

Imaging the rapidly developing brain: Current challenges for MRI studies in the first five years of life

Ted K. Turesky^{1,2}, Jolijn Vanderauwera^{3,4}, and Nadine Gaab^{1,2}

¹Laboratories of Cognitive Neuroscience, Division of Developmental Medicine, Department of Medicine, Boston Children’s Hospital, Boston, Massachusetts, U.S.

²Harvard Medical School, Boston, Massachusetts, U.S.

³Psychological Sciences Research Institute, Université Catholique de Louvain, Louvain-la-Neuve, Belgium

⁴Institute of Neuroscience, Université Catholique de Louvain, Louvain-la-Neuve, Belgium

October 22, 2020

Abstract

Rapid and widespread changes in brain anatomy and physiology in the first five years of life present substantial challenges for developmental MRI studies. One persistent challenge is that methods best suited to earlier developmental stages are suboptimal for later stages, which engenders a trade-off between using different, but age-appropriate, methods for different developmental stages or identical methods across stages. Both options have potential benefits, but also biases, as pipelines for each developmental stage can be matched on either whether methods are age-appropriate or the employed methods themselves, but not both. This review describes the data acquisition, processing, and analysis challenges that introduce these potential biases and attempts to elucidate decisions and recommendations that would optimize developmental comparisons.

1 Introduction

Early childhood is characterized by rapid brain development (Gilmore et al. 2018) and represents a critical time for establishing developmental trajectories of neural anatomy and physiology. Consequently, experiences are most influential during these early years, with the potential to shape and alter typical developmental trajectories (Nelson and Gabard-Durnam 2020). Magnetic resonance imaging (MRI) technology, which began as a tool to study disorders and aberrations in adult clinical populations, was brought to bear on typical infant and early childhood brain structure (Pfefferbaum et al. 1994; Raininko et al. 1994) and function (Born et al. 1998, 2000; Martin et al. 1999) over two decades ago. With the ability to study the brain at early stages of life came the opportunity to understand developmental changes and trajectories using a longitudinal time frame. However, substantial and rapid changes in brain anatomy and physiology present significant and unique challenges for data acquisition, processing, analysis, and interpretation as part of a developmental neuroimaging pipeline.

Understanding neurodevelopmental changes across early childhood requires accurate and precise measurements of brain structure and function at all developmental stages (or age groups), and in a manner that does not favor a particular developmental stage when they are compared. On the one hand, past critiques have highlighted general methodological concerns for developmental studies (Madhyastha et al. 2018; Vijayakumar et al. 2018; or

the "Methodological Challenges in Developmental Neuroimaging: Contemporary Approaches and Solutions" Issue (Vol. 33, 2018) of *Developmental Cognitive Neuroscience* journal). However, these do not address the neurodevelopmental processes unique to children under five years of age, which are qualitatively and quantitatively different and often considerably faster than those transpiring after year five. On the other hand, critiques specifically for infant studies do not address developmental comparisons (Mongerson et al. 2017; Makropoulos et al. 2018; Zhang et al. 2019; Dubois et al. 2020). Indeed, there are methodological challenges that are unique to studies on early development that warrant special consideration (Vasung et al. 2019). In the present article, we first outline macroscopic anatomical and physiological changes that can impact developmental neuroimaging studies of early childhood. Grounded in these developmental changes, we then review the unique challenges associated with conducting developmental MRI studies in infants (i.e., children ≤ 1 year) and young children (1-5 years) and highlight various approaches that have been developed to overcome these challenges. We conclude by outlining where the field continues to struggle and what future methodological developments are needed. Overall, the aim of this review is to provide a resource for trainees and junior investigators, as well as for established investigators who are starting to conduct neuroimaging in this age range.

2 Developmental Changes in Brain Anatomy and Physiology

Proportionally, the largest neuroanatomical growth transpires in the first year of life and tapers thereafter (Matsuzawa et al. 2001; Knickmeyer et al. 2008; Groeschel et al. 2010; Sanchez et al. 2011; Deoni et al. 2012; Geng et al. 2012; Gilmore et al. 2012; Li et al. 2013; Lyall et al. 2015). Namely, total intracranial volume, cortical gray matter volume, subcortical volume, and cerebellar volume more than double in the first year (Knickmeyer et al. 2008; Gilmore et al. 2012). A smaller, yet significant, increase is observed in cortical thickness (Lyall et al. 2015), gray matter surface area (Li et al. 2013; Lyall et al. 2015), white matter volume (Knickmeyer et al. 2008), and white matter fractional anisotropy (FA; Geng et al. 2012), a measure of the structural organization of white matter pathways. The relative growth for the majority of these brain

