

# BMJ Open Protocol for Multi-modality MEDical imaging sTudy bAsed on KaiLuan Study (META-KLS): rationale, design and database building

Jing Sun,<sup>1</sup> Ying Hui,<sup>1</sup> Jing Li,<sup>1</sup> Xinyu Zhao,<sup>2</sup> Qian Chen,<sup>1</sup> Xiaoshuai Li,<sup>1</sup> Ning Wu,<sup>3</sup> Mingze Xu,<sup>4,5</sup> Wenjuan Liu ,<sup>1</sup> Rui Li,<sup>1</sup> Pengfei Zhao,<sup>1</sup> YunTao Wu,<sup>6</sup> Aijun Xing,<sup>6</sup> Huijing Shi,<sup>7</sup> Shun Zhang,<sup>8</sup> Xiaoliang Liang,<sup>8</sup> Yongxin Wang,<sup>9</sup> Han Lv ,<sup>1</sup> Shouling Wu,<sup>6</sup> Zhenchang Wang<sup>1</sup>

**To cite:** Sun J, Hui Y, Li J, *et al.* Protocol for Multi-modality MEDical imaging sTudy bAsed on KaiLuan Study (META-KLS): rationale, design and database building. *BMJ Open* 2023;**13**:e067283. doi:10.1136/bmjopen-2022-067283

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2022-067283>).

JS, YH and JL contributed equally.

Received 09 August 2022  
Accepted 01 February 2023



© Author(s) (or their employer(s)) 2023. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

## Correspondence to

Dr Han Lv;  
[chrislvhan@126.com](mailto:chrislvhan@126.com),  
Professor Shouling Wu;  
[drwusl@163.com](mailto:drwusl@163.com) and  
Dr Zhenchang Wang;  
[cjr.wzhch@vip.163.com](mailto:cjr.wzhch@vip.163.com)

## ABSTRACT

**Introduction** Multi-modality medical imaging study, especially brain MRI, greatly facilitates the research on subclinical brain disease. However, there is still a lack of such studies with a wider age span of participants. The Multi-modality MEDical imaging sTudy bAsed on KaiLuan Study (META-KLS) was designed to address this issue with a large sample size population.

**Methods and analysis** We aim to enrol at least 1000 subjects in META-KLS. All the participants without contraindications will perform multi-modality medical imaging, including brain MRI, retinal fundus photograph, fundus optical coherence tomography (OCT) and ultrasonography of the internal carotid artery (ICA) every 2–4 years. The acquired medical imaging will be further processed with a standardised and validated workflow. The clinical data at baseline and follow-up will be collected from the KaiLuan Study. The associations between multiple risk factors and subclinical brain disease are able to be fully investigated. Researches based on META-KLS will provide a series of state-of-the-art evidence for the prevention of neurological diseases and common chronic diseases.

**Ethics and dissemination** The Kailuan Study and META-KLS have been approved by the Medical Ethics Committee of Kailuan General Hospital (IRB number: 2008 No. 1 and 2021002, respectively). Written informed consent will be acquired from each participant. Results are expected to be published in professional peer-reviewed journals beginning in 2023.

**Trial registration number** NCT05453877.

## INTRODUCTION

Subclinical brain disease is a major concern in general population<sup>1,2</sup> as well as in patients with high blood pressure,<sup>3</sup> diabetes<sup>4</sup> and cerebral small vessel disease (CSVD).<sup>5</sup> Brain imaging study, especially brain MRI, is promising in research on subclinical brain disease. Brain MRI has been collected in several study cohorts, such as the Rotterdam Scan Study<sup>6</sup> and the PolyvasculaR Evaluation for

## STRENGTHS AND LIMITATIONS OF THIS STUDY

- ⇒ The META-KLS (Multi-modality MEDical imaging sTudy bAsed on KaiLuan Study) will include a broad age range of participants, enabling the detection of features of subclinical brain disease.
- ⇒ Multi-modality medical imaging may facilitate research on prevention and early detection of neurological diseases and common chronic diseases.
- ⇒ The homogeneity of study population reduces potential confounding effects.
- ⇒ The cumulative exposure and trajectory of risk factors can be assessed.
- ⇒ The findings from one community population in China may not be generalisable to other populations.

Cognitive Impairment and vaScular Events (PRECISE) study,<sup>7</sup> in order to analyse neurological disease or cardiovascular/cerebrovascular disease and cognitive function in elder population with ages over 50 years. Nevertheless, more researches are still needed in population-based cohort study with a wider range of age.

## KaiLuan Study: a population-based cohort study with a wide range of age

The KaiLuan Study is an ongoing prospective occupational cohort study conducted in the Kailuan community of Tangshan in Northern China, which is a large, modern industrial city southeast of Beijing.<sup>8,9</sup> To date, more than 170 000 adults aged 18–98 years have been enrolled since 2006.<sup>8</sup> Demographic questionnaires and clinical and laboratory examinations were prospectively collected every 2 years from 11 local hospitals.<sup>10,11</sup> This cohort was designed to explore the risk factors for the development or progression of common diseases, including cardiovascular/cerebrovascular disease, diabetes, metabolic

syndrome, etc.<sup>12 13</sup> Ageing, life style, nutrition, etc, are also hot topics in the KaiLuan Study.<sup>14 15</sup> Based on this population-based cohort, multiple factors related to subclinical brain disease in subjects with a wide range of age are able to be analysed.

### Multi-modality Medical imaging sTudy bAsed on KaiLuan Study (META-KLS)

Pathological changes of brain that are related to subclinical brain disease are already present years before clinical significance or being visualised on medical imaging, even in the younger age.<sup>16</sup> State-of-the-art multi-modality medical imaging in a large sample size of population with a wide range of age is essential for clinicians and researchers. Aware of this potential benefit, since December 2020, we have been retrospectively collected the clinical data and will also prospectively enrol the subjects with multi-modality medical imaging based on the KaiLuan Study. Theoretically, there is no limitation on the number of participants that enrolled in META-KLS. All of the participants of the KaiLuan Study without contraindications are able to be recruited in this cohort. According to the fact, we aim to enrol at least 1000 subjects in this cohort.

All of the enrolled participants will be asked to perform a multi-modality medical imaging, including brain MRI scan, retinal fundus photograph, fundus optical coherence tomography (OCT) and ultrasonography of the internal carotid artery (ICA). Imaging collection will be performed every 2–4 years as follow-ups up to 30 years. Meanwhile, the clinical data of the enrolled subjects can also be collected by extracting data from the KaiLuan Study. As such, risk factors and features about subclinical brain disease are able to be fully investigated based on this cohort.

The purpose of this protocol was to give a general outline of META-KLS.

## METHODS AND ANALYSIS

### Study population

We aim to enrol at least 1000 subjects in the META-KLS.

### Inclusion criteria

Inclusion criteria are as follows: (1) the source population of the META-KLS originates from the KaiLuan Study, completed  $\geq 1$  visits with the medical history recorded and physiological indicators measured; (2) without contraindications of brain MRI scan, retinal fundus photograph and OCT; and (3) can cooperate with the multi-modality medical imaging acquirement and clinical data collection.

