



HAL
open science

Strengths and challenges of longitudinal non-human primate neuroimaging

Xiaowei Song, Pamela García-Saldivar, Nathan Kindred, Yujiang Wang, Hugo Merchant, Adrien Meguerditchian, Yihong Yang, Elliot A Stein, Charles W Bradberry, Suliann Ben Hamed, et al.

► To cite this version:

Xiaowei Song, Pamela García-Saldivar, Nathan Kindred, Yujiang Wang, Hugo Merchant, et al.. Strengths and challenges of longitudinal non-human primate neuroimaging. *NeuroImage*, 2021, pp.118009. 10.1016/j.neuroimage.2021.118009 . hal-03196946

HAL Id: hal-03196946

<https://amu.hal.science/hal-03196946v1>

Submitted on 13 Apr 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Journal Pre-proof

Strengths and challenges of longitudinal non-human primate neuroimaging

Xiaowei Song , Pamela García-Saldivar , Nathan Kindred ,
Yujiang Wang , Hugo Merchant , Adrien Meguerditchian ,
Yihong Yang , Elliot A. Stein , Charles W. Bradberry ,
Suliann Ben Hamed , Hank P. Jedema , Colline Poirier

PII: S1053-8119(21)00286-X
DOI: <https://doi.org/10.1016/j.neuroimage.2021.118009>
Reference: YNIMG 118009



To appear in: *NeuroImage*

Received date: 21 July 2020
Revised date: 16 March 2021
Accepted date: 23 March 2021

Please cite this article as: Xiaowei Song , Pamela García-Saldivar , Nathan Kindred , Yujiang Wang , Hugo Merchant , Adrien Meguerditchian , Yihong Yang , Elliot A. Stein , Charles W. Bradberry , Suliann Ben Hamed , Hank P. Jedema , Colline Poirier , Strengths and challenges of longitudinal non-human primate neuroimaging, *NeuroImage* (2021), doi: <https://doi.org/10.1016/j.neuroimage.2021.118009>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2021 Published by Elsevier Inc.
This is an open access article under the CC BY-NC-ND license
(<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

Highlights

- Strengths and challenges of longitudinal non-human primate MRI are described.
- Statistical power calculation of longitudinal and cross-sectional designs are provided.
- The impact of template choice on grey matter estimation is demonstrated.
- Recommendations for designing and analysing such studies are provided.

Journal Pre-proof

Strengths and challenges of longitudinal non-human primate neuroimaging

Xiaowei Song¹, Pamela García-Saldivar², Nathan Kindred³, Yujiang Wang⁴, Hugo Merchant², Adrien Meguerditchian⁵, Yihong Yang⁶, Elliot A. Stein⁶, Charles W. Bradberry¹, Suliann Ben Hamed⁷, Hank P. Jedema^{1}, Colline Poirier^{3*}*

1. Preclinical Pharmacology Section, Intramural Research Program, NIDA, NIH, Baltimore, MD 21224, USA.
2. Instituto de Neurobiología, UNAM, Campus Juriquilla. Boulevard Juriquilla No. 3001 Querétaro, Qro. 76230 México.
3. Biosciences Institute & Centre for Behaviour and Evolution, Faculty of Medical Sciences, Newcastle University, United Kingdom.
4. CNNP Lab (www.cnnp-lab.com), Interdisciplinary Complex Systems Group, School of Computing, Newcastle University, United Kingdom.
5. Laboratoire de Psychologie Cognitive, UMR7290, Université Aix-Marseille/CNRS, Institut Language, Communication and the Brain 13331 Marseille, France.
6. Neuroimaging Research Branch, Intramural Research Program, NIDA, NIH, Baltimore, MD 21224, USA.
7. Institut des Sciences Cognitives Marc Jeannerod, UMR 5229, Université de Lyon – CNRS, France.

*Corresponding authors:

colline.poirier@ncl.ac.uk

hank.jedema@nih.gov

Abstract

Longitudinal non-human primate neuroimaging has the potential to greatly enhance our understanding of primate brain structure and function. Here we describe its specific strengths, compared to both cross-sectional non-human primate neuroimaging and longitudinal human neuroimaging, but also its associated challenges. We elaborate on factors guiding the use of

different analytical tools, subject-specific versus age-specific templates for analyses, and issues related to statistical power.

Keywords: Magnetic resonance imaging; development; ageing; templates; simulation; non-human primate

Journal Pre-proof

1. Introduction

Non-human primate (NHP) neuroimaging is a field progressively establishing itself as a crucial complement to human neuroimaging (Phillips et al., 2014; Roelfsema and Treue, 2014). The use of NHP models not only allows performing experiments that cannot be done in humans but also can shed light on the evolution of the primate brain. NHP neuroimaging has mostly been dominated by between-subject, cross-sectional experimental designs. However, within-subject, longitudinal designs are generally more powerful, but also more challenging. Here we compare longitudinal NHP neuroimaging with both cross-sectional NHP neuroimaging and longitudinal human neuroimaging. We describe its inherent strength in terms of statistical power and its specific strengths in developmental and ageing studies, as well as in interventional studies. We then describe specific challenges, encompassing data acquisition, image processing and statistical analyses.

2 Strengths of longitudinal approaches

2.1 Inherent strength: Cohort size and sample size

NHP neuroimaging operates within a stringent regulatory framework. While regulations vary between countries, a common principle applied globally is the '3Rs' (*Replacement, Reduction and Refinement*) (Mitchell et al., this issue). As a consequence, animal numbers need to be kept to a minimum. Longitudinal designs present the crucial advantage of having greater statistical power than cross-sectional ones (Liu and Liang, 1997, Fitzmaurice et al., 2012, Tustison et al 2019), offering a powerful way to reduce the number of animals used.

Statistical power, the probability to detect a true effect, depends on the size of the true effect, the statistical threshold, the sample size, and the amount of variability in the response variable. For a fixed effect size, fixed statistical threshold, and fixed sample size, statistical power therefore depends on the amount of variability in the response variable, with more variability leading to less power to detect statistical differences that can be attributed to the true effect. By focusing on within-subject differences, longitudinal studies circumvent the problem of inter-individual variability due to genes and gene X environment interactions, resulting in lower variability, and therefore, increased power. Importantly, longitudinal studies have greater statistical power than cross-sectional studies not only for a fixed number of subjects, but also for a fixed number of scans (fig. 1). While longitudinal studies during childhood and adulthood both benefit from increased

power (compared to a cross-sectional design), studies during adulthood benefit more, since inter-individual variability tends to be larger in adults than in juveniles (as individual gene X environment interactions increase with time).

Data sharing offers the opportunity to further increase the statistical power of longitudinal NHP studies by increasing the cohort size. The NHP neuroscience community has recently started to share neuroimaging datasets (Milham, 2018, 2020). While only cross-sectional datasets have been shared so far, the potential of longitudinal dataset sharing is even greater. Indeed, by focusing on within-subject effects, sharing of longitudinal data is less prone to problems linked to differences in data quality and scanning parameters between sites. An approach especially promising is the sharing of control group data in interventional studies to increase the sample size in developmental and ageing studies of healthy individuals. Another potential source of longitudinal neuroimaging datasets is the routinely acquired scans used to assess the health of NHPs (Basso et al., this issue; Balezeau et al., this issue). An increased sample size, combined with sharing of metadata (e.g. genetic information, early life history; for more details about metadata, see Poirier et al., this special issue) will allow future investigation of individual differences in brain development and ageing.

2.2 Specific strength: Developmental studies

Longitudinal MRI in NHPs is of considerable interest for investigating brain development from birth to adulthood or even prior to birth, thanks to foetal imaging. One major objective of such studies is to establish developmental trajectories in NHP species and to compare them across NHP species and with humans. Another goal is to develop experimental NHP models of human neurodevelopmental pathologies, thanks to either genetic mapping or pharmacological, environmental, or behavioural interventional approaches.

Several NHP longitudinal neurodevelopmental MRI databases have recently emerged (see Table 1). For instance, several rhesus macaque longitudinal MRI databases from postnatal to early adulthood can be found from different primate facilities mostly located in the USA (Scott et al. 2016; Young et al. 2017; Xia et al., 2020; Malkova, Heuer & Saunders 2006) but also in France (Rayson, et al., unpublished data) and in China (Liu et al., 2015), potentially describing distinct developmental trajectories of different sub-populations of rhesus macaques. Although less represented in comparison to rhesus macaques, similar longitudinal MRI brain data have also been

collected in baboons, including early postnatal (e.g. Becker et al., 2021) and foetal MRI brain images (Kochunov et al. 2010; Liu et al., 2008, 2010), as well as in marmosets, from infancy to adulthood (Sawiak et al., 2018, Seki et al., 2017). Such in vivo non-invasive approaches include structural T1- and/or T2-weighted imaging (T1w and/or T2w) and to a lesser extent, Diffusion Tensor Imaging (DTI) or Resting State functional MRI (RS-fMRI). To our knowledge, no such longitudinal neurodevelopmental database exists for squirrel monkeys (*Saimiri sp.*) or mouse lemurs (*Microcebus sp.*), two genera of increasing relevance to the field of neuroscience (Fritz et al., 2020; Rojo et al., 2021).

Table 1. Longitudinal brain development MRI databases in non-human primates.

Species	Centre	Publication	T1w/ T2w	DTI	RS- fMRI	Initial cohort Size	Age Range
<i>Callithrix jacchus</i> (marmoset, average lifespan, 5-7 years, max 12)	Brain Science Institute RIKEN, Wako, Japan	Seki et al., 2017	x			23*	1-30 months
	Cambridge University, UK	Sawiak et al. 2018	x			41*	3-27 months
<i>Macaca mulatta</i> (macaque, average lifespan, 15 years, max 35, gestation time, 168 days)	California National Primate Research Centre, USA	Scott et al. 2016	x			48	1-52 weeks
	University of Wisconsin, USA	Young et al. 2017; Xia et al., 2020	x	x		37	10-64 months
	National Institutes of Health, Bethesda, USA	Malkova, Heuer & Saunders 2006	x			7*	1 week to 4 years
	CNRS centre, Lyon, France	Rayson, Froesel, et al, unpublished	x	x	x	21	1.5 / 2.5 / 3.5 years
	Primate Centre of Academy of Sciences, Kunming, China	Liu et al., 2015	x	x		14*	6, 7, 8, 18, 16 months
	Yerkes National Primate Research Centre, USA	Shi et al., 2017 (Atlas)	x	x		40*	2 weeks, 3, 6, 12 months
Oregon National Primate Research Centre, USA	Liu et al., 2020	x	x		18*	85, 110, 135 gestation days	
<i>Papio anubis</i> (baboon, average lifespan, 25, max 37,	CNRS Primate Centre, Rousset, France	Becker et al., 2021	x	x	x	30	0-2; 8-10; 24-26 & 48 months
	Southwest National Primate Research Centre, USA	Kochunov et al. 2010	x			7*	17-25 gestation weeks

<i>gestation time, 26 weeks)</i>	<i>Columbia University, New York, USA</i>	<i>Liu et al., 2008, 2010</i>	<i>x</i>	<i>5 *</i>	<i>8-26 gestation weeks</i>
--------------------------------------	---	-----------------------------------	----------	------------	---------------------------------

* Number of subjects scanned varies across time points.

Alongside these developmental longitudinal datasets, a handful of NHP developmental atlases exist. Regarding the macaque, Shi et al. (2017) have produced a developmental atlas from birth to 12 months of age, providing a parcellation of cortical and subcortical structures as well as brain fibre tract data. Liu et al. (2020) have recently issued an atlas of the foetal macaque brain at 85, 110 and 135 days of gestation, providing T2-based, fractional anisotropy (FA) and apparent diffusion coefficient (ADC) templates. These developmental macaque atlases can be co-registered to the multiple adult anatomical, histological and functional macaque atlases available to date. Regarding marmosets, Seki et al. (2017) and Sawiak et al. (2018) provide age-specific brain templates. In these latter age-specific brain templates, cortical brain regions are defined based on co-registration with adult marmoset histological (Majka et al., 2016) and anatomical sections (Paxinos et al., 2012). There is clearly a need for expansion of such atlases, covering longer macaque age ranges, and including other species such as baboons, squirrel monkeys, or mouse lemurs.