measures diminishes substantially in year two (Knickmeyer et al. 2008; Geng et al. 2012; Gilmore et al. 2012; Li et al. 2013; Lyall et al. 2015), and afterward for gray and white matter (Courchesne et al. 2000; Matsuzawa et al. 2001; Groeschel et al. 2010) and FA (Reynolds et al. 2019). In contrast, relative growth in cortical white matter increases (Knickmeyer et al. 2008). Importantly, these growth rates also differ by brain regions (Gilmore et al. 2012) and white matter tracts (Deoni et al. 2012; Geng et al. 2012). Please see Figure 1 for sample MRI images depicting relatively rapid brain growth in early compared with later development. Nevertheless, comprehensive examinations of brain anatomy are currently limited to the first two years following birth (Table 1). Further work is needed to extend developmental trajectories to five years.

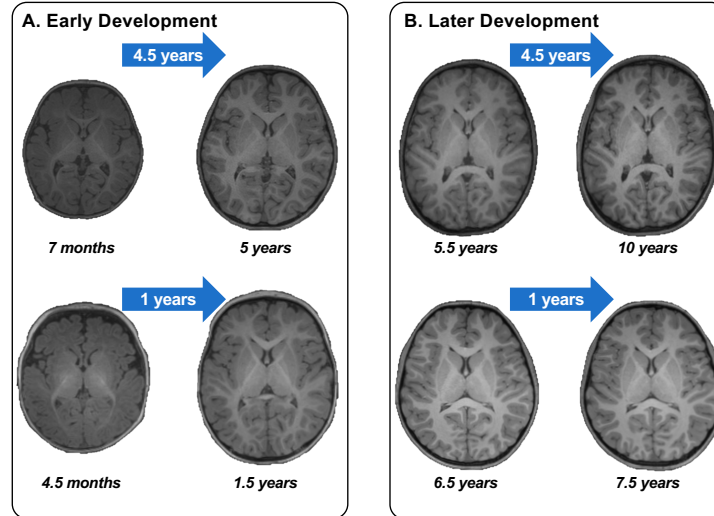


Figure 1. Sample structural MRI images depicting brain growth during (A) early and (B) later development. Each quadrant shows brain images from a child at two developmental stages as collected longitudinally. Please note the substantial anatomical changes during early development, especially in the first year of life, compared to later development.

In addition to developmental changes in brain tissue, anatomical structures that support the brain also change. For instance, head circumference increases most in the first year of life, while this increase is significant but smaller in the second and third years of life (Centers for Disease Control and Prevention 2001). Average skull thickness also increases most in year one, although it still increases substantially in years two and three (Table 1; Li et al. 2015a). Meanwhile, angiogenesis (i.e., the growth and branching of blood vessels) vastly expands microvasculature established prenatally; e.g., the dense layer of veins in vascular layer 3 is pruned and supplanted by a plexus of capillaries (Norman and O’Kusky 1986), which scales with local metabolic activity and neural and glial (e.g., astrocyte) demands for oxygen. Considerable microvascular remodeling also occurs during this time, involving vessel sprouting, pruning, and lengthening (Harb et al. 2013; for a review of postnatal microvascular development of the brain, please see Coelho-Santos and Shih 2019). Downstream, developmental changes in vasculature can affect functional MRI and near infrared spectroscopy experimentation (for a review, please see Vasung et al. (2019)), as the measures collected for these studies derive from neurovascular and gliovascular interactions; i.e., neurons and glial cells consume and deplete oxygen to function and this oxygen is replenished by oxygenated hemoglobin via blood ves-

sels (Yu et al. 2010). Finally, heart rate, respiration, and water content all decrease between infancy and adulthood (Mongerson et al. 2017; Bastiani et al. 2019).

These anatomical and physiological changes have profound implications for conducting structural and functional developmental MRI studies for two main reasons. First, methods that were developed for adults and older children need to be adapted to the developmental stage of the infant or young child to optimize the accuracy and precision of brain measures (e.g., tissue volume as a measure of brain structure, or functional activation or functional connectivity as measures of brain function). Second, practices that are optimal for a child of one age may be suboptimal for a child of a different age and induce spurious, age-related differences that are driven by methodological inconsistencies, rather than true brain development. Ultimately, this leads to a trade-off and challenging decision for investigators of whether to optimize methods for the neuroanatomy and neurophysiology unique to the specific developmental stage, or preserve consistent methods across stages. In the case of the former, developmental stages are matched on whether methods are optimal for children of that age; in the case of the latter, comparisons of developmental stages are matched on the methods themselves. Both options have potential benefits, but also introduce biases.

In the following sections, we highlight and describe the acquisition, processing, and analysis steps involved in developmental MRI datasets that may introduce these biases (please see Figure 2 for a summary). To ease imple-

mentation for any trainees and investigators entering the field, we have structured these sections largely according to an experiment pipeline typical of an early developmental MRI study.