### Exclusion criteria

Individuals should be excluded if any of the following is present: (1) known clinical history of congenital or acquired organic diseases (such as aortic stenosis); (2) known clinical history of psychiatric diseases (such as schizophrenia); (3) known history of drug abuse; (4) pregnant and lactating women; and (5) intracranial tumour accidentally detected on MR images.

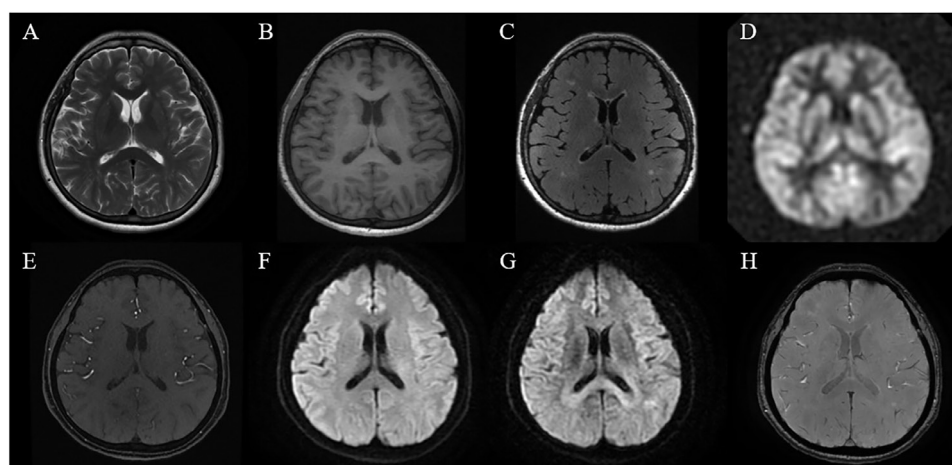
### Types of disease involved

Participants in the META-KLS will be followed up for a variety of neurological disorders including but not limited to cerebrovascular disease (eg, cerebral small vessel disease (CSVD), transient ischaemic attack, ischaemic stroke, cerebral haemorrhage), traumatic brain injury and neurodegenerative disease.

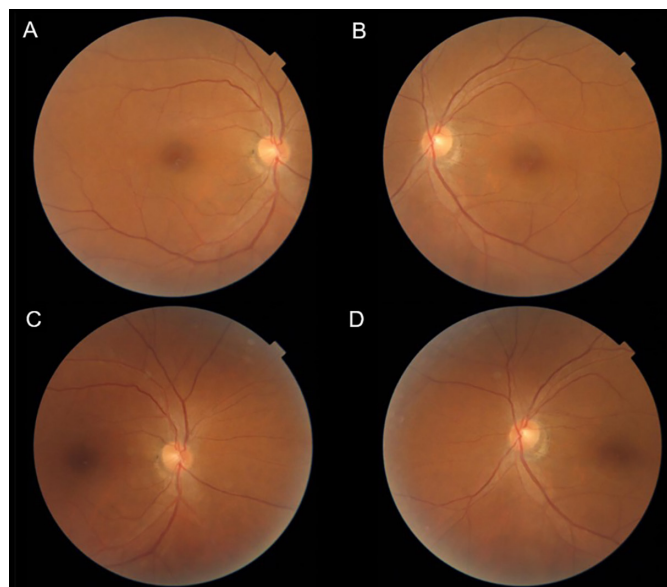
### Data collection

#### Multi-modality medical imaging

The brain MRI are collected using a 3.0 T scanner and an eight-channel phased-array coil since December 2020 (GE 750W; General Electric Medical Systems, Milwaukee, Wisconsin, USA). Data acquisition should include but not limited to the following morphological and functional sequences: T2-weighted image (T2WI), three-dimensional



**Figure 1** Depiction of the brain MRI in the Multi-modality Medical imaging sTudy bAsed on KaiLuan Study (META-KLS). (A) T2-weighted image (T1WI); (B) three-dimensional (3D) brain volume image for high-resolution T1WI; (C) fluid-attenuated inversion recovery image; (D) 3D arterial spin labelling; (E) 3D time-of-flight magnetic resonance angiography; (F) diffusion weighted imaging; (G) diffusion tensor imaging; (H) image of susceptibility-weighted angiography.



**Figure 2** The retinal fundus photographs. (A, B) Centred in macula for the right and left eyes; (C, D) centred in optic disc for the right and left eyes.

(3D) brain volume for high-resolution T1-weighted image (T1WI), 3D fluid-attenuated inversion recovery (FLAIR), 3D arterial spin labelling (ASL, PLD=1.5 s and 2.5 s), 3D time-of-flight magnetic resonance angiography (TOF-MRA) of head and neck, diffusion weighted imaging (DWI, b=0 and b=1000), diffusion tensor imaging (DTI), susceptibility-weighted angiography (SWAN). 3D T2WI and 3D ASL (PLD=2.0 s) sequences have been collected since March 2022. Ocular and otorhinolaryngologic structures can also be evaluated on MR images. The brain MRI acquisition protocol in META-KLS is described in online supplemental table 1.

Brain imaging data are collected in digital imaging and communications in medicine (DICOM) format. [Figure 1](#) takes one participant as an example to show the different sequences acquired.

Retinal fundus photographs and OCT will also be collected from each participant in the same day of

brain MRI scan. The bilateral retinal fundus images are captured by a fundus camera (CX-1, Topcon, Topcon Corporation, Tokyo, Japan). The fundus image photography is centred in macula and optic disc, respectively ([figure 2](#)). The OCT images are acquired in the 12-line 9 mm radial macula pattern and optic disc pattern from Topcon Deep Range Imaging OCT Triton device (Topcon, Tokyo, Japan). We will apply the eye-tracking system (follow-up mode) to enable measurement at approximately the same location among different subjects ([figure 3](#)).

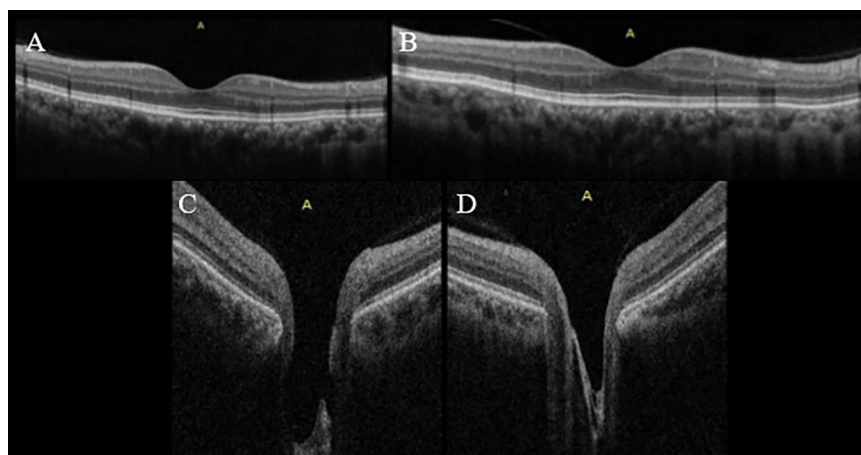
Since June 2022, an ultrasonic equipment (VINNO M86, VINNO TECHNOLOGY, Suzhou, China) has been used to perform ultrasound examinations on the bilateral ICAs for the evaluation of carotid lumen diameter, plaques, intima-media thickness (IMT), carotid-femoral pulse wave velocity (cfPWV), the compliance coefficient, distensibility coefficient and elastic modulus.