Using MRI in NHPs to study brain development provides critical benefits over human MRI datasets. First, while developmental trajectories in humans are well documented for healthy children aged 4 years old and up (Giedd et al., 1999; Sowell et al., 2004; Raznahan et al. 2012), it has been extremely challenging to collect high quality motion-free images from children below 4 years old. As a result, this has limited our knowledge about brain development in this very dynamic and critical age range. Few human studies have been able to scan longitudinally subjects below the age of two years old (Knickmeyer et al., 2008, Gilmore et al., 2012, Nie et al., 2014). This issue can be addressed when scanning nonhuman primate subjects, thanks to close anaesthesia monitoring, allowing one to obtain motion-free images at all age classes from birth into adulthood.

Second, as is the case for other types of longitudinal studies in humans, recruiting and following the same homogeneous infant cohort in order to obtain individual MRI images at several time-points remains a critical challenge, especially when the study is expected to cover a broad age range, for example from birth into adulthood. As a result, the majority of paediatric neuroimaging studies are cross-sectional, that is to say not necessarily tracking all subjects, at all time points (e.g., Evans, 2006, Jernigan et al., 2016, Van Essen et al., 2013). In NHPs, there are theoretically no

such limitations, neither in terms of access to subjects, control of age homogeneity nor for the follow-up of intra-individual scanning at all predefined time points across development. The only limitation pertains to the capacity of reproduction in primate breeding centres, i.e. sample size.

Third, in order to secure the well-being of fully awake human infants and achieve motion-free brain images, the duration of data acquisition is often limited inside the MRI scanner, constraining the number of MRI multimodal sequences and their parameters (e.g. spatial resolution). In in-vivo NHP MRI acquisitions, securing the welfare and health of subjects by proper anaesthesia and physiological monitoring procedures under veterinary control allows a considerable increase in the time of a given MRI session. This provides unique opportunities for testing and developing more and much longer MRI sequences for better brain image quality and multimodal potentialities at this early stage of brain development. Additionally, this compensates for the decrease in the Signal-to-Noise Ratio of the MRI acquisitions incurred by the smaller NHP brain size (Kochunov and Duff Davis, 2010).

Fourth, there is now ample evidence that neurodevelopmental trajectories can be dramatically impacted by genetic (van Dyck et al., 2017) and non-genetic factors, thought to act via epigenetic mechanisms (Lester and Marsit, 2018). These non-genetic factors include social factors such as social deprivation during a so-called social critical period (Feldman, 2015; Sheridan and McLaughlin, 2014) and physical factors such as sensory deprivation, anaesthesia (Disma et al., 2018), antipsychotic or analgesic maternal drug history (Hjorth et al., 2019, EUROmediCAT Steering Group, 2015, Nordeng et al., 2017), and metabolic and nutritional history (Barks et al., 2019). However, the precise mechanisms and factors by which neurodevelopment is impacted are still a matter of intense research. In this respect, NHPs represent a unique and extremely valuable model in which social and physical environmental factors as well as genetic and epigenetic mechanisms can be controlled or modified from conception into adulthood. Such experimental designs can thus be used for highly controlled interventional developmental studies in which a group of animals is submitted to a specific physical or social early rearing history experimental manipulation and its brain developmental trajectory is compared to the one of a normal control group.

2.3 Specific strength: Ageing studies

The impact of ageing on brain structure and function is a topic of great interest which has been investigated mainly by a cross-sectional approach in humans and NHPs. Taking advantage of huge

datasets and the development of algorithms to automatically analyse thousands of MRI scans, recent studies in humans have documented age-related brain changes consistent across subjects (Cole et al., 2018; Kaufmann et al., 2019; Madan and Kensinger, 2018; Walhovd et al., 2011; Wang et al., 2016). However, cross-sectional studies of ageing suffer from two main limitations which can be overcome by using a longitudinal approach. First, cross-sectional studies do not allow distinguishing true ageing effects from cohort effects. Because cross-sectional ageing studies compare brain scans from groups of subjects born in different years (different cohorts), they cannot distinguish true ageing effects from brain differences induced by factors unrelated to ageing that changed between cohorts' birth years (e.g. nutrition of mothers during the prenatal period). Longitudinal studies allow the investigation of within-subject effects, which are immune to cohort effects, revealing true ageing effects. Secondly, cross-sectional studies, which can only reveal between-subject effects, treat individual differences unrelated to age as a source of noise. This approach is therefore ill-suited to investigate the source of individual differences in ageing effects. By capturing ageing processes in their within-subject effects, longitudinal studies allow for investigating individual variability by testing the interaction of between- and within-subject effects. To take full advantage of these methodological strengths, longitudinal studies must include multiple time points, covering the specific life period under investigation in each subject.

Despite these important advantages, longitudinal MRI studies are notoriously difficult to perform in humans. For pragmatic reasons, they can suffer from an important drop-out of participants and usually cover a short period of subjects' lives (usually less than 10 years) (Fjell et al., 2017; Kuo et al., 2020; Scahill et al., 2003; Storve et al., 2014; Wu et al., 2013). NHPs therefore offer a unique opportunity to investigate true ageing effects in the primate brain. The fact that their longevity is shorter than humans (one year in a macaque life is equivalent to 3-4 years in a human life) allows researchers to collect data at a much faster speed and to cover longer periods of their lives (relative to their lifespan). In addition, given that NHPs spend their whole life in a controlled environment, it limits the sources of variation in brain function and structure over time which are unrelated to ageing.

2.4 Specific strength: Interventional studies

The above paragraphs indicating the structural changes across the entire lifespan, underscore the critical need to control for age in longitudinal studies involving experimental manipulations such as disease models, rearing conditions, or training. Cross-sectional interventional studies in humans are limited by design, to comparisons between groups of subjects with different conditions, for

example the comparison of a group having a certain disease with a “normal” control group. Between relatively large groups, such cross-sectional studies may be able to identify a difference between groups if the variance within groups is sufficiently small. If a between-group difference can be observed using a cross-sectional comparison (i.e. between a patient and a control group), the critical question remains whether these observed group differences were due to pre-existing traits or a consequence of the disorder (Jedema et al 2020). Due to the fact that imaging data are not normally obtained prior to becoming a patient, this question remains very difficult to address in humans, although large prospective longitudinal human studies such as the ABCD study (Volkow et al 2018) and the IMAGEN study (Mascarell Maricic et al 2020) will help to address this issue. In substance use disorder studies, the question that arises with all cross-sectional studies is whether any observed difference in the drug abuse group is the consequence of prolonged drug use, or reflects a pre-existing condition that conferred greater vulnerability to drug use or even a combination of these factors.

In contrast to cross-sectional studies, longitudinal studies make within-subject comparisons, thereby reducing the impact of between-subject differences and nuisance variables. When combined with a proper age-matched control group, these studies control for any pre-existing conditions and permit causal attribution of the group by time interaction to the experimental manipulation. Given the similarity in brain structure, circuitry, and anatomical assignment of function as well as similarities in drug kinetics and adaptations with long term treatment, longitudinal NHP studies provide critically important preclinical data for complex human conditions. In addition, in longitudinal NHP imaging studies there is much better experimental control over the entire lifespan of the experimental and control groups compared to clinical populations (for example, substance abusers using a combination of multiple, different drugs versus a college student control group).

The value of NHPs to interventional longitudinal studies goes beyond preclinical studies and can serve to shed light on fundamental neuroscience research questions. For instance, NHPs can be trained to perform sophisticated cognitive tasks using operant conditioning techniques (Buffalo et al., 2019; Gámez et al., 2018). Macaques have been taught to categorize complex stimuli from all modalities (Freedman and Assad, 2016; Mendoza et al., 2018), to use abstract and dynamic rules to guide their behaviour (Miller and Cohen, 2001; Bartolo and Averbeck, 2020), to solve complex spatial cognition and game theory tasks (Georgopoulos et al., 1989; Crowe et al., 2005; Lee and Seo, 2016), and to learn to control brain machine interfaces with novel decoding requirements

(Sadler et al., 2014; Athalye et al., 2017; Golub et al., 2018). In addition, NHPs are an excellent model to obtain an evolutionary perspective on the neurobiology of cognitive traits originally considered human specific, such as language and music (Petkov and Ten Cate, 2019; Balezeau et al., 2020; Mendoza and Merchant, 2014; Merchant and Honing, 2014). Since the training periods in these elaborate tasks can take many months, the use of longitudinal imaging at different learning stages can reveal important structural changes in the grey and white matter of the underlying brain circuits. Because NHPs learn these complex tasks more slowly, there are more opportunities to scan them at different stages of learning and thus there is a better possibility to study the correlation between learning and structure. In contrast, because humans learn the complex tasks “too” quickly, they usually can only be scanned at stages before learning or after learning is complete.

Other applications of interventional longitudinal studies include investigating the impact of early rearing conditions onto brain development and cognition (Howell et al., 2019), as well as studying adult cortical plasticity following acute controlled brain damage or limb immobilization and subsequent rehabilitation or remediation strategies, whether pharmacological, behavioural, surgical or technological (e.g. brain-machine interfaces). The advent of genetically modified NHPs, whether marmosets (Kumita et al., 2019, Yoshimatsu et al., 2019) or macaques (Liu et al., 2016), opens the way to study gene control over both behaviour and brain structural and functional organization, in normal and pathological neurodevelopment, ageing and psychiatric/neurological disorder onset (Park and Silva, 2019). The goal is then to generalize multilevel or latent growth modelling approaches classically applied to the characterization of longitudinal brain trajectories (Mills and Tamnes, 2014) for comparison across the experimental and control groups.

3 Challenges of longitudinal MRI datasets

Despite numerous benefits, NHP longitudinal neuroimaging studies come with their own challenges. Some challenges are specific to particular applications, while others are generic. Here we describe what we consider to be the main challenges, related to experimental design, data acquisition, image processing, and statistical analyses.

3.1 Experimental design

Despite a reduction of the drop-out challenge characteristic of human longitudinal studies, NHP longitudinal studies are constrained by a limited number of subjects, for ethical (see section 2.1) and pragmatic reasons. In discovery science, when the biologically relevant minimum effect size is

not always known, power calculations are difficult to perform, making it difficult to justify exploratory studies involving large numbers of animals. The lack of universally accepted power calculators suitable for longitudinal designs only exacerbates this problem.

In addition, the number of available subjects is limited by the cost of NHP housing as well as the capacity of primate breeding centres. This last limitation is inflated in interventional neurodevelopment experiments which most often include, in addition to a test group, a control group to which the test group is compared. If the experimental intervention affects health parameters (e.g. increasing sensitivity to infections), the test group size might decrease with time. As a result, larger groups need to be provided at the beginning of the study, in anticipation of these losses. Growing evidence indicates that independent validation of the reported main effects is crucial when analysing longitudinal datasets (Herting et al., 2018). This can be done by multiple statistical approaches, such as independent tests on two independent subsamples of the data or leave one out or k-fold statistical cross validation procedures. The ability to perform such analyses is fully tied to cohort size.

For developmental studies, a related challenge is the ability to scan all individuals within the same age windows. This is especially critical in the very early months following birth in which brain morphology (in terms of myelination, overall brain size and gyrification) can change dramatically on a weekly basis. This can be complicated by the fact that births often happen at very short intervals in primate breeding centres, which in turn requires consistent access to the MRI facility. The presence of an MRI facility at the same location as the breeding centre tends to (partially) mitigate this problem.

3.2 Data acquisition

All longitudinal NHP studies require hardware and software to be held constant during the whole experiment, which can sometimes last years. However, NHP neuroimaging approaches are often characterised by customized equipment and acquisition protocols in a state of constant refinement (Milham et al., 2018, 2020). The need for stability of longitudinal studies can therefore be in conflict with the drive for methodological and technical improvement, especially when imaging facilities and equipment are shared by multiple research groups. Moreover, to assure stable image quality, the head of the animal needs to be consistently positioned in the scanner field across scans and various custom radiofrequency coils for NHPs have been presented in recent

years to help address this issue (Belcher et al 2013, Gilbert et al 2019, Quan et al 2020). A potential risk is that experimenters become more proficient over time at positioning the heads of animals relative to the coils and the magnet, which can have an impact on image quality, especially when one uses non-standardized arrangements of surface coils. For awake NHP imaging, an additional problem is the impact of animals' movements on image quality: if awake animals are not already well accustomed to the scanning environment and radio-frequency noise, the amplitude and frequency of their movements can decrease over time, resulting in a progressive improvement in data quality. It is therefore important to assess data quality *a posteriori*, and to identify outliers or check that data quality is not correlated with the longitudinal design (i.e. data quality increasing with time) (fig. 2).