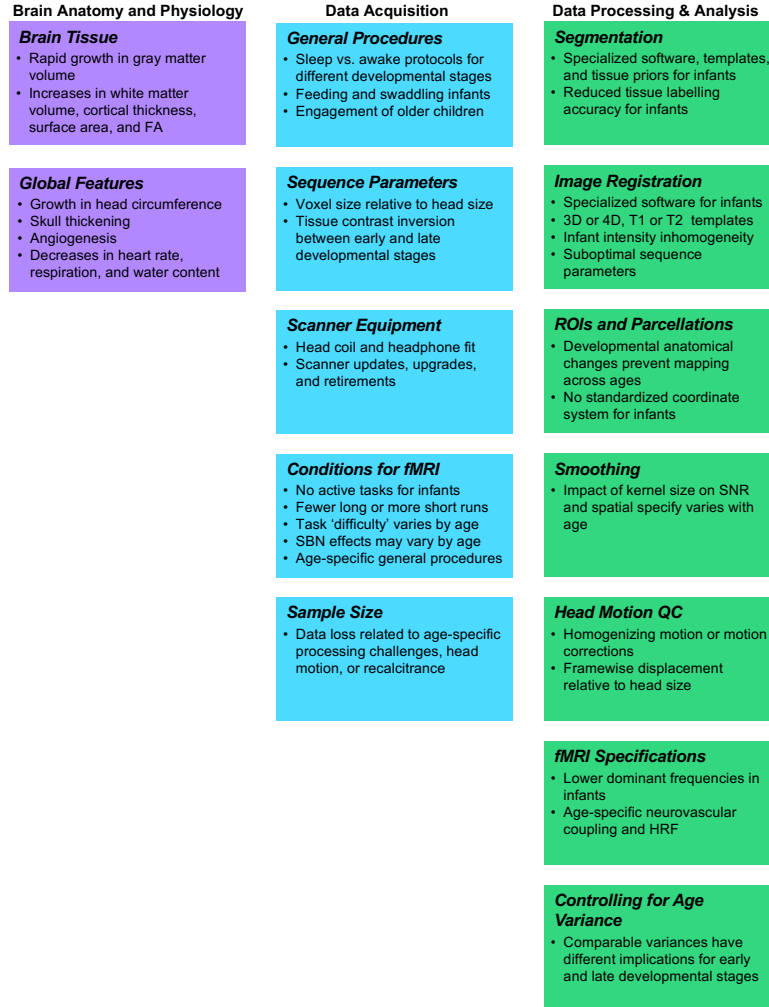


Figure 2. Challenges and options that may introduce bias when scanning children at different developmental stages between birth and age five.

3 Data Acquisition

Acquiring data for developmental MRI studies in children ages 0-5 years presents numerous challenges. The first set of challenges relates to the procedures for scanning young children. Infants and toddlers require specialized procedures and the absence of these can stymie quality data acquisition (Raschle et al. 2012a). Second, sequence parameters and scanner equipment that are optimal for infants may differ from those that are optimal for children between one and five years, given the current understanding of widespread and rapid developmental changes in brain and head anatomy and physiology. Third, angiogenesis and other developmental changes that affect neurovasculature have profound implications for acquiring functional MRI (fMRI) data, the signal for which depends on neurovascular coupling (Vasung et al. 2019). Each step in the data acquisition pipeline carries with it a trade-off between implementing age-appropriate data acquisition practices or using consistent methods across developmental stages. Lastly, as data processing is particularly challenging for children 0-5 years, data loss is likely to be high. This should be considered when estimating sample sizes for developmental stages in this age range.

3.1 General Procedures

Procedures for pediatric imaging require careful consideration to maximize compliance and data quality. One reason that procedures for MRI data acquisition in developmental neuroimaging for ages 0-5 are challenging

is because scanning infants and toddlers requires adapted data acquisition procedures (Raschle et al. 2012a). One example of an age-specific procedure relates to whether the child is asleep or awake during the scan. For the youngest children (i.e., birth to three years old), a natural sleep protocol is often used because children of this age can sleep in the scanner (Liu et al. 2008; Fransson et al. 2009; Gao et al. 2009; Doria et al. 2010; Graham et al. 2016; Howell et al. 2019; Turesky et al. 2019, 2020), which reduces motion and increases the likelihood of collecting usable data. Common techniques include a form of “feed and wrap,” in which infants are fed prior to and swaddled during the scan (Almli et al. 2007; Ibrahim et al. 2015), while more involved techniques can rely on vacuum immobilization (Dean et al. 2014; Hughes et al. 2017; for techniques used to induce natural sleep in infants during MRI, please see Raschle et al. 2012a; Antonov et al. 2017). Notably, this natural sleep procedure is in general most successful in the youngest children, although this is subject to large inter-individual variability. Infants have also been scanned while awake (Deen et al. 2017), more recently during cognitive tasks (Ellis et al. 2020), though scanning awake infants is rarer due to higher likelihood of motion. For typical developing children closer to five years of age, the natural sleep protocol is used less often and supplanted by detailed protocols that are specifically designed to engage the child, such as watching movies during structural sequences or rebranding functional tasks as games (Raschle