### Clinical data

#### Questionnaire assessments

All participants in the META-KLS will be assessed at baseline and follow-ups at 11 hospitals for clinical data collection. Demographics information (eg, date of birth, gender, marital status, education, family income), lifestyle (eg, smoking status, physical activities, salt intake, alcohol consumption), clinical history (eg, hypertension, diabetes, dyslipidaemia, cardiovascular/cerebrovascular disease and cancer) and family history (cardiovascular/cerebrovascular disease), medication use (eg, antihypertensive, cholesterol-lowering and glucose-lowering medications), labour intensity, neurological and mental health rating scales will be collected via face-to-face standardised questionnaires.

Specifically, smoking status is divided into three categories: 'never', 'former' and 'current'. Current smoker was defined as smoking at least one cigarette/day on average in a recent year.<sup>15</sup> Physical activities are evaluated based on responses to questions and combined with lifestyles of occupational and discretionary physical activities, and was



**Figure 3** The images of fundus optical coherence tomography. (A, B) Centred in macula for the right and left eyes; (C, D) centred in optic disc for the right and left eyes.

categorised as 'very active', 'moderately active' or 'inactive'. The use of antihypertensive, cholesterol-lowering and glucose-lowering medications within the past 2 weeks of the interview is self-reported.<sup>17</sup>

The commonly used neurological and mental health rating scales are collected to assess the severity of cognitive impairment (Montreal Cognitive Assessment, MoCA) and psychiatric disorders (eg, Patient Health Questionnaire-9 (PHQ-9) for depression and Generalised Anxiety Disorder 7-item (GAD-7) for anxiety), and to assess the ability of daily living (Activity of Daily Living Scale, ADL), individual self-efficacy (General Self-Efficacy Scale, GSES), overall stress (Perceived Stress Scale, PSS) and sleep quality (Athens Insomnia Scale (AIS) and Pittsburgh Sleep Quality Index (PSQI)).

### Physical examinations

Physical examinations (eg, body weight, height, waist circumference, hip circumference, heart rate, blood pressure at sitting and standing position, brachial ankle pulse wave velocity (BaPWV)) are conducted by trained physicians and nurses during each survey. Body weight is measured to the nearest 0.1 kg using a calibrated platform scale. Height is measured to the nearest 0.1 cm using a platform scale altimeter. Body mass index ( $\text{kg}/\text{m}^2$ ) is calculated as body weight divided by the square of height. Waist circumference is accurate to 0.1 cm at the midpoint between the subcostal margin and the iliac crest.<sup>18</sup> Hip circumference is accurate to 0.1 cm around the thighs and through the greater trochanter in the standing position.<sup>18</sup>

We are paying special attention to the blood pressure and BaPWV measurement in this cohort. Blood pressure measurements are taken between 07:00 and 09:00 on the examination day. Smoking and drinking of caffeinated or alcoholic beverages are prohibited for at least 3 hours, and exercise is prohibited for at least 30 min prior to measurement. All participants are required to rest for at least 5 min in the seated position. The reading of the first-time phase of kirschmann sound on the calibrated mercury sphygmomanometer is recorded as systolic blood pressure, and the reading of the fifth time phase of kirschmann sound is recorded as diastolic blood pressure. The mean of three measurements with 1–2 min interval is used for that visit.<sup>19</sup> Blood pressure at standing position is measured immediately after standing, 1 min after standing and 3 min after standing. It is taken with the arm supported at the elbow and the cuff at the heart level.

BaPWV is measured between 07:00 and 09:00 using BP-203 RPE III networked arteriosclerosis detection device since December 2020 (Omron Health Medical (China)).<sup>20</sup> After being seated for about 5 min in the examination room with temperature controlled between 22°C and 25°C, participants in thin clothes are asked to lay down on the examination table in supine position and remain quite during the measurement. Cuffs are wrapped on both arms and ankles. The lower edge of the arm cuff is positioned 2–3 cm above the transverse striation of the cubital fossa, and the lower edge of the

ankle cuff is positioned 1–2 cm above the superior aspect of medial malleolus. The ECG electrodes are placed on both wrists. One heart sound detector was placed at the left edge of the sternum. Measurement was repeated for twice for each subject. The result of second measurement was recorded.

### Laboratory assessments

Blood samples are taken from the anterior elbow vein from 07:00 to 09:00 after an overnight fast and transfused into vacuum tubes containing EDTA. Tubes are centrifuged at 3000g for 10 min at room temperature. After separation, plasma samples will be used within 4 hours. Laboratory test results (eg, total cholesterol, triglycerides, high-density lipoprotein, low-density lipoprotein, platelet count, neutrophil count, uric acid, C reactive protein, creatinine, fasting blood glucose, etc) are derived via an auto analyser (Hitachi 747; Hitachi, Tokyo, Japan) at the central laboratory at the same day.<sup>21</sup> Fasting blood glucose is measured using the hexokinase/glucose-6-phosphate dehydrogenase method (Mind Bioengineering, Shanghai, China) with an upper limit of detection of 30.07 mmol/L. Total cholesterol (TC) and triglycerides are measured enzymatically (interassay coefficient of variation <10%; Mind Bioengineering, Shanghai, China). The collected blood and urine samples can also be measured for exposure omics, metabolomics, whole genome and epigenetic indicators, and the content of heavy metal.

The potential risk factors to be investigated will be all kinds of clinical information collected above, including questionnaire assessments, physical examinations and laboratory assessments.

### Data processing

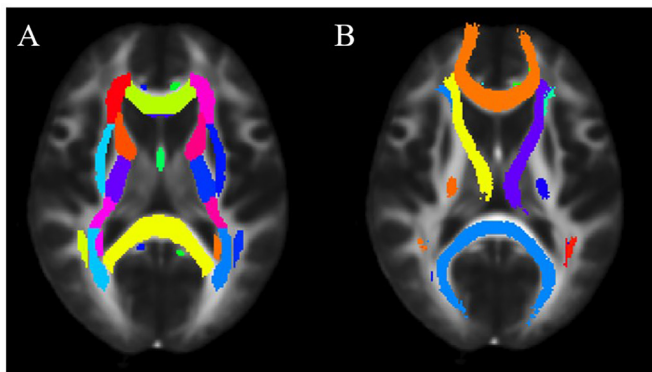
A standardised and validated processing workflow is being developed for imaging data to enable the accurate, objective, high-efficiency and reproducible analysis. We aim to extract relevant parameters related with brain microstructure and macrostructure, brain perfusion, possible cerebrovascular and eye disease from multi-modality medical imaging.

### Brain microstructure based on DTI analysis

The preprocessing of each participant's DTI data will be performed using MRtrix3<sup>22</sup> with the following steps: (1) image denoising,<sup>23</sup> (2) removing Gibbs ringing artefacts,<sup>24</sup> (3) motion and eddy current distortion correction,<sup>25</sup> (4) bias field correction<sup>26</sup> and (5) skull stripping.<sup>27</sup> The voxel-wise calculation of the diffusion metrics will be performed using FSL,<sup>28 29</sup> which yield maps of the fractional anisotropy (FA), mean diffusivity (MD), axial diffusivity (AD) and radial diffusivity (RD).