Another challenge specific to developmental studies in anaesthetised NHPs is the choice of the anaesthetic drugs. Anaesthesia allows minimizing head and body motion (which itself can alter the homogeneity of the magnetic field) and scanning for longer periods of time, thus making it possible to acquire multimodal imaging data. However, anaesthesia also has the potential to critically interfere with early brain development, due to the fact that most anaesthetics enhance GABAergic neurotransmission and depress NMDA neuromodulation, two key players of cortical plasticity. In addition, most anaesthetics show a high degree of neurotoxicity on repetitive exposure, including isoflurane, increasing apoptosis of neurons and oligodendrocytes in immature animals if administered within the critical neurodevelopmental period in which synaptogenesis is maximal (Brambrink et al., 2012; Coleman et al., 2017; Hays & Deshpande, 2011; Noguchi et al., 2017; Schenning et al., 2017). Dexmedetomidine, which provides an optimal trade-off between neurotoxicity and sedation, may be considered as sufficient for short scanning sessions in the youngest individuals (Koo et al., 2014) and has been successfully used to monitor resting state activity under anaesthesia (Fukuda et al 2013, Brynildsen et al, 2017). In any case, it is recommended that the possible behavioural effects of repeated anaesthesia be tracked by an ethologist. Likewise, providing for a control group in which scanning (and thus anaesthesia) will only start after early childhood development is over should be envisioned. Whether scans on this control group should start before or after the important cortical changes associated with puberty is a matter for discussion. Alternatively, control groups could consist of independent individuals with no prior anaesthesia history, for each target age. The question of the interference of anaesthesia with neurodevelopmental trajectories thus turns out to be a scientific question in itself, of crucial impact on the neurodevelopmental longitudinal studies discussed here.

3.3 Image processing

3.3.1 Analyses in population-based standard space

For many group analyses, MR images need to be moved into a generic population-based reference space via a co-registration process referred to as Normalization. Such transformation of original images can be performed using linear or non-linear deformation and allows for comparison of brain images at the voxel level. Normalization is also needed to co-register individual brains onto an atlas of interest (cytoarchitectonics, myeloarchitectonics, fiber tracking etc.), in order to identify comparable (cortical) areas across subjects.

A number of longitudinal anatomical data processing methods or pipelines have been proposed, including, but not limited to, Ashburner and Ridgway (2013), Reuter et al. (2012), Holland and Dale (2011), and Avants et al. (2010). The pipelines usually include two normalization steps between individual scans of different subjects taken at multiple time points: one normalization step from individual scans to a subject-or age-specific template, and a second normalization step from these templates to a final common reference space for NHPs (such as INIA-19, Rohlfing et al 2019, NMT, Seidlitz et al. 2017), or a study specific average template that is independent of subject or age. The two normalization steps may iterate several times for the convergence of normalization parameters.

Longitudinal data processing aims to preserve within-subject changes for final statistical analyses. However, without special attention, many of the processing steps can introduce bias and influence the within-subject variance preservation. Several approaches for reducing bias have been introduced, such as symmetrical normalization to reduce interpolation bias (Avants et al., 2010; Holland and Dale, 2011; Nakamura et al., 2011; Reuter et al., 2011) and building robust, unbiased, within-subject templates based on median or mean values and simultaneous co-registration of all time-points (Ashburner and Ridgway, 2013; Reuter et al., 2012).

Most non-linear registration procedures attempt to minimize the difference between intensity values of each individual scan and the study-based template with a wide range of degrees-of-freedom parameters. Compared to linear registration, this difference-minimization procedure can improve the registration, but it can also introduce bias of the individual scan toward the template, affecting the final statistical tests or the final detection sensitivity.

While there is general agreement that using a subject-specific template is best for longitudinal studies during adulthood (Iglesias et al., 2016; Reuter et al., 2011, 2012; Ashburner et al., 2013),

no such consensus has been reached for developmental studies where both subject-specific and age-specific templates have been used in combination with different normalization procedures (Scott et al., 2016; Ball et al., 2019) to characterize the drastic changes during these more dynamic phases of the lifespan (fig. 3).

Here, we demonstrate using a simulation experiment of VBM data that the template choice is important when the pipeline contains both linear and non-linear registration steps to a final common space (see online Supplementary text). Specifically, the grey matter density (GMD) estimate after non-linear registration to the age-specific template is closer to the ground truth compared to the GMD estimate after non-linear registration to a subject-specific template (fig. 4). In addition, the variance of the GMD estimate following registration to the subject specific template is larger. Given the greater deviation from the ground truth as well as the larger variance, the GMD estimate after registration to the age-specific template will have greater sensitivity to detect real changes during neurodevelopment. Similar results are expected in longitudinal studies during old age, a life period also characterized by notable brain changes (Fig 3 phase C). In addition, our simulation demonstrates that the template choice does not seem to impact pipelines containing dual non-linear registrations steps to a final common template, although this may reduce overall sensitivity to detect real changes.

3.3.2 Analyses in subject-specific space

Some analyses do not require images to be normalized into a generic population-based reference space. For instance, measures of whole brain volume, total amount of grey matter, regional cortical volume (e.g. temporal lobe volume, occipital lobe) and subcortical volume can be extracted for each subject at each time point (Mills and Tamnes, 2014) without voxel-to-voxel correspondence. Overall and regional cortical thickness, surface areas and gyrification measures (e.g. gyrification index) can also be assessed (Mills and Tamnes, 2014). In this case, images can be left in their native space, decreasing potential for biases. Nevertheless, for longitudinal studies during early and mid-adulthood, when within-subject changes over time are expected to be small (compared to between-subject differences), specific data processing can improve detection of longitudinal changes in native space. For instance, FreeSurfer longitudinal pipeline computes subject-specific tissue probability maps to improve the segmentation results, and subject-specific meshes to improve surface extraction.

3.3.3 Limitations of human longitudinal pipelines

To date, longitudinal pipelines have been primarily developed for human MRI data. Such workflows include ANTs Longitudinal Cortical Thickness Pipeline (Tustison et al., 2017), FreeSurfer Longitudinal Pipeline (Reuter et al., 2011) and The Longitudinal registration and Longitudinal method (LL method) (Aubert-Broche, 2013). However, human pipelines directly applied to NHP data can generate errors at many steps (skull-stripping, subcortical labelling, segmentation of grey and white matter, particularly in visual and precentral cortex). These errors require manual intervention, which is time consuming and has the potential to bias the analysis. Another limitation of applying human pipelines is their primary focus on voxel-based or surface-based changes (for instance, FreeSurfer focuses on surface-based analyses, SPM and ANTs on voxel-based analyses). In NHP neuroimaging, both types of analyses are useful and ideally, a workflow for longitudinal NHP data would allow both. A new approach to automatically process cross-sectional and longitudinal NHP data is described elsewhere in this special issue (Garcia-Saldivar et al., this issue). Briefly, PREEMACS (pipeline for PREprocessing and Extraction of the MACaque brain Surface) is a set of tools taken from several image processing programs commonly used for human data and customized for rhesus monkey imaging, avoiding manual intervention. While this approach focuses on surface-based analyses, it should be possible to combine its pre-processing steps (including the automatic skull extraction and the debiasing approach) with voxel-based pipelines such as SPM or CAT (Gaser and Dahnke, 2016) to obtain voxel-based analyses without, or only minimal, manual intervention.

3.4 Statistical analyses

3.4.1 Separating between- and within-subject effects

Independent of the strategy used to process images (cf. section 3.2), longitudinal studies need to be analysed with statistical approaches specific to longitudinal designs, taking into account the dependency of repeated measures within subjects. In developmental and ageing studies, the null hypothesis is that there is no within-subject time effect (i.e. no age effect). In interventional studies, the null hypothesis is that within-subject effects are similar in different groups (i.e. no interaction between within-subject time effect and between-subject group effect). Testing within-subject effects can become an issue when the covariate of interest changes not only over time (i.e. within subjects) but also between subjects (within one group). Such a situation arises when researchers track age-related changes in subjects of different ages at the beginning of the study, a strategy often used to increase the sample size in ageing studies (e.g. Fjell et al., 2017; Storsve et

al., 2014; Tamnes et al., 2017; Wu et al., 2013). In this specific case, it is important to separate within- from between-subject effects in the statistical analysis. Only within-subject effects will reveal true and accurate ageing effects, since between-subject effects are potentially biased by cohort effects (cf. section 2.2) and by individual differences of genetic origin (especially when sample sizes are modest).

3.4.2 Parameter estimation

The next issue relates to the estimation of the fixed within-subject effects parameters. Here we briefly review three popular statistical tools used to analyse longitudinal data: Linear Mixed Effects modelling (LME) (Chen et al., 2013), Sandwich Estimator (SwE) models (Guillaume et al., 2014), and Permutation Analysis for Linear Models (PALM) (Winkler et al., 2014, 2016).

Linear mixed effects modelling (LME)

Linear mixed effects modelling can take into account multiple factors, including within-subject repeated measures factors, between-subject factors, or a mixture of both. LME has great flexibility in its modelling and estimation of the variance-covariance structures for both random effects and residuals. Traditional general linear models (such as ANOVA/ANCOVA) have much stricter assumptions on the variance-covariance structure and cannot model the covariance structure when there is random deviation of subjects at the levels of a random-effects factor. Due to the flexibility of LME modelling, it has been used extensively in a variety of experimental designs and for the description of all these various possibilities, that are outside the scope of this review, we refer the reader to detailed descriptions on LME modelling (Fitzmaurice et al., 2012, Chen et al. 2013, and Bates et al., 2015). Briefly, the response $\mathbf{y}_i \in \mathbb{R}^{N_i}$ (for subject i , and response vector length N_i) can be described as the sum of fixed effect $\boldsymbol{\beta} \in \mathbb{R}^P$ and random effect $\boldsymbol{\gamma}_i \in \mathbb{R}^Q$, i.e., $\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\beta} + \mathbf{Z}_i \boldsymbol{\gamma}_i + \boldsymbol{\epsilon}_i$ (with $\mathbf{X}_i \in \mathbb{R}^{N_i \times P}$ and $\mathbf{Z}_i \in \mathbb{R}^{N_i \times Q}$). The random effect $\boldsymbol{\gamma}_i$ follows a normal distribution with variance $\boldsymbol{\Psi} \in \mathbb{R}^{Q \times Q}$, i.e., $\boldsymbol{\gamma}_i \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Psi})$, and a within-subject residual $\boldsymbol{\epsilon}_i \in \mathbb{R}^{N_i}$ following a normal distribution with variance $\boldsymbol{\Sigma}_i \in \mathbb{R}^{N_i \times N_i}$. Overall, the response vector is modelled as a multi-variate normal distribution, $\mathbf{y}_i \sim \mathcal{N}(\mathbf{X}_i \boldsymbol{\beta}, \mathbf{Z}_i \boldsymbol{\Psi} \mathbf{Z}_i^T + \boldsymbol{\Sigma}_i)$. By taking advantage of the available LME estimation package *lme4* (Bates et al., 2015), it is possible to deal with various variance-covariance structures for the between-subject $\boldsymbol{\Psi}$ matrix and the within-subject $\boldsymbol{\Sigma}_i$ matrix, from the simplest diagonal matrix (i.e., no correlation among factors) to only requiring symmetry and positive definite. A method to use Akaike Information Criteria (AIC) or other criteria to select

among multiple available LME variance-covariance matrix structures has been presented (Müller et al 2013).

Practically for neuroimaging studies, 3dLME (Chen et al., 2013) and lme4 (Bates et al., 2015), can be used to model a variety of complex experimental designs and directly test the factor of interest without requiring implementation detail. For example, 3dLME allows the user just to specify the component of interest without building any design/contrast matrix. 3dLME can be applied in situations with a few response variables up to tens of thousands of voxels in voxel-wise fMRI datasets. However, such great modelling flexibility usually requires prior knowledge to choose certain models, which is not trivial. In addition, structural datasets with many voxels, as seen in typical high resolution NHP scans, can quickly exceed the capacity of even large computation clusters.

Sandwich estimator (SwE)

The Sandwich estimator (SwE) method focuses on population averages of longitudinal data and it has some adjustments for small sample size. SwE models implicitly handle within-subject covariance as opposed to LME models where users are required to explicitly specify or choose a variance-covariance matrix structure, and different random effects (such as including random subject-level intercepts/slopes or not), which is difficult in practice. From an optimization viewpoint, SwE separates within- and between-subject covariance such that the estimation can be as efficient as a generalized least square (GLS) method (thereby avoiding non-converging iterations as in the LME method).