et al. 2011; Raschle et al. 2012b; Vanderauwera et al. 2017; Yu et al. 2018). Some developmental studies from birth to age five have been performed with all children asleep in the scanner (Gabard-Durnam et al. 2018), while others allow older children (≥ 3 years) to remain awake (depending on caregiver preference) and younger children (< 3 years) to sleep (Howell et al. 2019), which can introduce bias.

Moreover, this age and sleep/awake-related bias in head motion may be compounded when examining certain clinical populations, who tend to exhibit greater motion, such as children with attention-deficit/hyperactivity disorder or children with autism spectrum disorder. Taken together, mitigating head motion across early development may require different procedures prior to and during data acquisition. However, as discussed in greater detail below, this practice also introduces confounds for functional neuroimaging as sleep state and sleep stage can affect these measures (Horovitz et al. 2009; Picchioni et al. 2013).

3.2 Sequence Parameters

Sequence parameters are the set of radiofrequency pulse and gradient properties that the MRI scanner uses to acquire images, which vary considerably depending on whether one wishes to acquire structural, functional, or diffusion images and also on the population examined (Saunders et al. 2007). MRI scanners acquire images at the level of the voxel, a three-dimensional unit in the shape of a cube or rectangular prism. Analogous to pixels on a computer screen, each voxel occupies a small fraction of a

larger three-dimensional grid or volume and contains a corresponding intensity. The voxel size, which dictates the resolution of an image, can be set through the following sequence parameters: field-of-view, matrix size (i.e., number of samples in each direction), and slice thickness. Meanwhile, signal-to-noise ratio (SNR) in the image, which reflects desired signal intensity compared with level of background noise (Edelstein et al. 1986), can also be affected by these sequence parameters (Redpath 1998), as voxel size and SNR are inversely proportional (Runge et al. 2013). Resolution and SNR are critical image features because—along with tissue contrast, scan time, and presence of artifacts—they directly contribute to overall image quality. Accordingly, sequence parameters for developmental MRI studies need to be set such that voxel sizes and intensities are acquired without bias toward one or more developmental stage. This proves difficult when considering that the optimal set of sequence parameters (e.g., to maximize resolution and SNR) for one developmental stage may be suboptimal for another. Therefore, broadly, a second set of challenges for early developmental neuroimaging relates to whether to employ age-appropriate sequence parameters that optimize image quality for each developmental stage or identical sequence parameters for all stages.

In the most salient case, brains of infants are considerably smaller than those of older children. This has important implications when considering which voxel sizes to employ and whether to optimize for each developmental stage or standardize across stages. First, the same voxel size used for younger and older children

will represent a much larger proportion of brain tissue at the earlier developmental stages, which biases the resolution in favor of older children and cannot be overcome through normalization or smoothing (please see Data Processing and Analysis section below). Furthermore, the substantial growth of individual brain structures across early development (Gilmore et al. 2012) means that fewer voxels will capture each brain structure in the earlier compared with later developmental stages (Keil et al. 2011). Second, due to the smaller brain structures in younger compared with older children (Gilmore et al. 2012), single voxels are more likely to include multiple brain structures and tissue types, reducing the accuracy of segmentation of tissue types (discussed in greater detail below) and brain structures. And third, the smaller head size in children can cause a substantial drop in contrast between tissues compared with adults (Prastawa et al. 2005).

Rapid developmental changes in brain anatomy and physiology also cause brain tissue imaged with the same MRI sequence to look vastly different for children of different developmental stages. For instance, T1-weighted-sequences, which are commonly employed in MRI studies for their superior tissue contrast and clear visualization of brain structures, show white matter as brighter (i.e., higher intensity) than gray matter in adults and children over age one (Gilmore et al. 2018). By comparison, prior to six months, with increasing myelination (Deoni et al. 2012) and decreasing water content (Counsell and Rutherford 2002), gray matter exhibits higher intensity than white matter on T1-weighted sequences; gray and white

matter exhibit more comparable intensity between ages 8-12 months (Paus et al. 2001). While these differences may meaningfully characterize developmental changes in brain tissue, they also cause challenges for adequately measuring this brain tissue (e.g., in terms of volume or thickness). For instance, contrast between gray and white matter increases across developmental stages and can lead to better tissue labelling estimation in older compared to younger children (Fig. 1; also, please see Segmentation subsection below).