### Diffusion metrics calculation

For each subject, FA map will be aligned to the FMRIB58\_FA template using a linear and nonlinear alignment method (FLIRT and FNIRT). The other diffusion metric maps (MD, AD, RD) are subsequently transformed to the Montreal Neurological Institute (MNI) space with



**Figure 4** Brain atlas applied for diffusion metrics calculation. (A) Johns Hopkins University (JHU) labels atlas; (B) JHU tracts atlas.

non-linear registration. These normalised maps are resampled to 2mm isotropic. The regional diffusion metrics are calculated by averaging the values with each region of the WM atlases. For voxel-based analysis, the spatially normalised diffusion metric map will be smoothed with a Gaussian kernel (6mm full-width at half maximum (FWHM)). Johns Hopkins University (JHU) labels atlas and JHU tractography atlas were applied (figure 4 and online supplemental tables 2 and 3).

#### *Tract-based spatial statistics (TBSS) analysis*

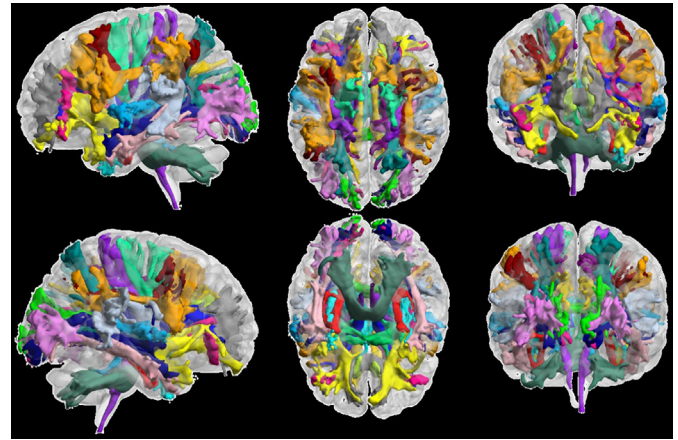
The mean FA maps for all subjects are created and refined to create a mean FA skeleton that represented the centres of neural tracts common to all subjects. Finally, all individual parameter values (FA, MD, AD and RD) are projected onto the population skeleton. The diffusion metrics along the skeleton are calculated by averaging the values within regions of interest (ROI) (figure 5 and online supplemental table 4). The results of left and right sided cerebrum, cerebellum and brainstem are also recorded.

#### *White matter tracts analysis*

XTRACT pipeline implemented in FSL will be performed to extract a set of tracts in the subject's native space automatically using probabilistic diffusion tractography.<sup>30</sup> Briefly, the approach include the following steps: (1) the probabilistic diffusion model is fitted based on the preprocessed DTI data with the FSL bedpostx toolbox;<sup>31</sup> (2) probabilistic tractography is run with the precalculated non-linear field coefficients to extract forty-two white matter tracts between the predefined 42 pair of seeds and targets; (3) the mean, median and standard



**Figure 5** White matter skeleton overlay the brain lobes.

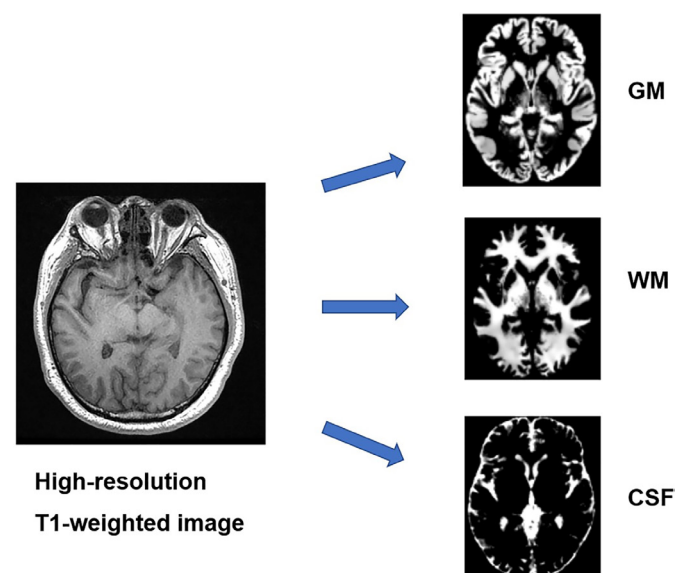


**Figure 6** Forty-two white matter tracts.

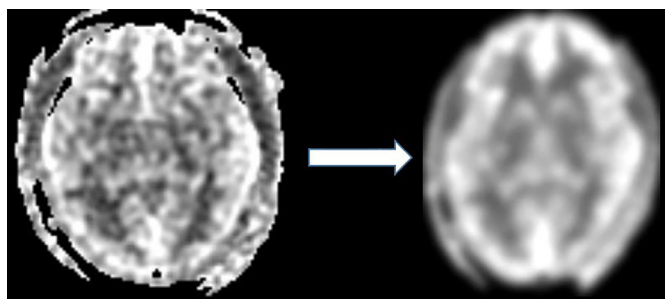
division of volume, length, probability, FA, MD, AD and RD along the forty-two white matter tracts are calculated for further statistics. To account for better coverage due to the relatively large slice thickness, the seed and target masks are altered by increasing 3mm in the Z direction (figure 6 and online supplemental table 5).

#### *Brain macrostructure based on high-resolution T1WI*

All of the high-resolution T1WI will be checked to exclude images with artefacts that caused by head motion, susceptibility artefacts and instrument malfunction. We will apply CAT12 package (<http://www.neuro.uni-jena.de>) based on Statistical Parametric Mapping (SPM)12 for the structural data analysis with the following steps: (1) Reorientation. The images are reoriented to place the anterior commissure at the origin and the anterior-posterior commissure in the horizontal plane. (2) Segmentation. The images are segmented into grey matter (GM), white matter (WM) and cerebrospinal fluid (CSF) regions (figure 7). (3)



**Figure 7** The high-resolution T1-weighted image is segmented into grey matter (GM), white matter (WM) and cerebrospinal fluid (CSF).



**Figure 8** Original and smoothed cerebral blood flow maps.

Normalisation. The individual WM and GM components are normalised into the standard MNI space using the diffeomorphic anatomical registration through Cat12 default algorithm. The normalised GM/WM component is modulated to calculate the relative GM/WM volume by multiplication by the non-linear part of the deformation field at the normalisation step.

Based on the preprocessed images, we can calculate the value of total intracranial volume, total GM, total WM and total CSF of all subjects. Second, using the ROI signal extractor in the DPABI package (<http://rfmri.org/dpabi>), we can extract the average intensity value of all voxels of 116 sub-brain regions of each subject according to the AAL\_90 atlas (online supplemental table 6), and then, multiply by the size of each voxel to obtain the absolute volume value of each region.