Specifically, for the i^{th} individual ($i = 1, \dots, M$) with N_i observations, SwE models $\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \boldsymbol{\epsilon}_i^*$, in which $\mathbf{y}_i \in \mathbb{R}^{N_i}$ contains N_i observations, $\mathbf{X}_i \in \mathbb{R}^{N_i \times P}$ is the design matrix, $\boldsymbol{\beta} \in \mathbb{R}^P$ denotes P fixed effects, and $\boldsymbol{\epsilon}_i^*$ indicates an individual marginal error term. The fixed effects parameter $\boldsymbol{\beta}$ is estimated as $\hat{\boldsymbol{\beta}} = (\sum_i^M \mathbf{X}_i^T \mathbf{X}_i)^{-1} \sum_i^M \mathbf{X}_i^T \mathbf{y}_i$. Its covariance ($\mathbf{S} \in \mathbb{R}^{P \times P}$) is estimated as $\mathbf{S} = \text{var}\{\hat{\boldsymbol{\beta}}\} =$

$$\underbrace{(\sum_i^M \mathbf{X}_i^T \mathbf{X}_i)^{-1}}_{\text{Bread}} \underbrace{(\sum_i^M \mathbf{X}_i^T \hat{\mathbf{V}}_i \mathbf{X}_i)}_{\text{Meat}} \underbrace{(\sum_i^M \mathbf{X}_i^T \mathbf{X}_i)^{-1}}_{\text{Bread}}, \text{ in which } \hat{\mathbf{V}}_i \in \mathbb{R}^{N_i \times N_i} \text{ denotes the covariance of the } i^{\text{th}}$$

subject's fitting error/residual ($\mathbf{e}_i = \mathbf{y}_i - \mathbf{X}_i \hat{\boldsymbol{\beta}}$), i.e., $\hat{\mathbf{V}}_i = \mathbf{e}_i \mathbf{e}_i^T$. Such subject-wise error covariance is chosen to be substituted with group-wise homogeneous error covariance estimation for each group, i.e., a group g (having subjects index set as I_g) share the same fitting error covariance

$$\hat{\mathbf{V}}_g = \frac{1}{m} \sum_{i \in I_g} \mathbf{e}_i \mathbf{e}_i^T. \text{ For the inference on the parameter } \hat{\boldsymbol{\beta}}, \text{ SwE uses a Wald test } T =$$

$$(\mathbf{C}\hat{\boldsymbol{\beta}})^T (\mathbf{C}\mathbf{S}\mathbf{C}^T)^{-1} (\mathbf{C}\hat{\boldsymbol{\beta}}) / q \text{ for the null hypothesis } H_0: \mathbf{C}\boldsymbol{\beta} = 0, \text{ in which } q \text{ denotes the rank of the}$$

contrast matrix $\mathbf{C} \in \mathbb{R}^{P \times P}$. All the estimations are based on maximum likelihood estimation which requires a large sample number. The Wald test statistic may have a heavier tail with small samples rather than the usual χ_q^2 distribution. To adjust for small samples in the estimation, SwE can accommodate a correction factor for the raw residual \mathbf{e}_i before estimating $\hat{\mathbf{V}}_i$ but this requires the choice of an approximate statistic and F distribution as the null distribution: $\frac{v-q+1}{vq} (\mathbf{C}\hat{\boldsymbol{\beta}})^\top (\mathbf{C}\mathbf{S}\mathbf{C}^\top)^{-1} (\mathbf{C}\hat{\boldsymbol{\beta}}) \sim F(q, v-q+1)$, in which v denotes a degree of freedom parameter to be estimated and is advised to use group-wise homogeneous estimation.

Permutation Analysis of Linear Models (PALM)

The maximum likelihood estimators are asymptotical, require large samples, and assume certain distributions (usually Gaussian) in inference. These conditions are frequently not met, especially with small sample sizes generally seen in longitudinal NHP data. As a result, these methods might inflate the power or induce a higher false positive rate (FPR). Alternatively, permutation methods provide more precise control of the FPR. Among the permutation-based tools, PALM can accommodate longitudinal data with each subject as a whole permutation block. The precise control of FPR is the result of building the null distribution with many permutations. With certain distribution assumptions (such as Generalized Pareto Distribution), the required number of permutations can be reduced (500 permutations for a minimal P-value of 0.2%), compared to general permutation-based tests (~5000 permutations). Furthermore, comparisons between smaller numbers of subjects is possible. For example, there can be 252 permutations with a minimal P-value being 0.4%, with as few as 5 subjects for 2 groups with the assumption of different variances for 2 longitudinal time points.

In detail, given a linear model $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \boldsymbol{\epsilon}$ in which $\mathbf{y} \in \mathbb{R}^M$ is the observed data for M subjects, $\mathbf{X} \in \mathbb{R}^{M \times P}$ denotes the design matrix with regressors of interest, $\mathbf{Z} \in \mathbb{R}^{M \times Q}$ denotes the design matrix with nuisance regressors, $\boldsymbol{\beta} \in \mathbb{R}^P$, $\boldsymbol{\gamma} \in \mathbb{R}^Q$ denotes corresponding regression coefficients of interest and nuisance, and $\boldsymbol{\epsilon} \in \mathbb{R}^M$ is the residue. The null hypothesis is that there is no difference for a given contrast (such as a contrast between the estimation and 0, or a contrast between two subsets for a comparison between 2 groups), i.e., $\mathcal{H}_0: \mathbf{c}^\top \hat{\boldsymbol{\beta}} = 0$ for only one contrast $\mathbf{c} \in \mathbb{R}^P$, or $\mathbf{C}^\top \hat{\boldsymbol{\beta}} = 0$ for S contrasts matrix $\mathbf{C} \in \mathbb{R}^{P \times S}$ ($1 \leq S \leq P$). The permutability or exchangeability is that under the null assumption, the joint probability distribution function for the involved variables is not different between the 2 contrasts and can be found through permutations.

Anatomical data, such as volume, density, and thickness, are all bounded numbers (i.e., there is a physical upper bound and lower bound), making it difficult to meet the infinite distribution value range assumption required for the SwE method. Combined with better control of the FPR, the use of PALM to approximate the distribution is preferred for small sample sizes over the assumption of a specific distribution as in SwE.

Advantages and disadvantages of the three statistical approaches are summarized in table 2.

Table 2: Advantages and disadvantages of each approach.

	AFNI's 3dLME	SwE	PALM
Advantages	Most flexible and can deal with missing data. Intuitive on data input, no need for design matrix and contrast matrix.	Finer control in the model over within-subject variance for longitudinal design	Exact control over FPR, not dependent on specific distribution. Supports almost all neuroimaging file formats and classical multivariate inference such as multivariate analysis of variance (MANOVA) or multivariate analysis of covariance (MANCOVA), making joint analyses possible (Winkler et al., 2016, 2018).
Disadvantages	Anatomical measurements, which have a finite value range violate the assumption of infinity value range distribution. Voxel-wise calculations are computationally costly, often exceeding computing resources with high resolution structural scans. Flexible modelling makes model selection difficult. Asymptotic estimation requires large sample size.	Anatomical measurements, which have a finite value range violate the assumption of Infinity value range distribution. Even with small number correction, its maximum likelihood asymptotic estimation requires large sample size.	Requires careful design and contrast matrix for longitudinal studies. Permutations are computationally costly.

5. Conclusion

Longitudinal primate neuroimaging offers unique opportunities and promises to open exciting new research ventures. However, it is also associated with specific challenges. Some of these challenges can be circumvented but need to be anticipated. Considering the human and financial investment necessary to acquire longitudinal datasets, and in line with the '3Rs', sharing such datasets should be encouraged. However, considering the associated dilution of primacy of data origination, which is important to investigators' career progression, it should also be properly rewarded.

Data and Code Availability Statement

Data were used for illustration purpose only. The simulation code is available here:
<https://github.com/dawnsong/longitudinalVBMTemplateBias>

Declaration of competing interest

The authors declare no competing interests.

CRedit authorship contribution statement

Xiaowei Song: Formal analysis; Investigation; Writing - original draft.

Pamela García Saldivar: Formal analysis, Visualization, Writing – original draft, Writing – review and editing.

Nathan Kindred: Visualization, Writing – review and editing.

Yujiang Wang: Writing - review & editing.

Hugo Merchant: Visualization, Writing – original draft, Writing – review and editing.

Adrien Meguerditchian: Writing – original draft.

Yihong Yang: Writing - review & editing.

Elliot A. Stein: Writing - review & editing.

Charles W. Bradberry: Writing – review and editing

Suliann Ben Hamed: Writing – original draft, Writing – review and editing.

Hank P. Jedema: Conceptualisation, Writing – original draft, Writing – review and editing.

Colline Poirier: Conceptualisation, Writing – original draft, Writing – review and editing.

Acknowledgements

Colline Poirier was supported by a Newcastle University Research Fellowship. Nathan Kindred was supported by a DiMeN MRC studentship. Suliann Ben Hamed was supported by funding from the Centre National pour la Recherche Scientifique (CNRS) and LABEX CORTEX funding (ANR-11-LABX-0042) from the Université de Lyon, within the program Investissements d’Avenir (ANR-11-IDEX-

0007) operated by the French National Research Agency (ANR). Adrien Meguerditchian was supported by funding from the European Research Council under the European Union's Horizon 2020 research and innovation program grant agreement No 716931 - GESTIMAGE - ERC-2016-STG. Hugo Merchant was supported by Consejo Nacional de Ciencia y tecnología A1-S-8430. Pamela García Saldivar was supported by Consejo Nacional de Ciencia y tecnología studentship number 280464. The contributions of Xiaowei Song, Hank P. Jedema, Elliot Stein, Yihong Yang, and Charles W. Bradberry were supported by the Intramural Research Program of the National Institute on Drug Abuse of the National Institutes of Health.

References

Arcaro, M. J., Schade, P. F., & Livingstone, M. S., 2019. Body map proto-organization in newborn macaques. *Proc. Natl. Acad. of Sci. U.S.A.* 116 (49), 24861–24871. <https://doi.org/10.1073/pnas.1912636116>

Ashburner, J., & Ridgway, G. R., 2013. Symmetric diffeomorphic modeling of longitudinal structural MRI. *Front. Neurosci.* 6, 197. <https://doi.org/10.3389/fnins.2012.00197>

Athalye, V. R., Ganguly, K., Costa, R. M., & Carmena, J. M., 2017. Emergence of Coordinated Neural Dynamics Underlies Neuroprosthetic Learning and Skillful Control. *Neuron* 93(4), 955–970.e5. <https://doi.org/10.1016/j.neuron.2017.01.016>

Aubert-Broche, B., Fonov, V. S., García-Lorenzo, D., Mouiha, A., Guizard, N., Coupé, P., Eskildsen, S. F., & Collins, D. L., 2013. A new method for structural volume analysis of longitudinal brain MRI data and its application in studying the growth trajectories of anatomical brain structures in childhood. *NeuroImage* 82, 393–402. <https://doi.org/10.1016/j.neuroimage.2013.05.065>

Avants, B., Cook, P.A., McMillan, C., Grossman, M., Tustison, N.J., Zheng, Y., Gee, J.C., 2010. Sparse Unbiased Analysis of Anatomical Variance in Longitudinal Imaging, in: Jiang, T., Navab, N., Pluim, J.P.W., Viergever, M.A. (Eds.), *Medical Image Computing and Computer-Assisted*

Intervention – MICCAI 2010, Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, pp. 324–331. https://doi.org/10.1007/978-3-642-15705-9_40

Balezeau, F., Wilson, B., Gallardo, G., Dick, F., Hopkins, W., Anwander, A., Friederici, A. D., Griffiths, T. D., & Petkov, C. I., 2020. Primate auditory prototype in the evolution of the arcuate fasciculus. *Nat. Neurosci.* 23(5), 611–614. <https://doi.org/10.1038/s41593-020-0623-9>

Balezeau, F. et al. MRI monitoring of macaque monkeys in neuroscience: Case studies, resource and normative data comparisons. *NeuroImage*, this issue. <https://doi.org/10.1016/j.neuroimage.2021.117778>

Ball, G., & Seal, M. L., 2019. Individual variation in longitudinal postnatal development of the primate brain. *Brain Struct. Funct.* 224 (3), 1185–1201. <https://doi.org/10.1007/s00429-019-01829-5>

Barks, A., Hall, A. M., Tran, P. V., & Georgieff, M. K., 2019. Iron as a model nutrient for understanding the nutritional origins of neuropsychiatric disease. *Pediatr. Res.* 85 (2), 176–182. <https://doi.org/10.1038/s41390-018-0204-8>

Basso, M., Tansey, G., Frey, S., Leopold, D., Pope, W., Poirier C., Kastner, S. Using MRI to enhance the care and well-being of laboratory rhesus macaque monkeys. *Neuroimage*, this special issue.