Thus, if the aim were to produce images with corresponding relative tissue intensities, different developmental stages would require different scan sequences. For structural scans, children over one year would undergo T1-weighted sequences and infants under one year would be scanned with T2-weighted sequences (Paus et al. 2001; Gilmore et al. 2018), which rely on different types of proton relaxations (after protons are perturbed by radiofrequency pulses). In addition, intensity inhomogeneity, in general, caused by heterogeneous static and radiofrequency fields, becomes more problematic in infants when considering ongoing myelination, as well as higher variability in tissue intensity and low-intensity contrast (Prastawa et al. 2005).

Overall, when deciding on sequence parameters, developmental neuroimaging studies encounter a central trade-off between optimizing these parameters for the neuroanatomy and neurophysiology unique to the specific developmental stages, or using the same parameters across stages. Both approaches introduce benefits and biases. The first option, optimizing sequence

parameters to a specific developmental stage, can increase the quality of images acquired for children of that age and improve the accuracy with which brain measures are estimated. By comparison, using suboptimal parameters could lead to reduced image quality, which in turn could lead to overestimation of volumetric, surface, and thickness estimates of one tissue type and underestimation of the other (relative to manual estimations). In this sense, tailoring sequences for each developmental stage, to more accurately estimate brain measures, can reduce bias in developmental comparisons. The second option, preserving scanning parameters across stages, is more common in developmental research. This approach may increase comparability across age ranges because no differences can be attributed to differences in sequence parameters, but it may also introduce bias from reduced image quality in specific developmental stages that are scanned using parameters that are not optimized for their unique biology. It is possible, though, that the bias introduced by suboptimal parameters is more severe for developmental studies of early childhood and has been hitherto underestimated when considering the early childhood population; further research will be needed to determine whether this is the case.

3.3 Scanner Equipment

In addition to scanning procedures and acquisition parameters, investigators face challenges in choosing MRI scanner equipment for studying early development. For instance, head coils designed to optimize infant imaging are too small for older children. Al-

though head coils designed for older children can accommodate a smaller infant head, the SNR, encoding performances, and consequent image quality will all be lower compared to using an age-appropriate head coil (Keil et al. 2011) because the gap between the head and head coil is wider. In the last decade, several laboratories and vendors have developed equipment that optimizes pediatric MR imaging for specific ages; however, the use of age-specific head coils introduces other challenges. First, they can be difficult to obtain, as these are often custom-built and costly (Keil et al. 2011). Second, if the head coil used for one developmental stage yields higher SNR compared with the head coil for another stage, then confounds are introduced. And third, even if SNR is comparable, the equipment used for each group is not consistent, which forces a trade-off between standardizing the head coil or the fit of the head coil across stages. Importantly, if investigators attempt to employ sequences with smaller voxel sizes to capture the smaller structures in infants (as described in the previous section), then SNR and consequent image quality would inherently decrease. The improvements in SNR from an age-appropriate head coil could offset the loss in SNR from the reduced voxel size. Additionally, age-appropriate head coils are also likely to improve comfort for children (Keil et al. 2011), which can increase time tolerated inside the scanner while also reducing head motion. As for head coils, investigators may need multiple head-phones for functional MRI studies to accommodate differing head sizes.

If conducting longitudinal studies, then scanner updates, upgrades, and retirements also pose challenges for

conducting multiple longitudinal measures of one individual. For example, infants may undergo MRI with one scanner; but by the time they are five years old, that scanner may have gone through considerable upgrades. Or perhaps the institution conducting the study replaces the scanner used for the infants with a newer scanner with superior image quality. As the field moves toward using larger datasets, one scanner at one site may be insufficient to acquire all of the scans planned for a developmental study. The use of different scanners – for example, at different scanning sites – can alleviate this challenge, but doing this might also bias comparisons between developmental stages if each site scans only children of particular ages. Future work needs to assess whether current guidelines for conducting multi-site studies (e.g., Glover et al. 2012) and practices for integrating datasets acquired from multiple sites or scanners (i.e., harmonization; Karayumak et al. 2019; Pomponio et al. 2020) can be applied to early developmental studies without introducing additional confounds.