#### Brain perfusion based on ASL

We will obtain the maps of CBF by pairwise subtraction of the ASL control and label images. The CBF images of each participant are co-registered to a PET-perfusion template in the standard space of MNI by SPM12 with resampling to 2 mm×2 mm×2 mm. Then, we smooth the CBF images with an 8 mm FWHM Gaussian kernel (figure 8). The CBF of the whole brain, GM, WM and the different hemisphere of the results are recorded. The CBF values of additional 24 sub-brain regions are also recorded (online supplemental table 7).

#### CSVD burden assessment

CSVD burden assessment included the following four factors.

##### White matter hyperintensity (WMH)

WMH is a type of white matter lesion shown as increased brightness on T2 images such as FLAIR, without cavitation.<sup>32</sup> The Fazekas rating scale is used to assess the severity of periventricular and deep WMH.<sup>33</sup>

##### Lacunes

Lacunes are defined as round or ovoid lesions 3–15 mm in diameter that exhibit the signal features similar to CSF on all sequences and appear as a hyperintense rim on FLAIR sequences.<sup>32</sup> The numbers and locations of Lacunes are recorded.

##### Cerebral microbleeds (CMBs)

CMBs are defined as lesions with very low signal intensity with associated blooming on T2\*-weighted, gradient-recalled echo (GRE) images or sensitivity-weighted images.<sup>32 34</sup> The numbers and locations of CMBs are recorded.

##### Enlarged perivascular spaces (EPVS)

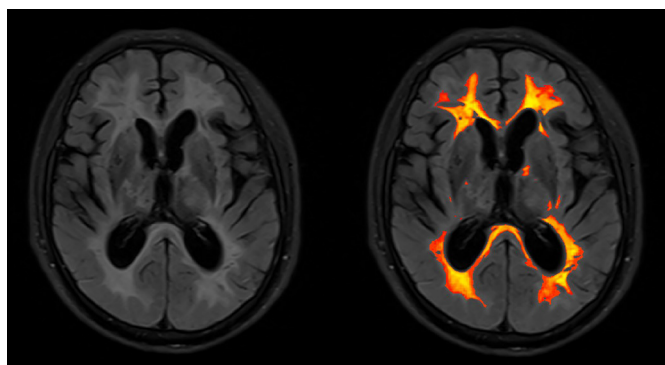
EPVS is the round, ovoid or linear space less than 3 mm in diameter with CSF signal on all sequences in the basal ganglia or centrum semiovale.<sup>32</sup> EPVS is rated on a previously described, validated semiquantitative scale from 0 to 4.<sup>35</sup>

The brain imaging assessment will be performed by two well-trained neuroradiologists and further confirmed by a third neuroradiologist with more than 10 years of experience to resolve discordance, who is blinded to the clinical information of the subjects. The total MRI burden of CSVD is rated on an ordinal scale from 0 to 4 according to the widely accepted validated method that developed by Wardlaw's group,<sup>35</sup> by counting the presence of each of the four MRI features of CSVD. A point is awarded for presence of one or more lacunes or any CMB. If there are moderate to severe (grade 2–4) EPVS in the basal ganglia, the presence of EPVS is scored as 1 point. The presence of (early) confluent deep WMH (Fazekas score 2–3) or irregular periventricular WMH extending into the deep white matter (Fazekas score 3) is recorded 1 point for the feature of WMH. The total CSVD score can provide a complete estimate of the overall impact of CSVD on the brain in a simple and pragmatic manner. It may have the potential for risk stratification and efficacy assessment in clinical trials of interventions to prevent the progress of CSVD.

The volume of WMH is also calculated based on 3D FLAIR images by using the Lesion Prediction Algorithm (LPA),<sup>36</sup> a widely used unsupervised pipeline implemented in the LST toolbox for the SPM software. LPA is a binary classifier in the form of a logistic regression model. The parameters of this model fit are used to segment lesions by providing an estimate for the lesion probability for each voxel (figure 9). The lesion probability map is then converted to the WMH volumes (mL) and the fraction of total brain volume occupied by WMH (%) within both the deep white matter region and periventricular region (figure 10).

##### Retina layer thickness analysis based on OCT

Both the optic nerve head (ONH) and macular OCT images are analysed with the automatic OCT layer segmentation algorithm (Retinal Image Analysis Lab, Iowa Institute for Biomedical Imaging, Iowa City, Iowa, USA).<sup>37–39</sup> The accuracy and reproducibility of the Iowa Reference Algorithms have been reported in patients with diabetic macular oedema and in healthy volunteers.<sup>40 41</sup> The feasibility has also been demonstrated in Asian populations with Iowa Reference Algorithms.<sup>42</sup> In this study, the following pipeline will be employed.



**Figure 9** The fluid-attenuated inversion recovery image was segmented into white matter hyperintensity lesion probability map.

First, 10 retinal layer (11 boundary) segmentation of an OCT image is obtained on both ONH and macular OCT images (figure 11), including (1) retinal nerve fibre layer (RNFL); (2) ganglion cell layer (GCL); (3) inner plexiform layer (IPL); (4) inner nuclear layer (INL); (5) outer plexiform layer (OPL); (6) outer nuclear layer (ONL); (7) photoreceptor inner/outer segments (IS/OS); (8) inner/outer segment junction to inner boundary of outer segment photoreceptor/retinal pigment epithelium complex (IS/OSJ to IB\_RPE); (9) outer segment photoreceptor/retinal pigment epithelium complex (OPR); (10) retinal pigment epithelium (RPE).

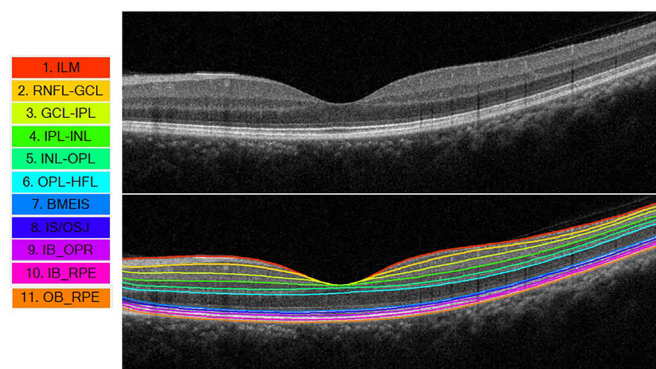
Second, for macular OCT images, mean and standard deviation of retinal thickness of 10 retinal layers are calculated within nine regions including the foveal subfield, as well as the inner and outer rings of a standard Early Treatment Diabetic Retinopathy Study (ETDRS) grid (figure 12).

At last, for ONH OCT images, mean and SD of retinal thickness of 10 retinal layers are calculated within six regions divided by ellipse grid (figure 13).

Furthermore, fundus photography images are used to identify retinal lesions (eg, (retinal nerve fibre layer) RNFL defects) and to perform automatic segmentation and quantitative measurement of retinal vessels and evaluating optic disc appearance in the posterior pole.

### New horizons for data processing

The automatic calculation method for the number and (or) volume of lacunes, CMBs, EPVS will be fully discussed and further developed based on the collected brain MRI data. Automatic segmentation and data extraction from



**Figure 11** Ten retinal layer (11 boundary) segmentation of an optical coherence tomography image. GCL, ganglion cell layer; INL, inner nuclear layer; IPL, inner plexiform layer; OPL, outer plexiform layer; OPR, outer segment photoreceptor/retinal pigment epithelium complex; RPE, retinal pigment epithelium.