Bates, D., Mächler, M., Bolker, B.M., Walker, S.C., 2015. Fitting Linear Mixed-Effects Models using lme4. *Journal of Statistical Software*, 1 67. <https://doi.org/10.18637/jss.v067.i01>

Becker, Y., Sein, J., Velly, L., Giacomino, L., Renaud, L., Lacoste, R., Anton, J.-L., Nazarian, B., Berne, C., & Meguerditchian, A. (2021). Early Left-Planum Temporale Asymmetry in Newborn Monkeys (*Papio anubis*): A Longitudinal Structural MRI Study at Two Stages of Development. *NeuroImage*, 227, 117575. <https://doi.org/10.1016/j.neuroimage.2020.117575>

Belcher, A. M., Yen, C. C., Stepp, H., Gu, H., Lu, H., Yang, Y., Silva, A. C., & Stein, E. A., 2013. Large-scale brain networks in the awake, truly resting marmoset monkey. *J. Neurosci.* 33(42), 16796–16804. <https://doi.org/10.1523/JNEUROSCI.3146-13.2013>

Brambrink, A. M., Back, S. A., Riddle, A., Gong, X., Moravec, M. D., Dissen, G. A., Creeley, C. E., Dikranian, K. T., & Olney, J. W., 2012. Isoflurane-induced apoptosis of oligodendrocytes in the neonatal primate brain. *Ann. Neurol.* 72(4), 525–535. <https://doi.org/10.1002/ana.23652>

Bryndtsen, J. K., Hsu, L. M., Ross, T. J., Stein, E. A., Yang, Y., & Lu, H., 2017. Physiological characterization of a robust survival rodent fMRI method. *Magn. Reson. Imaging* 35, 54–60. <https://doi.org/10.1016/j.mri.2016.08.010>

Buffalo, E. A., Movshon, J. A., & Wurtz, R. H., 2019. From basic brain research to treating human brain disorders. *Proc. Natl. Acad. Sci. U.S.A.* 116(52), 26167–26172. Advance online publication. <https://doi.org/10.1073/pnas.1919895116>

Chen, G., Saad, Z.S., Britton, J.C., Pine, D.S., Cox, R.W., 2013. Linear mixed-effects modeling approach to fMRI group analysis. *NeuroImage* 73, 176–190. <https://doi.org/10.1016/j.neuroimage.2013.01.047>

Cole, J. H., Ritchie, S. J., Bastin, M. E., Valdés Hernández, M. C., Muñoz Maniega, S., Royle, N., Corley, J., Pattie, A., Harris, S. E., Zhang, Q., Wray, N. R., Redmond, P., Marioni, R. E., Starr, J. M., Cox, S. R., Wardlaw, J. M., Sharp, D. J., & Deary, I. J., 2018. Brain age predicts mortality. *Mol. Psychiatry* 23 (5), 1385–1392. <https://doi.org/10.1038/mp.2017.62>

Coleman, K., Robertson, N. D., Dissen, G. A., Neuringer, M. D., Martin, L. D., Cuzon Carlson, V. C., Kroenke, C., Fair, D., & Brambrink, A. M., 2017. Isoflurane Anesthesia Has Long-term Consequences on Motor and Behavioral Development in Infant Rhesus Macaques. *Anesthesiology* 126(1), 74–84. <https://doi.org/10.1097/ALN.0000000000001383>

Crowe, D. A., Averbeck, B. B., Chafee, M. V., & Georgopoulos, A. P., 2005. Dynamics of parietal neural activity during spatial cognitive processing. *Neuron* 47(6), 885–891. <https://doi.org/10.1016/j.neuron.2005.08.005>

Disma, N., O'Leary, J. D., Loepke, A. W., Brambrink, A. M., Becke, K., Clausen, N. G., De Graaff, J. C., Liu, F., Hansen, T. G., McCann, M. E., Salorio, C. F., Soriano, S., Sun, L. S., Szmuk, P., Warner, D. O., Vutskits, L., & Davidson, A. J., 2018. Anesthesia and the developing brain: A way forward for laboratory and clinical research. *Paediatr. Anaesth.* 28 (9), 758–763. <https://doi.org/10.1111/pan.13455>

EUROmedicAT Steering Group, 2015. EUROmedicAT Recommendations for European Pharmacovigilance concerning safety of medication use in pregnancy. *Pharmacoepidemiol. Drug Saf.* 24 Suppl 2, 3–7. <https://doi.org/10.1002/pds.3866>

Esteban, O., Birman, D., Schaer, M., Koyejo, O. O., Poldrack, R. A., & Gorgolewski, K. J., 2017. MRIQC: Advancing the automatic prediction of image quality in MRI from unseen sites. *PLoS one* 12(9), e0184661. <https://doi.org/10.1371/journal.pone.0184661>

Evans, A. C., & Brain Development Cooperative Group, 2006. The NIH MRI study of normal brain development. *Neuroimage* 30 (1), 184–202. <https://doi.org/10.1016/j.neuroimage.2005.09.068>

Feldman R., 2015. Sensitive periods in human social development: New insights from research on oxytocin, synchrony, and high-risk parenting. *Dev. Psychopathol.* 27 (2), 369–395. <https://doi.org/10.1017/S0954579415000048>

Fitzmaurice, G.M., Laird, N.M., Ware, J.H., 2012. *Applied Longitudinal Analysis*. vol. 998. John Wiley & Sons.

Fjell, A. M., Sneve, M. H., Grydeland, H., Storsve, A. B., Amlien, I. K., Yendiki, A., & Walhovd, K. B., 2017. Relationship between structural and functional connectivity change across the adult lifespan: A longitudinal investigation. *Hum. Brain Mapp.* 38(1), 561–573. <https://doi.org/10.1002/hbm.23403>

Freedman, D. J., & Assad, J. A., 2016. Neuronal Mechanisms of Visual Categorization: An Abstract View on Decision Making. *Annu. Rev. Neurosci.* 39, 129–147. <https://doi.org/10.1146/annurev-neuro-071714-033919>

Fritz R.G., Zimmermann E., Meier M., Mestre-Francés N., Radespiel U., & Schmidtke D., 2020. Neurobiological substrates of animal personality and cognition in a nonhuman primate (*Microcebus murinus*). *Brain Behav.* 2020 Sep; 10(9): e01752. <https://doi.org/10.1002/brb3.1752>

Fukuda, M., Vazquez, A. L., Zong, X., & Kim, S. G., 2013. Effects of the α_2 -adrenergic receptor agonist dexmedetomidine on neural, vascular and BOLD fMRI responses in the somatosensory cortex. *Eur. J. Neurosci.* 37(1), 80–95. <https://doi.org/10.1111/ejn.12024>

Gámez, J., Yc, K., Ayala, Y. A., Dotov, D., Prado, L., & Merchant, H., 2018. Predictive rhythmic tapping to isochronous and tempo changing metronomes in the nonhuman primate. *Ann. N.Y. Acad. Sci.* 10.1111/nyas.13671. Advance online publication. <https://doi.org/10.1111/nyas.13671>

Garcia-Saldivar, P., Garimella, A., Garza-Villareal, E., Mendez F., Concha L., Merchant H., 2020. PREEMACS: Pipeline for Preprocessing and Extraction of the Macaque Brain Surface, Neuroimage.

Gaser, C., Dahnke, R., 2016. CAT-a computational anatomy toolbox for the analysis of structural MRI data. Presented at the 22nd Annual Meeting of the Organization for Human Brain Mapping, Geneva, Switzerland.

Georgopoulos, A.P., Lurito J.T., Petrides, M., Schwartz, A.B., Massey, J.T., 1989. Mental rotation of the neuronal population vector. *Science* 243, 234-236. doi:10.1126/science.2911737

Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., Zijdenbos, A., Paus, T., Evans, A. C., & Rapoport, J. L., 1999. Brain development during childhood and adolescence: a longitudinal MRI study. *Nat. Neurosci.* 2 (10), 861–863. <https://doi.org/10.1038/13158>

Gilbert, K. M., Schaeffer, D. J., Gati, J. S., Klassen, L. M., Everling, S., & Menon, R. S., 2019. Open-source hardware designs for MRI of mice, rats, and marmosets: Integrated animal holders and radiofrequency coils. *J. Neurosci. Methods* 312, 65–72. <https://doi.org/10.1016/j.jneumeth.2018.11.015>

Gilmore, J. H., Shi, F., Woolson, S. L., Knickmeyer, R. C., Short, S. J., Lin, W., Zhu, H., Hamer, R. M., Styner, M., & Shen, D., 2012. Longitudinal development of cortical and subcortical gray matter from birth to 2 years. *Cereb. Cortex* 22(11), 2478–2485. <https://doi.org/10.1093/cercor/bhr327>

Golub, M. D., Sadtler, P. T., Oby, E. R., Quick, K. M., Ryu, S. I., Tyler-Kabara, E. C., Batista, A. P., Chase, S. M., & Yu, B. M., 2018. Learning by neural reassociation. *Nat. Neurosci.* 21(4), 607–616. <https://doi.org/10.1038/s41593-018-0095-3>

Guillaume, B., Hua, X., Thompson, P. M., Waldorp, L., Nichols, T. E., & Alzheimer's Disease Neuroimaging Initiative, 2014. Fast and accurate modelling of longitudinal and repeated measures neuroimaging data. *Neuroimage* 94, 287–302. <https://doi.org/10.1016/j.neuroimage.2014.03.029>

Hays, S.R., Deshpande, J.K., 2011. Newly postulated neurodevelopmental risks of pediatric anesthesia. *Curr Neurol Neurosci Rep.* 11(2), 205-210. doi: 10.1007/s11910-010-0177-4.

Herting, M. M., Gautam, P., Chen, Z., Mezher, A., & Vetter, N. C., 2018. Test-retest reliability of longitudinal task-based fMRI: Implications for developmental studies. *Dev. Cogn. Neurosci.* 33, 17–26. <https://doi.org/10.1016/j.dcn.2017.07.001>

Hjorth, S., Bromley, R., Ystrom, E., Lupattelli, A., Spigset, O., & Nordeng, H., 2019. Use and validity of child neurodevelopment outcome measures in studies on prenatal exposure to psychotropic and analgesic medications - A systematic review. *PLoS One* 14 (7), e0219778. <https://doi.org/10.1371/journal.pone.0219778>

Holland, D., & Dale, A.M., 2011. Nonlinear registration of longitudinal images and measurement of change in regions of interest. *Medical Image Analysis, Special section on IPMI 2009* 15, 489–497. <https://doi.org/10.1016/j.media.2011.02.005>

Howell, B. R., Ahn, M., Shi, Y., Godfrey, J. R., Hu, X., Zhu, H., Styner, M., & Sanchez, M. M., 2019. Disentangling the effects of early caregiving experience and heritable factors on brain white matter development in rhesus monkeys. *Neuroimage* 197, 625–642. <https://doi.org/10.1016/j.neuroimage.2019.04.013>

Iglesias, J. E., Van Leemput, K., Augustinack, J., Insausti, R., Fischl, B., Reuter, M., & Alzheimer's Disease Neuroimaging Initiative, 2016. Bayesian longitudinal segmentation of hippocampal substructures in brain MRI using subject-specific atlases. *Neuroimage* 141, 542–555. <https://doi.org/10.1016/j.neuroimage.2016.07.020>

Jedema, H. P., Song, X., Aizenstein, H. J., Bonner, A. R., Stein, E. A., Yang, Y. and Bradberry, C. W. (2020). Long-Term Cocaine Self-administration Produces Structural Brain Changes That Correlate With Altered Cognition. *Biol Psychiatry* 89, 376. <https://doi.org/10.1016/j.biopsych.2020.08.008>

Jernigan, T. L., Brown, T. T., Hagler, D. J., Jr, Akshoomoff, N., Bartsch, H., Newman, E., Thompson, W. K., Bloss, C. S., Murray, S. S., Schork, N., Kennedy, D. N., Kuperman, J. M., McCabe, C., Chung, Y., Libiger, O., Maddox, M., Casey, B. J., Chang, L., Ernst, T. M., Frazier, J. A., ... Pediatric Imaging, Neurocognition and Genetics Study, 2016. The Pediatric Imaging, Neurocognition, and Genetics (PING) Data Repository. *Neuroimage* 124 (Pt B), 1149–1154. <https://doi.org/10.1016/j.neuroimage.2015.04.057>

