weighted images, diffusion-weighted images (DWI), magnetisation-transfer (MT) images, and BOLD EPI images during rest, a sensorimotor task and movie-watching, as well as MEG data during rest and the same sensorimotor task. The depth of this neurocognitive phenotyping is currently unparalleled, and provides a testing ground for the development of new multimodal analysis methods.

In the following section, we briefly describe the three Cam-CAN data-collection stages. Acquisition and analysis of the MRI and MEG data in Stage 2 are described more fully in Section 2.

1.2. Data-collection stages

1.2.1. Stage 1

In Stage 1[1], 2681 participants were interviewed in their homes to acquire demographic information; measures of cognitive, mental and physical health; and lifestyle information. Tests of vision, hearing, balance, and speeded response times were administered, and participants completed detailed self-report questionnaires about their physical activity and life experiences. Neuropsychological tests included general cognitive assessments (MMSE, Folstein et al., 1975; ACE-R, Mioshi et al., 2006a, 2006b), tests of memory (logical memory from WMS-III UK, Wechsler, 1999), and verbal intelligence (Spot the Word, Baddeley et al., 1993).

Measures taken in Stage 1 additionally served to screen participants for participation in Stage 2: To continue, participants were required to be willing to continue, be cognitively healthy (MMSE > 24), to meet hearing, vision, and English language ability criteria necessary for completing experimental tasks, and to be free of MRI or MEG contraindications and neurological or serious psychiatric conditions.

1.2.2. Stage 2

In Stage 2, participants (target N = 700: 50 men, 50 women from each age decade) were recruited to attend testing sessions at the Medical Research Council (UK) Cognition and Brain Sciences Unit (MRC-CBSU) in Cambridge, UK. Owing to recruitment problems for the youngest decade (18–27), only 56 (27 men) were tested from this decade. In this stage, structural and functional MRI scans, MEG recordings, and cognitive task data were collected over three separate sessions. Structural and functional MRI scans collected in Stage 2 are listed in Tables 1 and 2; MEG recordings are listed in Table 3; and cognitive behavioural tasks are listed in Table 4. Physiological data (height, weight, and blood pressure) were also collected, and a saliva sample was taken for future genetic analysis. As the data from Stage 2 are the focus of the repository described in this paper, these data and derived measures are described in more detail in Section 2.

[1] In the data repository, Stages 1, 2, and 3 are referred to as CC3000, CC700, and CC280, respectively.

1.2.3. Stage 3

In Stage 3, a subset of participants (target N = 280: 20 men, 20 women from each decade) were recruited to attend further MRI and MEG sessions within 3 years of their assessment in Stage 2. Over three sessions, structural MRI and physiological measures were collected, along with fMRI and MEG data on a variety of cognitive tasks. Structural MRI scans (all participants) included a repeat T1-weighted structural image, as well as T2-weighted FLAIR, and arterial spin labelling (ASL). Functional MRI tasks (target N = 140 each) investigated emotion regulation, emotional memory, fluid intelligence, picture naming, response selection and inhibition, sentence comprehension, and visual short-term memory; repeat resting-state data and field maps for distortion correction were also collected. Height, weight, and blood pressure, which were measured in Stage 2, were re-measured at the Stage 3 MRI session. MEG tasks investigated incidental memory, oddball processing, picture naming, response selection and inhibition, sentence comprehension, and word recognition, as well as a resting state. Hearing, vision, and cognitive status (MMSE), which were also measured in Stage 1, were re-evaluated in the MEG session.

2. Database details

2.1. Purpose

Using semi-automated Matlab and Linux shell scripts, raw data were pulled from various sources (testing laptops, MRI and MEG data servers) into a central location. Once there, further automated scripts identified new raw data and submitted them to the appropriate processing scripts. Behavioural data were analysed by custom Matlab scripts; MRI and MEG data were processed using Automatic Analysis (aa 4.2; Cusack et al., 2014) pipelines and modules which called relevant functions from neuroimaging analysis software and toolboxes (SPM12, Wellcome Department of Imaging Neuroscience, London, UK; FSL, Smith et al., 2004; Freesurfer, Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Boston, MA, USA; in-house code).

The remit of the database is limited to the Cam-CAN initial study cohort project. Therefore, upon completion of data collection and preprocessing analyses, the repository will effectively be static (i.e., no further raw data will be added). However, as existing pipelines are improved, new data releases are made, and as new pipelines are developed, new processed data will become available. Further, any future studies that use the same cohort and adopt a similar data-sharing policy will likely be incorporated into the database (e.g., there are plans to re-test the Stage 2 cohort on a subset of the same cognitive and neuroimaging measures in the future, to provide longitudinal data).

2.2. Contents

The Stage 2 repository contains MRI, MEG, and behavioural data from 656 participants aged 18–87 years old. All data are labelled with unique project IDs. Data were quality-control checked by semi-automated scripts monitored by the Cam-CAN methods team. All analysis scripts (including aa modules and recipes) are stored in the repository and can therefore be viewed by any user. Matlab scripts are also available to query the repository and compile data cross-referenced by participant identifiers.

2.2.1. MRI

2.2.1.1. MRI data collection. All MRI datasets were collected at a single site (MRC-CBSU) using a 3 T Siemens TIM Trio scanner with a 32-channel head coil. Participants were scanned in a single 1-hour session. Before scanning, physiological measurements were taken, and two behavioural experiments were run. In the scanner, memory foam cushions were used for comfort and to minimise head movement. Instructions and visual stimuli for functional tasks were back-projected onto a screen