3.4 Conditions for fMRI

Functional MRI (fMRI) captures neural responses by acquiring multiple volumes while participants perform a task or rest (often awake and mind-wandering, or asleep). Experimental conditions need to be designed to enable participant compliance that is consistent across developmental stages. Meeting this goal is especially difficult when examining developmental changes between birth and age five, foremost because infants can perform only passive tasks such

as processing auditory or visual stimuli (e.g., faces, bodies, and scenes) while awake (Deen et al. 2017) or, while asleep, viewing flickering lights through closed eyelids (Born et al. 1998) or listening to sounds (Anderson et al. 2001). Other studies acquired fMRI data while children were asleep without imposing a task (i.e., a resting-state fMRI paradigm; Fransson et al. 2007, 2011; Doria et al. 2010; Gao et al. 2015a; Graham et al. 2015a; Pruett et al. 2015; Graham et al. 2016; Mitra et al. 2017; Rudolph et al. 2018; Turesky et al. 2019). In theory, passive task or resting-state paradigms can also be used for older children, making them of interest for developmental studies. In practice, however, as children emerge from early infancy, they do not require as much sleep and have more difficulty remaining motionless in the scanner during passive tasks and resting-state, introducing greater motion into the older compared with earlier developmental stages. For task paradigms, fewer long functional runs could mitigate motion in infants and toddlers because these offer a consistent environment in which children can sleep without interruption. However, for older children, dividing a long functional run into multiple shorter functional runs reduces the amount of time the child needs to remain motionless and still facilitate collection of a sufficient number of data points; e.g., our group has probed phonological processing in children using separate experimental and control conditions (Yu et al. 2018). In contrast, resting-state paradigms, which are mainly used to examine functional connectivity, require at least four minutes of continuous data acquisition (after any data removal due to scrubbing; Power et al.

2012) for functional connectivity estimates to stabilize, according to studies in adolescents (Satterthwaite et al. 2013) and young adults (Van Dijk et al. 2010). If a comparable minimum duration is needed for infants (Mongerson et al. 2017), then separating this paradigm into multiple runs and concatenating volumes cannot be a solution. Lastly, even if children are given identical tasks across stages, task difficulty can vary inversely by age. Initially, the task may be too difficult, but then too easy by age five, with task difficulty introducing another source of bias.

Additionally, most fMRI experiments rely on the blood-oxygen-level-dependent (BOLD) signal to measure brain activation. The BOLD signal, which reflects the ratio of oxygenated to deoxygenated hemoglobin, is affected by cerebral blood flow, cerebral blood volume, and cerebral metabolic rate of oxygen (Ogawa and Lee 1990; Raichle 1998). Therefore, angiogenesis and microvascular refinement in early development could directly affect the BOLD signal (please see Developmental Changes of Brain Anatomy section and Coelho-Santos and Shih (2019) and Vasung et al. (2019) for reviews). The BOLD signal can also be affected by other, distal factors that may differ by developmental stage. For instance, as described above, infant scans are mainly conducted while infants are asleep, whereas older children are not generally instructed to sleep. If children of earlier developmental stages are asleep and later developmental stages are awake, then resulting developmental effects could instead reflect differences in consciousness, as these—with differential influences on cerebral metabolic pro-

cesses—affect the BOLD signal and subsequent functional measures differently (Horovitz et al. 2009; Picchioni et al. 2013). Evincing this, one report has shown that resting-state functional connectivity in sleeping infants is qualitatively more similar to resting-state functional connectivity in adults during sleep relative to adults during wakefulness (Mittra et al. 2017). In addition, identical patterns of resting-state functional connectivity do not underlie all sleep stages. This variability may be monitored using concurrent EEG or by tracking physiological measures such as heart rate or respiration. However, EEG acquisition in developmental populations has its own set of challenges, especially when performed concurrently with MRI. Further characterization of these unique challenges is, however, beyond the scope of the present review.

Another factor known to impact consciousness during acquisition pertains to scanner background noise (SBN), which can be extremely loud at over 100 decibels. SBN can interfere with sleep, which can negatively impact head motion, but it also requires special consideration in the context of fMRI as it has been shown to attenuate auditory cortex responses (Gaab et al. 2007a, 2007b, 2008) and these effects may vary by age. Ear protection has effectively attenuated SBN in younger children (Zahr and de Traversay 1995). However, keeping ear protection fastened can be more challenging in infants and toddlers compared with older children. Furthermore, it is important to consider that often only one set of headphones are available at an imaging center and their fit may differ depending on the age of the child. Certain experimental designs, such as sparse

sampling (whereby auditory stimuli are played during silent delays in MRI volume acquisitions; Gaab et al. 2003; Perrachione and Ghosh 2013), which reduces effects of SBN (Gaab et al. 2007b), can be alternatives. Therefore, investigators should consider whether these adjustments can be made to their data acquisition protocols in a manner that does not limit their ability to address their research questions. However, these sparse protocols increase overall scanning duration and contain long periods of silence, which can be suboptimal for certain age ranges as it can lead to increased head movement.