Kaufmann, T., van der Meer, D., Doan, N. T., Schwarz, E., Lund, M. J., Agartz, I., Alnæs, D., Barch, D. M., Baur-Streubel, R., Bertolino, A., Bettella, F., Beyer, M. K., Bøen, E., Borgwardt, S., Brandt, C. L., Buitelaar, J., Celius, E. G., Cervenka, S., Conzelmann, A., Córdova-Palomera, A., ... Westlye, L. T., 2019. Common brain disorders are associated with heritable patterns of apparent aging of the brain. *Nat. Neurosci.* 22 (10), 1617–1623. <https://doi.org/10.1038/s41593-019-0471-7>

Knickmeyer, R. C., Gouttard, S., Kang, C., Evans, D., Wilber, K., Smith, J. K., Hamer, R. M., Lin, W., Gerig, G., & Gilmore, J. H., 2008. A structural MRI study of human brain development from birth to 2 years. *J. Neurosci.* 28 (47), 12176–12182. <https://doi.org/10.1523/JNEUROSCI.3479-08.2008>

Kochunov, P., Castro, C., Davis, D., Dudley, D., Brewer, J., Zhang, Y., Kroenke, C. D., Purdy, D., Fox, P. T., Simerly, C., & Schatten, G., 2010. Mapping primary gyrogenesis during fetal development in primate brains: high-resolution in utero structural MRI of fetal brain development in pregnant baboons. *Front. Neurosci.* 4, 20. <https://doi.org/10.3389/fnins.2010.00020>

Kochunov, P., & Duff Davis, M., 2010. Development of structural MR brain imaging protocols to study genetics and maturation. *Methods* 50(3), 136–146. <https://doi.org/10.1016/j.ymeth.2009.08.002>

Koo, E., Oshodi, T., Meschter, C., Ebrahimnejad, A., & Dong, G., 2014. Neurotoxic effects of dexmedetomidine in fetal cynomolgus monkey brains. *J. Toxicol. Sci.* 39 (2), 251–262. <https://doi.org/10.2131/jts.39.251>

Kumita, W., Sato, K., Suzuki, Y., Kurotaki, Y., Harada, T., Zhou, Y., Kishi, N., Sato, K., Aiba, A., Sakakibara, Y., Feng, G., Okano, H., & Sasaki, E., 2019. Efficient generation of Knock-in/Knock-out marmoset embryo via CRISPR/Cas9 gene editing. *Sci. Rep.* 9 (1), 12719. <https://doi.org/10.1038/s41598-019-49110-3>

Kuo, P.-L., Schrack, J. A., Shardell, M. D., Levine, M., Moore, A. Z. et al., 2020. A roadmap to build a phenotypic metric of ageing: insights from the Baltimore Longitudinal Study of Aging. *J. Intern. Med.* 287(4):373-394. <https://doi.org/10.1111/joim.13024>

Lee, D., & Seo, H., 2016. Neural Basis of Strategic Decision Making. *Trends Neurosci.* 39(1), 40–48. <https://doi.org/10.1016/j.tins.2015.11.002>

Lester, B. M., & Marsit, C. J., 2018. Epigenetic mechanisms in the placenta related to infant neurodevelopment. *Epigenomics* 10 (3), 321–333. <https://doi.org/10.2217/epi-2016-0171>

Liu, G., Liang, K.-Y., 1997. Sample Size Calculations for Studies with Correlated Observations. *Biometrics* 53, 937–947. <https://doi.org/10.2307/2533554>

Liu, C., Tian, X., Liu, H., Mo, Y., Bai, F., Zhao, X., Ma, Y., & Wang, J., 2015. Rhesus monkey brain development during late infancy and the effect of phencyclidine: a longitudinal MRI and DTI study. *Neuroimage* 107, 65–75. <https://doi.org/10.1016/j.neuroimage.2014.11.056>

Liu, F., Garland, M., Duan, Y., Stark, R. I., Xu, D., Dong, Z., Bansal, R., Peterson, B. S., & Kangarlu, A., 2008. Study of the development of fetal baboon brain using magnetic resonance imaging at 3 Tesla. *Neuroimage* 40 (1), 148–159. <https://doi.org/10.1016/j.neuroimage.2007.11.021>

Liu, F., Garland, M., Duan, Y., Stark, R. I., Xu, D., Bansal, R., Dong, Z., Peterson, B. S., & Kangarlu, A., 2010. Techniques for in utero, longitudinal MRI of fetal brain development in baboons at 3T. *Methods* 50 (3), 147–156. <https://doi.org/10.1016/j.ymeth.2009.03.019>

Liu, Z., Li, X., Zhang, J. T., Cai, Y. J., Cheng, T. L., Cheng, C., Wang, Y., Zhang, C. C., Nie, Y. H., Chen, Z. F., Bian, W. J., Zhang, L., Xiao, J., Lu, B., Zhang, Y. F., Zhang, X. D., Sang, X., Wu, J. J., Xu, X., Xiong, Z. Q., ... Qiu, Z., 2016. Autism-like behaviours and germline transmission in transgenic monkeys overexpressing MeCP2. *Nature* 530(7588), 98–102. <https://doi.org/10.1038/nature16533>

Liu, Z., Wang, X., Newman, N., Grant, K. A., Studholme, C., & Kroenke, C. D., 2020. Anatomical and diffusion MRI brain atlases of the fetal rhesus macaque brain at 85, 110 and 135 days gestation. *Neuroimage* 206, 116310. <https://doi.org/10.1016/j.neuroimage.2019.116310>

Madan, C. R., & Kensinger, E. A., 2018. Predicting age from cortical structure across the lifespan. *Eur.J. Neurosci.* 47 (5), 399–416. <https://doi.org/10.1111/ejn.13835>

Majka, P., Chaplin, T. A., Yu, H. H., Tolpygo, A., Mitra, P. P., Wójcik, D. K., & Rosa, M. G., 2016. Towards a comprehensive atlas of cortical connections in a primate brain: Mapping tracer injection studies of the common marmoset into a reference digital template. *J. Comp. Neurol.* 524(11), 2161–2181. <https://doi.org/10.1002/cne.24023>

Malkova, L., Heuer, E., & Saunders, R. C., 2006. Longitudinal magnetic resonance imaging study of rhesus monkey brain development. *Eur. J. Neurosci.* 24 (11), 3204–3212. <https://doi.org/10.1111/j.1460-9568.2006.05175.x>

Mascarell Maricic, L., H. Walter, A. Rosenthal, S. Ripke, E. B. Quinlan, T. Banaschewski, G. J. Barker, A. L. W. Bokde, U. Bromberg, C. Buchel, S. Desrivieres, H. Flor, V. Frouin, H. Garavan, B. Itterman, J. L. Martinot, M. P. Martinot, F. Nees, D. P. Orfanos, T. Paus, L. Poustka, S. Hohmann, M. N. Smolka, J. H. Frohner, R. Whelan, J. Kaminski, G. Schumann, A. Heinz and I. consortium (2020). "The IMAGEN study: a decade of imaging genetics in adolescents." *Mol Psychiatry* 25(11): 2648-2671. <https://doi.org/10.1038/s41380-020-0822-5>

Mendoza, G., Méndez, J. C., Pérez, O., Prado, L., & Merchant, H., 2018. Neural basis for categorical boundaries in the primate pre-SMA during relative categorization of time intervals. *Nat. Commun.* 9(1), 1098. <https://doi.org/10.1038/s41467-018-03482-8>

Mendoza, G., & Merchant, H., 2014. Motor system evolution and the emergence of high cognitive functions. *Prog. Neurobiol.* 122, 73–93. <https://doi.org/10.1016/j.pneurobio.2014.09.001>

Merchant, H., & Honing, H., 2014. Are non-human primates capable of rhythmic entrainment? Evidence for the gradual audiomotor evolution hypothesis. *Front. Neurosci.* 7, 274. <https://doi.org/10.3389/fnins.2013.00274>

Milham, M. P., Ai, L., Koo, B., Xu, T., Amiez, C., Balezeau, F., Baxter, M. G., Blezer, E., Brochier, T., Chen, A., Crosson, P. L., Damatac, C. G., Dehaene, S., Everling, S., Fair, D. A., Fleysher, L., Freiwald, W., Froudast-Walsh, S., Griffiths, T. D., Guedj, C., ... Schroeder, C. E., 2018. An Open Resource for Non-human Primate Imaging. *Neuron* 100(1), 61–74.e2. <https://doi.org/10.1016/j.neuron.2018.08.039>

Milham, M., Petkov, C.I., Margulies, D.S., et al. 2020. Accelerating the evolution of nonhuman primate neuroimaging: The PRIMatE data exchange (PRIME-DE) global collaboration workshop and consortium. *Neuron* 105, 600-603. doi: <https://doi.org/10.1016/j.neuron.2019.12.023>

Miller, E. K., & Cohen, J. D., 2001. An integrative theory of prefrontal cortex function. *Ann. Rev. Neurosci.* 24, 167–202. <https://doi.org/10.1146/annurev.neuro.24.1.167>

Mills, K. L., & Tamnes, C. K., 2014. Methods and considerations for longitudinal structural brain imaging analysis across development. *Dev. Cogn. Neurosci.* 9, 172–190. <https://doi.org/10.1016/j.dcn.2014.04.004>

Mitchell, A., Basso, M., Hartig, R., Jarrett, W., Kastner, S., Poirier, C., Roe, A., 2021. International primate neuroscience research regulation, public engagement and transparency opportunities. *Neuroimage*, this special issue. 10.1016/j.neuroimage.2020.117700

Müller, S., Scealy, J.L., Welsh, A.H., 2013. Model Selection in Linear Mixed Models. *Statist. Sci.* 28, 135–167. <https://doi.org/10.1214/12-STS410>

Nakamura, K., Fox, R., Fisher, E., 2011. CLADA: Cortical longitudinal atrophy detection algorithm. *NeuroImage* 54, 278–289. <https://doi.org/10.1016/j.neuroimage.2010.07.052>

Nie, J., Li, G., Wang, L., Shi, F., Lin, W., Gilmore, J. H., & Shen, D., 2014. Longitudinal development of cortical thickness, folding, and fiber density networks in the first 2 years of life. *Hum. Brain Mapp.* 35(8), 3726–3737. <https://doi.org/10.1002/hbm.22432>

Noguchi, K. K., Johnson, S. A., Dissen, G. A., Martin, L. D., Manzella, F. M., Schenning, K. J., Olney, J. W., & Brambrink, A. M., 2017. Isoflurane exposure for three hours triggers apoptotic cell death in neonatal macaque brain. *Br. J. Anaesth.* 119(3), 524–531. <https://doi.org/10.1093/bja/aex123>

Nordeng, H., Lupattelli, A., & Wood, M., 2017. Prenatal exposure to antidepressants and increased risk of psychiatric disorders. *BMJ* 358, j3950. <https://doi.org/10.1136/bmj.j3950>

Park, J. E., & Silva, A. C., 2019. Generation of genetically engineered non-human primate models of brain function and neurological disorders. *Am. J. of Primatol.* 81 (2), e22931. <https://doi.org/10.1002/ajp.22931>

Paxinos, G., Watson, C., Petrides, M., Rosa, M., Tokuno, H., 2012. The marmoset brain in stereotaxic coordinates. Amsterdam, London: Academic.