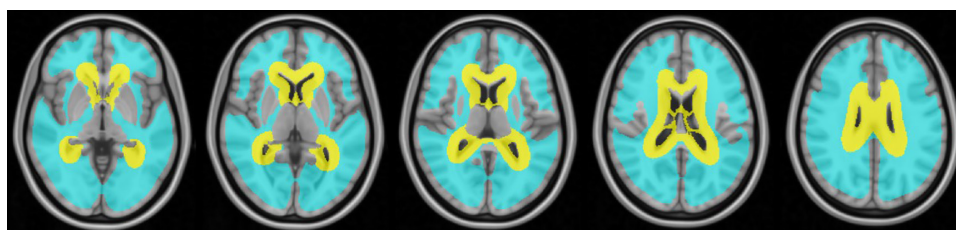
retinal fundus photograph and OCT will be developed or updated.

### Sample size determination

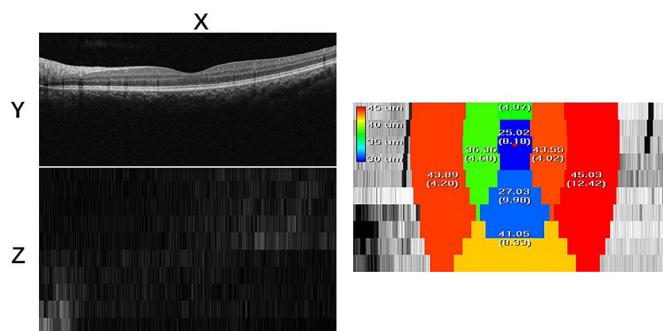
The PASS software (V.11.0 by NCSS, LLC) was used for sample size calculation. The statistician referenced hypertension, a binary categorical variable, for example, and calculated the approximate sample size based on 200 participants in META-KLS. We assume that the mean and SD of the hypertension group was  $616.7 \pm 45 \text{ mm}^3$ , and that of the non-hypertension group was  $608.2 \pm 43 \text{ mm}^3$ . Using the independent and two-sided t-test, which set the  $\alpha=0.05$ ,  $\beta=0.2$ , power=80%, and the sample size ratio between the two groups was 1:1, we calculated the sample size was 421 for each group. Assuming the 20% loss to the follow-up rate, the estimated minimum sample size needed for this study was 1012 individuals.

### Database building

Epidata V.3.1 ((Jens M. Lauritsen, Odense, Denmark) is used to input clinical data, including questionnaire assessments, physical examinations and laboratory assessments. Data entry will be performed in parallel by two independent registrants following double-blind entry rules and then exported as a SAS database. Cleaned imaging data following the standardised operation principle as previously described will also be stored in a separate SAS database. Finally, they are combined into one SAS database with a unique ID.



**Figure 10** Deep white matter region and periventricular region.



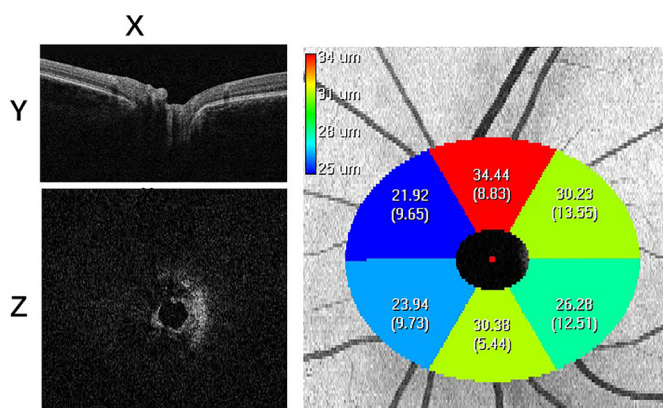
**Figure 12** ETDRS grid of the macular optical coherence tomography images.

The stenosis degree of ICA will be recorded as a categorical variable, and plaque size, IMT, cfPWV, etc, will be expressed as continuous numerical variables. After cleaning and quality control procedures, these data will also be entered into each subject's SAS database. In the future, the investigators are able to estimate the association between a variety of risk factors and the characteristics of ICA measured by ultrasound.

### Quality control of the database

The questionnaires are designed by epidemiology professionals. The participants will be interviewed face-to-face by doctors or nurses with uniform training. All hospitals are equipped with the same brand and model of height, weight and blood pressure measurement equipment. Laboratory tests are conducted in the central laboratory of Kailuan General Hospital according to standardised operation process.

Data governance will be conducted by all the principal investigators after professional training. The quality control group will verify the authenticity, integrity and accuracy of the clinical and imaging data based on the source information. Data modification will be recorded and informed to the principal investigators. After data verification, the database will be locked for further analysis.



**Figure 13** Ellipse grid of the optic nerve head optical coherence tomography images.

## Statistical analysis

Statistical analysis will be performed with SAS V.9.4 (SAS Institute, Cary, North Carolina, USA), SPSS V.25.0 (SPSS, Chicago, Illinois, USA) and R V.4.2.1 (R Foundation for Statistical Computing, Vienna, Austria). Continuous variables meeting normal distribution are expressed as mean±SD, and data with skewed distribution are presented as median (IQR). The threshold for statistical significance is two-sided  $p<0.05$ .

For group comparison, we will perform independent sample t-test, Mann-Whitney U test, Kruskal-Wallis, and analysis of variance test, etc. In addition, with multiple measurements during follow-ups, we are able to calculate the effect of cumulative exposure of risk factors on brain health. It is calculated as the sum of area under the trajectory during the time of follow-up (years) divided by the duration of follow-up. To analyse the variance (direction and magnitude) of factor changes, the slope can also be calculated by fitting a general linear model based on the average variance during follow-ups.

## Patient and public involvement

Participants will not be involved in the design, recruitment, conduct or dissemination of the study.

## DISCUSSION

The META-KLS study was designed to cope with the demographic changes that have led to an era of continuously increased life expectancy and an increasing proportion of the elderly in most populations.<sup>43</sup> Under such a background, subclinical brain disease has become a hot topic in current researches and has received great attention. Many risk factors have proven to be strongly associated with subclinical brain diseases. However, cross-sectional studies often cannot explain the changes in brain imaging. A predominant approach to finding such associations is a prospective cohort study, and multi-modality medical imaging is one of the best methods to analyse the presence and severity. Therefore, studying the long-term effect of these risk factors, such as cumulative exposure on brain health will be able to provide new insights into potential strategies for the early prevention and intervention of brain diseases.

The META-KLS is, to our knowledge, the first community-based database to comprehensively evaluate the association between a variety of risk factors and subclinical brain disease using multi-modality medical imaging. Researches based on this cohort will provide a series of state-of-the-art evidence for the prevention of neurological diseases and common chronic diseases.