Lastly, systematic differences in internal and external conditions among developmental stages prior to and throughout data acquisition may bias fMRI data acquisition. For instance, infant feeding as part of the “feed and wrap” technique (Almli et al. 2007; Ibrahim et al. 2015) could affect glucose metabolism, which would affect the BOLD signal. Also, anxiety and discomfort evoked by scanner background noise (Gaab et al. 2007a, 2007b; Raschle et al. 2012a) may disproportionately affect infants’ and older children’s responses to stimuli. Differences in scan state (environment) too may bias fMRI data acquisition as functional brain networks constituting functionally coupled brain regions (Buckner et al. 2013) exhibit moderate variability depending upon whether data are acquired while the participant is performing a task or at rest (Cole et al. 2014; Gratton et al. 2018). For example, with infants swaddled (Almli et al. 2007) or vacuum immobilized (Dean et al. 2014; Hughes et al. 2017) to induce sleep, and older children physically unconstrained, somatosensory inputs would

differ by developmental stage. Overall, different experimental conditions for different developmental stages may introduce bias during fMRI data acquisition.

3.5 Sample Size

The challenges described above have implications for proposed sample sizes, attrition rate, and power calculations. First, infants may have difficulty sleeping (or remaining asleep) in the scanner, which could cause severe and repeated motion artifacts and require that they be removed from the scanner. Second, data processing challenges unique to the infant stage (e.g., segmentation with inverted and reduced gray/white matter contrast and strict head motion thresholds; please see below) can reduce the number of datasets of sufficient quality to be included in the developmental analyses. In longitudinal studies, these challenges can disqualify participants from follow-up scans at later developmental stages if laboratories do not choose to impute in their statistical analyses. And third, head motion in older children who now may be awake in the scanner (please see Data Acquisition section above) may reduce the number of usable datasets at later stages, particularly stymieing longitudinal analyses.

Concordantly, at roughly 70% (e.g., Almli et al.), success rates for developmental studies from birth to five years are fairly low, and these rates vary by developmental stage. For instance, success rates tend to follow a U-shape across development, decreasing from birth to toddlerhood and then increasing closer to age 5. Specifically, children scanned at 2-4 weeks, 1 year, and

2 years had success rates of 60%, 50%, and 50% respectively (Knickmeyer et al. 2008); and between 2 and 5 years, higher rates have been reported at 72% (Thieba et al. 2018). A study by our group reported on this trend using cross-sectional data, showing 63% success rates at 3-7 months, 59% at 8-15 months, and 97% at 4-6 years (Raschle et al. 2012). Success rates also vary depending upon the definition of success. For example, Thieba and colleagues (2018) reported a 72% success rate for acquiring one out of three scans, but a 48% success rate for acquiring all three scans (Thieba et al. 2018).

For longitudinal studies, investigators should also expect attrition and consequently success rates lower than those cited above. An attrition rate of approximately 20% has been reported, though this figure can vary widely depending upon age, socioeconomic status, and psychiatric factors (Stange et

al. 2018), as well as the duration of time between data acquisitions. To mitigate this, maintaining strong relationships with families of children ages 0-5 years is extremely important – though uniquely challenging as this period of their lives is often accompanied with considerable flux, especially for those who have only recently become parents for the first time (please see Stange et al. (2018) for longitudinal neuroimaging retention practices and Teague et al. (2018) for general longitudinal retention practices). Guidelines in the field that resolve or mitigate these challenges will help to accurately estimate proposed sample sizes and attrition. However, in the interim, it may behoove investigators to begin with a higher initial sample size for developmental studies from birth to age five relative to the initial sample size used in developmental studies in children over five years of age.

4 Data Processing and Analysis

To quantitatively examine developmental stages, MRI data need to undergo rigorous pre- and post-processing. Data processing and analysis can comprise many steps, but we will limit our discussion to steps that are most likely to introduce bias to developmental comparisons. These are segmentation, image registration, defining regions-of-interests and parcellations, smoothing, head motion quality control (QC), and several fMRI-specific processing considerations. As with data acquisition, each

processing step introduces a trade-off between implementing age-appropriate data acquisition practices, or using similar or identical methods across developmental stages.

4.1 Segmentation

Basic tissue segmentation is the process of identifying and separating gray matter, white matter, and cerebrospinal fluid voxels, which can be important for quantifying tissue volumes, limiting subsequent analyses to certain tissue types, and/or registering

images (Ashburner and Friston 2005). Accurate tissue segmentation in an MR image, including extracting the brain from the skull (i.e., brain extraction or skull stripping), is considerably more challenging for infants compared with older ages (Makropoulos et al. 2018), which is a substantial concern as segmentation is the gateway to many subsequent analyses of brain structure (Devi et al. 2015). Unmyelinated axons, ongoing myelination, low tissue contrasts (Prastawa et al. 2005), and thin cortex (Lyll et al. 2015) in infants relative to older children and adults have limited the accuracy with which canonical segmentation tools can label gray/white matter tissue in infants (Makropoulos et al. 2018).