Petkov, C. I., & Ten Cate, C., 2019. Structured Sequence Learning: Animal Abilities, Cognitive Operations, and Language Evolution. *Top. Cogn. Sci.* 10.1111/tops.12444. Advance online publication. <https://doi.org/10.1111/tops.12444>

Phillips, K. A., Bales, K. L., Capitanio, J. P., Conley, A., Czoty, P. W., 't Hart, B. A., Hopkins, W. D., Hu, S. L., Miller, L. A., Nader, M. A., Nathanielsz, P. W., Rogers, J., Shively, C. A., & Voytko, M. L., 2014. Why primate models matter. *Am. J. Primatol.* 76(9), 801–827. <https://doi.org/10.1002/ajp.22281>

Poirier, C., Ben Hamed, S., Garcia-Saldivar, P., Kwok, S.C., Meguerditchian, A., Merchant, H., Wells, S., Rogers, J. & Fox, A.S. 2021. Beyond MRI: on the scientific value of combining non-human primate neuroimaging with metadata. *Neuroimage*, this issue. <https://doi.org/10.1016/j.neuroimage.2020.117679>

Quan, Z., Gao, Y., Qu, S., Wang, X., Friedman, R. M., Chernov, M. M., Kroenke, C. D., Roe, A. W., & Zhang, X., 2020. A 16-channel loop array for in vivo macaque whole-brain imaging at 3 T. *Magn. Reson. Imaging* 68, 167–172. <https://doi.org/10.1016/j.mri.2020.02.008>

Raznahan, A., Greenstein, D., Lee, N. R., Clasen, L. S., & Giedd, J. N., 2012. Prenatal growth in humans and postnatal brain maturation into late adolescence. *Proc. Natl. Acad. of Sci. U.S.A.* 109 (28), 11366–11371. <https://doi.org/10.1073/pnas.1203350109>

Reuter, M., & Fischl, B., 2011. Avoiding asymmetry-induced bias in longitudinal image processing. *Neuroimage* 57 (1), 19–21. <https://doi.org/10.1016/j.neuroimage.2011.02.076>

Reuter, M., Schmansky, N. J., Rosas, H. D., & Fischl, B., 2012. Within-subject template estimation for unbiased longitudinal image analysis. *Neuroimage* 61 (4), 1402–1418. <https://doi.org/10.1016/j.neuroimage.2012.02.084>

Roelfsema, P. R., & Treue, S., 2014. Basic neuroscience research with nonhuman primates: a small but indispensable component of biomedical research. *Neuron* 82(6), 1200–1204. <https://doi.org/10.1016/j.neuron.2014.06.003>

Rohlfing, T., Kroenke, C.D., Sullivan, E.V., Bowden, D.M., Grant, K.A., 2012. The INIA19 template and NeuroMaps atlas for primate brain image parcellation and spatial normalization. *Front. Neuroinform.* 6, 27. <https://doi.org/10.3389/fninf.2012.00027>

Royo, J., Forkel, S.J., Pouget, P., Thiebaut de Schotten, M. Saimiri model in Clinical Neuroscience Research, Preprint 10.13140/RG.2.2.34359.55200/1

Sadtler, P. T., Quick, K. M., Golub, M. D., Chase, S. M., Ryu, S. I., Tyler-Kabara, E. C., Yu, B. M., & Batista, A. P., 2014. Neural constraints on learning. *Nature* 512(7515), 423–426. <https://doi.org/10.1038/nature13665>

Sawiak, S. J., Shiba, Y., Oikonomidis, L., Windle, C. P., Santangelo, A. M., Grydeland, H., Cockcroft, G., Bullmore, E. T., & Roberts, A. C., 2018. Trajectories and Milestones of Cortical and Subcortical Development of the Marmoset Brain From Infancy to Adulthood. *Cereb. Cortex* 28(12), 4440–4453. <https://doi.org/10.1093/cercor/bhy256>

Scahill, R.I., Frost, C., Jenkins, R., L Whitwell, J. L., Rossor, M. N., Nick C Fox, N. C., 2003. A longitudinal study of brain volume changes in normal aging using serial registered magnetic resonance imaging. *Arch. Neurol.* 60(7), 989-994. <https://doi.org/10.1001/archneur.60.7.989>.

Schenning, K. J., Noguchi, K. K., Martin, L. D., Manzella, F. M., Cabrera, O. H., Dissen, G. A., & Brambrink, A. M., 2017. Isoflurane exposure leads to apoptosis of neurons and oligodendrocytes in 20- and 40-day old rhesus macaques. *Neurotoxicol. Teratol.* 60, 63–68. <https://doi.org/10.1016/j.ntt.2016.11.006>

Scott, J. A., Grayson, D., Fletcher, E., Lee, A., Bauman, M. D., Schumann, C. M., Buonocore, M. H., & Amaral, D. G., 2016. Longitudinal analysis of the developing rhesus monkey brain using magnetic resonance imaging: birth to adulthood. *Brain Struct. Funct.* 221 (5), 2847–2871. <https://doi.org/10.1007/s00429-015-1076-x>

Seidlitz, J., Sponheim, C., Glen, D., Ye, F.Q., Saleem, K.S., Leopold, D.A., Ungerleider, L., Messinger, A., 2017. A population MRI brain template and analysis tools for the macaque. *NeuroImage*. <https://doi.org/10.1016/j.neuroimage.2017.04.063>

Seki, F., Hikishima, K., Komaki, Y., Hata, J., Uematsu, A., Okahara, N., Yamamoto, M., Shinohara, H., Sasaki, E., & Okano, H., 2017. Developmental trajectories of macroanatomical structures in common marmoset brain. *Neuroscience* 364, 143–156. <https://doi.org/10.1016/j.neuroscience.2017.09.021>

Sheridan, M. A., & McLaughlin, K. A., 2014. Dimensions of early experience and neural development: deprivation and threat. *Trends Cogn. Sci.* 18 (11), 580–585. <https://doi.org/10.1016/j.tics.2014.09.001>

Shi, Y., Budin, F., Yapuncich, E., Rumple, A., Young, J. T., Payne, C., Zhang, X., Hu, X., Godfrey, J., Howell, B., Sanchez, M. M., & Styner, M. A., 2017. UNC-Emory Infant Atlases for Macaque Brain Image Analysis: Postnatal Brain Development through 12 Months. *Front. Neurosci.* 10, 617. <https://doi.org/10.3389/fnins.2016.00617>

Sowell, E. R., Thompson, P. M., Leonard, C. M., Welcome, S. E., Kan, E., & Toga, A. W., 2004. Longitudinal mapping of cortical thickness and brain growth in normal children. *J. Neurosci.* 24 (38), 8223–8231. <https://doi.org/10.1523/JNEUROSCI.1798-04.2004>

Storsve, A. B., Fjell, A. M., Tamnes, C. K., Westlye, L. T., Overbye, K., Aasland, H. W., & Walhovd, K. B., 2014. Differential longitudinal changes in cortical thickness, surface area and volume across the adult life span: regions of accelerating and decelerating change. *J. Neurosci.* 34(25), 8488–8498. <https://doi.org/10.1523/JNEUROSCI.0391-14.2014>

Suckling, J., J. Henty, C. Ecker, S. C. Deoni, M. V. Lombardo, S. Baron-Cohen, P. Jezzard, A. Barnes, B. Chakrabarti, C. Ooi, M. C. Lai, S. C. Williams, D. G. Murphy, E. Bullmore and M. A. Consortium (2014). Are power calculations useful? A multicentre neuroimaging study. *Hum Brain Mapp* 35(8): 3569-3577.

Tamnes, C. K., Herting, M. M., Goddings, A. L., Meuwese, R., Blakemore, S. J., Dahl, R. E., Güroğlu, B., Raznahan, A., Sowell, E. R., Crone, E. A., & Mills, K. L., 2017. Development of the Cerebral Cortex across Adolescence: A Multisample Study of Inter-Related Longitudinal Changes in Cortical Volume, Surface Area, and Thickness. *J. Neurosci.* 37(12), 3402–3412. <https://doi.org/10.1523/JNEUROSCI.3302-16.2017>

Tustison, N., Holbrook, N., Roberts, J., Avants, B., Cook, P., Stone, J., Gillen, D., Yassa, M., 2017. The ANTs Longitudinal Cortical Thickness Pipeline. The proceedings of The 13th International Conference on Alzheimer's & Parkinson's Diseases, Vienna, Austria. <https://doi.org/10.1101/170209>

Tustison, N. J., A. J. Holbrook, B. B. Avants, J. M. Roberts, P. A. Cook, Z. M. Reagh, J. T. Duda, J. R. Stone, D. L. Gillen, M. A. Yassa and I. Alzheimer's Disease Neuroimaging (2019). Longitudinal Mapping of Cortical Thickness Measurements: An Alzheimer's Disease Neuroimaging Initiative-Based Evaluation Study. *J Alzheimers Dis* 71(1): 165–183.

van Dyck, L. I., & Morrow, E. M., 2017. Genetic control of postnatal human brain growth. *Curr. Opin. Neurol.* 30 (1), 114–124. <https://doi.org/10.1097/WCO.0000000000000405>

Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E., Ugurbil, K., & WU-Minn HCP Consortium, 2013. The WU-Minn Human Connectome Project: an overview. *NeuroImage* 80, 62–79. <https://doi.org/10.1016/j.neuroimage.2013.05.041>

Volkow, N. D., G. F. Koob, R. T. Croyle, D. W. Bianchi, J. A. Gordon, W. J. Koroshetz, E. J. Perez-Stable, W. T. Riley, M. H. Bloch, K. Conway, B. G. Deeds, G. J. Dowling, S. Grant, K. D. Howlett, J. A. Matochik, G. D. Morgan, M. M. Murray, A. Noronha, C. Y. Spong, E. M. Wargo, K. R. Warren and S. R. B. Weiss (2018). "The conception of the ABCD study: From substance use to a broad NIH collaboration." *Dev Cogn Neurosci* 32: 4-7. <http://dx.doi.org/10.1016/j.dcn.2017.10.002>

Walhovd, K. B., Westlye, L. T., Amlien, I., Espeseth, T., Reinvang, I., Raz, N., Agartz, I., Salat, D. H., Greve, D. N., Fischl, B., Dale, A. M., & Fjell, A. M., 2011. Consistent neuroanatomical age-

related volume differences across multiple samples. *Neurobiol. Aging* 32 (5), 916–932.

<https://doi.org/10.1016/j.neurobiolaging.2009.05.013>

Wang, Y., Necus, J., Kaiser, M., & Mota, B., 2016. Universality in human cortical folding in health and disease. *Proc. Natl. Acad. of Sci. U.S.A.* 113 (45), 12820–12825.

<https://doi.org/10.1073/pnas.1610175113>

Winkler, A. M., Greve, D. N., Bjuland, K. J., Nichols, T. E., Sabuncu, M. R., Håberg, A. K., Skranes, J., & Rimol, L. M., 2018. Joint Analysis of Cortical Area and Thickness as a Replacement for the Analysis of the Volume of the Cerebral Cortex. *Cereb. Cortex* 28 (2), 738–749.

<https://doi.org/10.1093/cercor/bhx308>

Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M., & Nichols, T. E., 2014. Permutation inference for the general linear model. *Neuroimage* 92 (100), 381–397.

<https://doi.org/10.1016/j.neuroimage.2014.01.060>

Winkler, A.M., Webster, M.A., Brooks, J.C., Tracey, I., Smith, S.M., Nichols, T.E., 2016. Non-parametric combination and related permutation tests for neuroimaging. *Hum. Brain Mapp.* 37, 1486-1511.

<https://doi.org/10.1002/hbm.23115>

Wu, K., Taki, Y., Sato, K., Qi, H., Kawashima, R., & Fukuda, H., 2013. A longitudinal study of structural brain network changes with normal aging. *Front. Hum. Neurosci.* 7, 113.

<https://doi.org/10.3389/fnhum.2013.00113>

Xia, J., Wang, F., Wu, Z., Wang, L., Zhang, C., Shen, D., & Li, G., 2020. Mapping hemispheric asymmetries of the macaque cerebral cortex during early brain development. *Hum. Brain Mapp.* 41 (1), 95–106.

<https://doi.org/10.1002/hbm.24789>

Yoshimatsu, S., Okahara, J., Sone, T., Takeda, Y., Nakamura, M., Sasaki, E., Kishi, N., Shiozawa, S., & Okano, H., 2019. Robust and efficient knock-in in embryonic stem cells and early-stage embryos of the common marmoset using the CRISPR-Cas9 system. *Sci. Rep.* 9 (1), 1528.