Compared with the Rotterdam Scan<sup>6</sup> or PRECISE study,<sup>7</sup> META-KLS has shown significant strengths. The minimum age for enrolment in the Rotterdam and PRECISE study was 40 and 50 years, respectively,<sup>7 43</sup> while META-KLS will include a broader age range of volunteers aged between 18 and 98 years, enabling a more precise age-stratified analysis. The wider age range is also

conducive to the detection of early intracerebral lesions and even preclinical brain microstructural changes, thereby expanding our understanding of pathological changes before neurological symptoms present.

Furthermore, as this study is a subanalysis from the KaiLuan Study, we are able to continuously recruit participants subjecting to their will, which is an important supplement and guarantee for the research sample. The homogeneity of study participants can avoid potential confounding effects caused by racial/ethnic and health-care disparities.<sup>9</sup> We can also synchronise the KaiLuan Study for regular follow-ups and perform multi-modality medical imaging including brain MRI, retinal fundus photograph and OCT of the participants. In particular, the wide application of high-resolution imaging facilitates fast and accurate data processing and also enables automatic segmentation and data extraction based on artificial intelligence software.

Multi-modality medical imaging may also facilitate research on many kinds of diseases other than subclinical brain disease. Not only brain imaging but also the retinal fundus photograph and OCT may contribute early detection of common chronic diseases.<sup>44</sup>

Another strength of the META-KLS is that we collect clinical and laboratory examinations repeatedly, thus may reduce random error. In addition, we could perform long-term assessment of these clinical markers, such as the cumulative exposure and variation over periods.<sup>3 45</sup> Investigating and categorising the trajectories based on the baseline data and at least two follow-up visits may also provide insight into the potential association between long-term changes in risk factors and subclinical brain diseases. Regarding occupational exposure, subgroup analysis of workers with different levels of labour intensity is also a highlight of this study.

In total, the procedures strictly abide by ethical principles. The examinations will be performed at the will of the participants. The research committee will cover the costs of all examinations, including the collection of clinical data and multi-modality medical imaging. Besides, participants will be informed of the results and instructed to seek medical advice and develop effective early intervention if necessary.

### Limitations

This study has potential limitations. First, as we are more concerned about subclinical brain disease and the changes in brain imaging markers, the assessment of cognitive and neuropsychiatric function is insufficient. Second, CT will not be collected, as the radiation exposure of the population does not conform to the principles of ethics committee. It is also restricted by time, costs and medical resources. Third, this study lacks outcomes such as all-cause mortality or incident CVD events (eg, coronary heart disease or stroke). Third, as the participants are recruited from one community in China, the findings may not be generalisable to other populations and require external validation.

### ETHICS AND DISSEMINATION

The KaiLuan Study and META-KLS have been approved by the Medical Ethics Committee of Kailuan General Hospital (IRB number: 2008No. 1 and 2021002, respectively). These two trials were registered online (ChiCTR2000029767 on [chictr.org.cn](http://chictr.org.cn) and NCT05453877 on [ClinicalTrials.gov](http://ClinicalTrials.gov), respectively). The registration of META-KLS on [ClinicalTrials.gov](http://ClinicalTrials.gov) (<https://clinicaltrials.gov/>) is identical to this protocol. All the procedures will be performed in accordance with the principles of the Declaration of Helsinki. Written informed consent must be obtained from each participant. Results are expected to be first reported in professional peer-reviewed journals in 2023. The data can be accessed according to the request from corresponding authors.

### Author affiliations

<sup>1</sup>Department of Radiology, Beijing Friendship Hospital, Capital Medical University, Beijing, China

<sup>2</sup>Clinical Epidemiology & EBM Unit, Beijing Friendship Hospital, Capital Medical University; National Clinical Research Center for Digestive Diseases, Beijing, China

<sup>3</sup>Department of Medical Imaging, Yanjing Medical College, Capital Medical University, Beijing, China

<sup>4</sup>Center for MRI Research, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing, China

<sup>5</sup>Beijing Intelligent Brain Cloud Inc, Beijing, China

<sup>6</sup>Department of Cardiology, Kailuan General Hospital, Tangshan, Hebei, China

<sup>7</sup>Department of Rheumatology and Immunology, Kailuan General Hospital, Tangshan, Hebei, China

<sup>8</sup>Department of Psychiatry, Kailuan Mental Health Center, Tangshan, Hebei, China

<sup>9</sup>Department of MR, Kailuan General Hospital, Tangshan, Hebei, China

**Acknowledgements** The authors would like to thank all of the involved investigators, research coordinators, clinicians, nurses, technicians and staffs for their contribution to this cohort study.

**Contributors** HL, SW and ZW have full access to the dataset and take responsibility for the accuracy and integrity of the data. Concept and design: HL, JS, SW and ZW. Acquisition, analysis or interpretation of data: YH, JL and XL. Protocol drafting: JS, HL, YH, JL, XZ, QC, XL, NW and MX. Critical revision of the protocol for important intellectual content: HL, SW and ZW. Statistical analysis: XZ. Administrative, technical or material support: PZ, WL, RL, YW, AX, HS, SZ, XL and YW. All authors approved the submission of the final version of the protocol.

**Funding** This work was supported by Grant 61931013 (ZW), 62171297 (HL), 82272072 (JL) from the National Natural Science Foundation of China, No. BYESS2022073 (HL) from Beijing Association for Science and Technology, No. (2015) 160 from Beijing Scholars Program (ZW), No. ZYLX202101 from Beijing Hospitals Authority Clinical Medicine Development of Special Funding Support, No. 2021-135 from Beijing Municipal Health Commission-Beijing Key Clinical Discipline Funding.

**Competing interests** None declared.

**Patient and public involvement** Patients and/or the public were not involved in the design, or conduct, or reporting or dissemination plans of this research.