<https://doi.org/10.1038/s41598-018-37990-w>

Young, J. T., Shi, Y., Niethammer, M., Grauer, M., Coe, C. L., Lubach, G. R., Davis, B., Budin, F., Knickmeyer, R. C., Alexander, A. L., & Styner, M. A., 2017. The UNC-Wisconsin Rhesus Macaque Neurodevelopment Database: A Structural MRI and DTI Database of Early Postnatal Development. *Front. Neurosci.* 11, 29. <https://doi.org/10.3389/fnins.2017.00029>

Journal Pre-proof

Figure legend

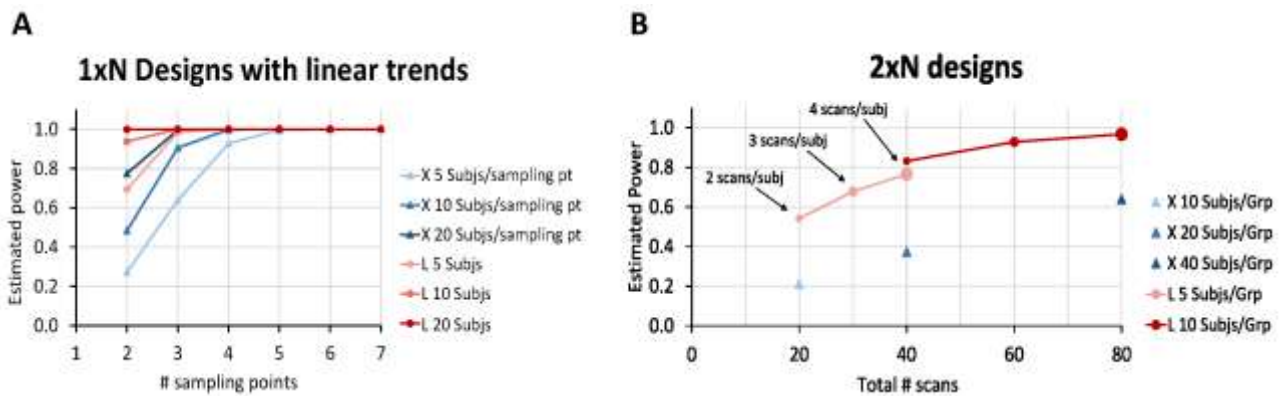


Figure 1. Power simulation of longitudinal (L) and cross-sectional (X) designs in two different scenarios: (A) a developmental/ageing study, where age varies within or between subjects; and (B) an interventional study with one control group and one treatment group (with equal number of subjects in each group). In the interventional scenario, the cross-sectional design with two groups is compared to a longitudinal design with two groups scanned once before treatment and once or several times after treatment (to study the sustainability of the treatment effect). The increased power for longitudinal experiments is illustrated using asymptotic statistical methods for simulated data with a fixed effect size and a fixed detection threshold, using the *longpower* package (<https://cran.r-project.org/web/packages/longpower/index.html>). It should be noted that practically, signal variability is not homogenous across brain regions and that different power value estimates would likely be obtained for different brain regions (Suckling et al 2014).

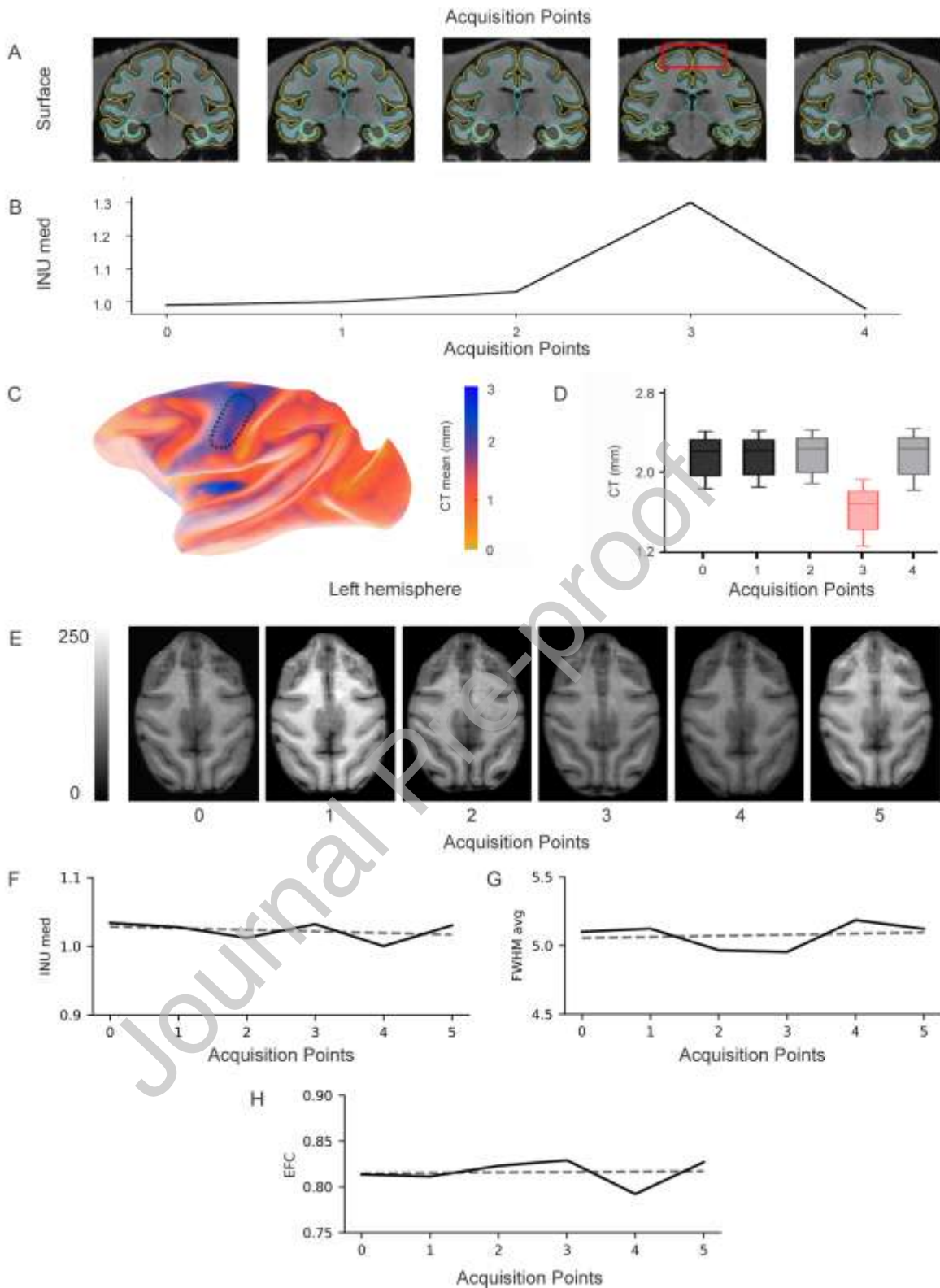


Figure 2. Quality control (QC) of T1w longitudinal images. A-D. Outlier identification. A. Brain surface estimation of an anesthetised macaque during five time points using PREEMACS Long. B. Estimation of intensity non-uniformity (INU) across acquisition points using MRIqc (Esteban et al., 2017) customized for the macaque by PREEMACS. C. Mean of the Brain Surface for the five acquisition points. The ROI defined by

the dotted line corresponds to the precentral gyrus (pcg). Colour bar of cortical thickness (CT) in mm. D. CT estimation of pcg for each acquisition point. Point 3 is considered an outlier with a pcg CT and an INU that were statistically different from the other time points. E-H. Lack of longitudinal trends on QC metrics associated with head motion during awake scanning. E. Longitudinal scans of an awake macaque (one scan every 6 months) F. Median of INU field (INU med) as extracted by the N4ITK algorithm (values closer to 1.0 are better) across six acquisition points. G. Full-Width Half Maximum Smoothness (FWHM) of the spatial distribution of the image intensity values in units of voxels (lower values are better) for each time point. H. Entropy Focus Criterion (EFC) as a function of time points. EFC uses the Shannon entropy of voxel intensities as an indication of ghosting and blurring induced by head motion (lower values are better). The dotted line (F-H) corresponds to the best linear model between the QC metrics and the six acquisition points, with a slope close to zero for the three QC metric, indicating no temporal trends in head motion.

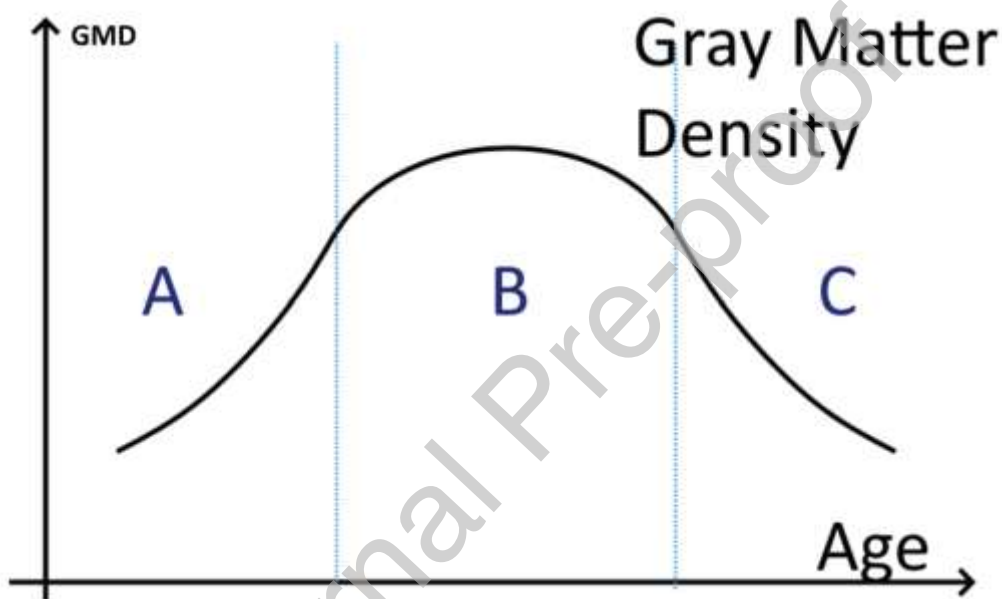


Figure 3. Anatomical measures such as Gray Matter Density (GMD), can change drastically across the lifespan, especially during early development (A) and late adulthood (C).

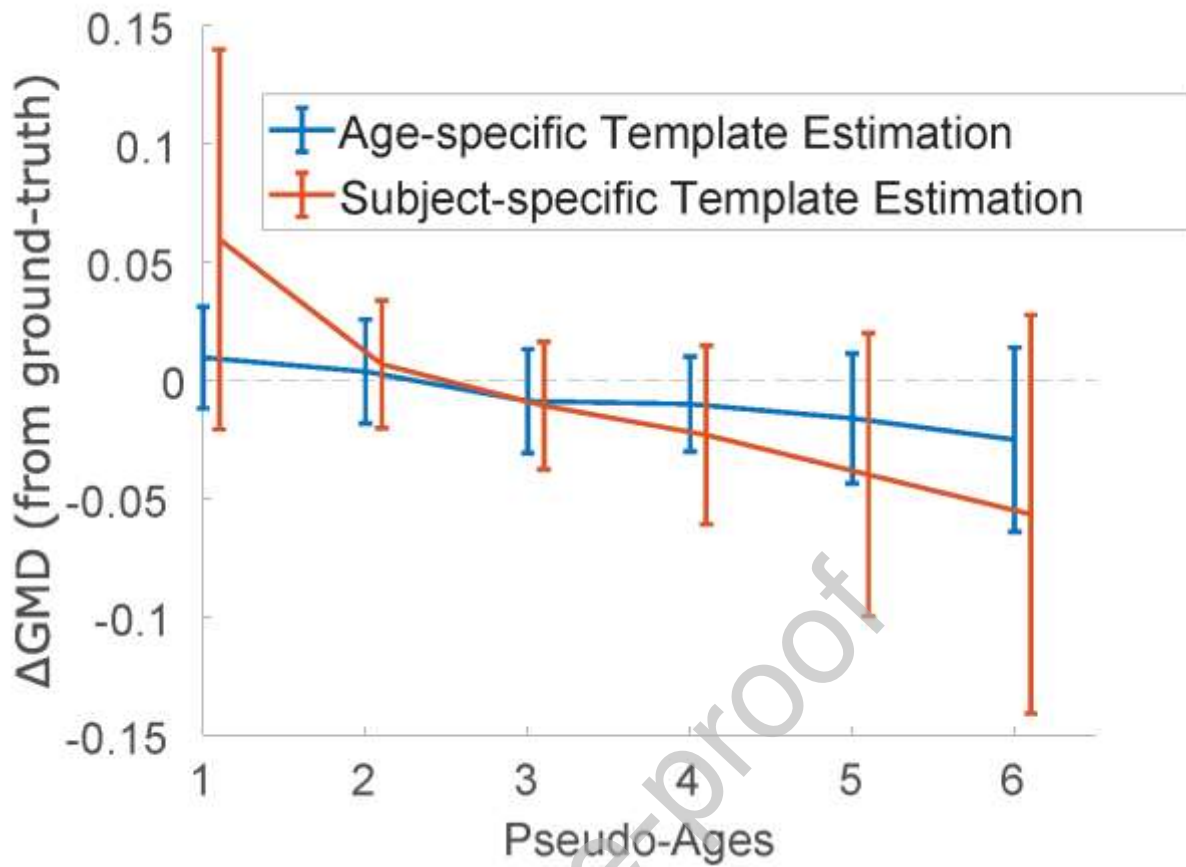


Figure 4. Simulation results showing the difference between grey matter density estimates and the ground truth following registration to an age-specific template (blue) and a subject-specific template (orange) in a pipeline combining linear and non-linear registration steps.