**Patient consent for publication** Obtained.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

#### ORCID iDs

Wenjuan Liu <http://orcid.org/0000-0001-6555-9532>

Han Lv <http://orcid.org/0000-0001-9559-4777>

## REFERENCES

- de Groot M, Cremers LGM, Ikram MA, et al. White matter degeneration with aging: longitudinal diffusion MR imaging analysis. *Radiology* 2016;279:532–41.
- Sedaghat S, Cremers LGM, de Groot M, et al. Lower microstructural integrity of brain white matter is related to higher mortality. *Neurology* 2016;87:927–34.
- Ma Y, Yilmaz P, Bos D, et al. Blood pressure variation and subclinical brain disease. *J Am Coll Cardiol* 2020;75:2387–99.
- Jing J, Zhou Y, Pan Y, et al. Reduced white matter microstructural integrity in prediabetes and diabetes: a population-based study. *EBioMedicine* 2022;82:104144.
- Liu J-Y, Zhou Y-J, Zhai F-F, et al. Cerebral microbleeds are associated with loss of white matter integrity. *AJNR Am J Neuroradiol* 2020;41:1397–404.
- Ikram MA, van der Lugt A, Niessen WJ, et al. The Rotterdam scan study: design and update up to 2012. *Eur J Epidemiol* 2011;26:811–24.
- Pan Y, Jing J, Cai X, et al. Polyvascular evaluation for cognitive impairment and vascular events (precise) – A population-based prospective cohort study: rationale, design and baseline participant characteristics. *Stroke Vasc Neurol* 2021;6:145–51.
- Zhao M, Song L, Sun L, et al. Associations of type 2 diabetes onset age with cardiovascular disease and mortality: the kailuan study. *Diabetes Care* 2021;44:1426–32.
- Wu S, An S, Li W, et al. Association of trajectory of cardiovascular health score and incident cardiovascular disease. *JAMA Netw Open* 2019;2:e194758.
- Zhang Q, Zhou Y, Gao X, et al. Ideal cardiovascular health metrics and the risks of ischemic and intracerebral hemorrhagic stroke. *Stroke* 2013;44:2451–6.
- Wong JC, Li J, Pavlova M, et al. Risk factors for probable REM sleep behavior disorder: a community-based study. *Neurology* 2016;86:1306–12.
- Jin C, Chen S, Vaidya A, et al. Longitudinal change in fasting blood glucose and myocardial infarction risk in a population without diabetes. *Diabetes Care* 2017;40:1565–72.
- Wang Z, Zhao X, Chen S, et al. Associations between nonalcoholic fatty liver disease and cancers in a large cohort in china. *Clin Gastroenterol Hepatol* 2021;19:788–96.
- Wu S, Jin C, Li S, et al. Aging, arterial stiffness, and blood pressure association in Chinese adults. *Hypertension* 2019;73:893–9.
- Huang S, Li J, Shearer GC, et al. Longitudinal study of alcohol consumption and HDL concentrations: a community-based study. *Am J Clin Nutr* 2017;105:905–12.
- Hedman AM, van Haren NEM, Schnack HG, et al. Human brain changes across the life span: a review of 56 longitudinal magnetic resonance imaging studies. *Hum Brain Mapp* 2012;33:1987–2002.
- Wu S, Huang Z, Yang X, et al. Prevalence of ideal cardiovascular health and its relationship with the 4-year cardiovascular events in a northern Chinese industrial City. *Circ Cardiovasc Qual Outcomes* 2012;5:487–93.
- Hu S, Li X, Sun Y, et al. Short-Term weight loss decreased the risk of chronic kidney disease in men with incident nonalcoholic fatty liver disease. *Obesity (Silver Spring)* 2022;30:1495–506.
- Zheng M, Zhang X, Chen S, et al. Arterial stiffness preceding diabetes: a longitudinal study. *Circ Res* 2020;127:1491–8.
- Chen S, Li W, Jin C, et al. Resting heart rate trajectory pattern predicts arterial stiffness in a community-based Chinese cohort. *Arterioscler Thromb Vasc Biol* 2017;37:359–64.
- Wang C, Yuan Y, Zheng M, et al. Association of age of onset of hypertension with cardiovascular diseases and mortality. *J Am Coll Cardiol* 2020;75:2921–30.
- Tournier J-D, Smith R, Raffelt D, et al. MRtrix3: a fast, flexible and open software framework for medical image processing and visualisation. *Neuroimage* 2019;202.
- Veraart J, Novikov DS, Christiaens D, et al. Denoising of diffusion MRI using random matrix theory. *Neuroimage* 2016;142:394–406.
- Kellner E, Dhital B, Kiselev VG, et al. Gibbs-ringing artifact removal based on local subvoxel-shifts. *Magn Reson Med* 2016;76:1574–81.
- Andersson JLR, Sotiropoulos SN. An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *Neuroimage* 2016;125:1063–78.
- Tustison NJ, Avants BB, Cook PA, et al. N4ITK: improved N3 bias correction. *IEEE Trans Med Imaging* 2010;29:1310–20.
- Dhollander T RD, Connelly A. Unsupervised 3-tissue response function estimation from single-shell or multi-shell diffusion MR data without a co-registered T1 image. ISMRM Workshop on Breaking the Barriers of Diffusion MRI; 2016:5.
- Basser PJ, Mattiello J, LeBihan D. Estimation of the effective self-diffusion tensor from the NMR spin echo. *J Magn Reson B* 1994;103:247–54.
- Jenkinson M, Beckmann CF, Behrens TEJ, et al. Fsl. *Neuroimage* 2012;62:782–90.
- Warrington S, Bryant KL, Khrapitchev AA, et al. XTRACT-standardised protocols for automated tractography in the human and macaque brain. *Neuroimage* 2020;217:116923.
- Hernández M, Guerrero GD, Cecilia JM, et al. Accelerating fibre orientation estimation from diffusion weighted magnetic resonance imaging using gpus. *PLoS One* 2013;8:e61892.
- Wardlaw JM, Smith EE, Biessels GJ, et al. Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. *Lancet Neurol* 2013;12:S1474–4422(13)70124–8:822–38..
- Fazekas F, Kleinert R, Offenbacher H, et al. Pathologic correlates of incidental MRI white matter signal hyperintensities. *Neurology* 1993;43:1683–9.
- Greenberg SM, Vernooij MW, Cordonnier C, et al. Cerebral microbleeds: a guide to detection and interpretation. *Lancet Neurol* 2009;8:165–74.
- Staals J, Makin SDJ, Doubal FN, et al. Stroke subtype, vascular risk factors, and total MRI brain small-vessel disease burden. *Neurology* 2014;83:1228–34.
- Schmidt P. *Bayesian inference for structured additive regression models for large-scale problems with applications to medical imaging* [PhD thesis]. LudwigMaximilians-Universität München, 2017 Available: <http://nbn-resolving.de/urn:nbn:de:bvb:19-203731>
- Li K, Wu X, Chen DZ, et al. Optimal surface segmentation in volumetric images -- a graph-theoretic approach. *IEEE Trans Pattern Anal Mach Intell* 2006;28:119–34.
- Garvin MK, Abramoff MD, Wu X, et al. Automated 3-D intraretinal layer segmentation of macular spectral-domain optical coherence tomography images. *IEEE Trans Med Imaging* 2009;28:1436–47.
- Antony B, Abramoff MD, Tang L, et al. Automated 3-D method for the correction of axial artifacts in spectral-domain optical coherence tomography images. *Biomed Opt Express* 2011;2:2403–16.
- Terry L, Cassels N, Lu K, et al. Automated retinal layer segmentation using spectral domain optical coherence tomography: evaluation of inter-session repeatability and agreement between devices. *PLoS One* 2016;11:e0162001.
- Sohn EH, Chen JJ, Lee K, et al. Reproducibility of diabetic macular edema estimates from SD-OCT is affected by the choice of image analysis algorithm. *Invest Ophthalmol Vis Sci* 2013;54:4184–8.
- Chua J, Tham YC, Tan B, et al. Age-Related changes of individual macular retinal layers among Asians. *Sci Rep* 2019;9:20352.
- Ikram MA, Brusselle GGO, Murad SD, et al. The Rotterdam study: 2018 update on objectives, design and main results. *Eur J Epidemiol* 2017;32:807–50.
- Zhang K, Liu X, Xu J, et al. Deep-learning models for the detection and incidence prediction of chronic kidney disease and type 2 diabetes from retinal fundus images. *Nat Biomed Eng* 2021;5:533–45.
- Wang X, Feng B, Huang Z, et al. Relationship of cumulative exposure to the triglyceride-glucose index with ischemic stroke: a 9-year prospective study in the kailuan cohort. *Cardiovasc Diabetol* 2022;21:66.