These age-related, neurobiological characteristics have engendered two potential challenges for developmental MRI studies. First, segmenting brains of different ages requires different tools, techniques, and/or parameters. Several labs have developed segmentation tools specifically for infants, such as Infant Brain Extraction and Analysis Toolbox (iBEAT; Dai et al. 2013; version 2.0 has recently been released: <https://ibeat.wildapricot.org/>) and Infant FreeSurfer (Zöllei et al. 2020; for a full list of neonate segmentation tools as of 2018, please see Makropoulos et al. 2018). In addition, investigators can also label tissue manually, which is time consuming and labor intensive (Makropoulos et al. 2018), but can be an effective alternative to automated tools (de Macedo Rodrigues et al. 2015). With regard to techniques, segmentation algorithms employing age-appropriate tissue priors—typically gray matter, white matter, and cerebrospinal fluid probability

maps in which voxel intensities reflect the likelihood that that voxel corresponds to a particular tissue class (Ashburner and Friston 2005)—also reduce misclassification of tissue, compared with segmentations using adult tissue priors (Altaie et al. 2008). Tissue priors spanning multiple stages beginning at the perinatal period are now available as part of 4D templates (Oishi et al. 2018; please see below for details). Similarly, some investigators have opted for a 4D longitudinal approach (which, as the name suggests, can be applied in longitudinal studies), in which segmentations are performed on data from older children (for whom segmentations are generally more accurate due to improved tissue contrasts), and then used as tissue priors to inform segmentation in infants (Shi et al. 2011; Gilmore et al. 2012). Segmentation without tissue priors (Dhankhar et al. 2010; e.g., using Advanced Normalization Tools (ANTs; <https://github.com/ANTsX/ANTs>) Atropos with k-means) would theoretically circumvent this trade-off, but introduces other bias as the algorithms might then rely more heavily on tissue contrast, which is inherently poorer in younger children (Prastawa et al. 2005), especially between 8 and 12 months (Paus et al. 2001). For brain extraction protocols that rely on a template with skull (such as the ANTs `antsBrainExtraction` module), age-appropriate templates would also reflect developmental increases in head circumference and skull thickness (Centers for Disease Control and Prevention 2001; Li et al. 2015a). Importantly, several laboratories have started to focus on the development of tools, templates, and tissue priors for segmentation as well as image registra-

tion and parcellation steps as outlined below.

Second, segmentation quality is age-dependent even with specialized tools and techniques. This can lead to under- or overestimation of tissue volumes, surface areas, and cortical thickness in the younger age group and therefore erroneous, age-related differences when compared to older age groups. Given that some toolkits (e.g., Statistical Parametric Mapping (SPM; <https://www.fil.ion.ucl.ac.uk/spm/>) 12) perform tissue labeling in concert with image registration as part of the unified segmentation framework (Ashburner and Friston 2005), segmentation quality may also affect analyses of brain structure or function that require normalizing to standard space (please see below). As with data acquisition (see above), segmentation procedures for developmental studies come with a careful trade-off between adapting approaches to the age-appropriate range or maintaining approaches for measurements across stages.

4.2 Image Registration

Registration is the process of aligning images to one another, and different data processing steps may depend upon it for various reasons. For instance, to control for head motion throughout a functional run, functional volumes will be realigned (via x, y, and z translations and pitch, roll, yaw rotations) such that the brain images in every volume overlap. To control for one individual’s head movement between scans of different modalities (e.g., structural and functional volumes), the image alignment procedure is often referred to as coregistration. In the special case of nor-

malization, all individual brain images are aligned or spatially normalized to a standard brain image (i.e., a template), which is important for group analyses to ensure that the location of any given brain area in one participant spatially corresponds to the location of the same brain area in another participant. Normalization can be especially challenging for developmental studies for two main reasons. First, the process of registering very young brains to templates (both through linear or non-linear procedures) can require specialized tools. For instance, the first version of iBEAT software relied on Hierarchical Attribute Matching Mechanism for Elastic Registration (Shen and Davatzikos 2002) for registering individual and standard images (Dai et al. 2013). In our experience, popular MRI data processing software packages such as SPM or the voxel-based morphometry toolkit implemented through it have suboptimally normalized individual infant images. That these tools couple normalization with segmentation is consistent with the limited adaptability of unified segmentation algorithms (Beare et al. 2016). In contrast, the ANTs algorithms could register individual and template images (Turesky et al. 2019) and, as it was not designed specifically for the younger children, can also process data effectively for older children, precluding bias due to different tools (we have made our code from this study available on github: <https://github.com/TeddyTuresky>). Crucially, accurate normalization (and segmentation) also depends upon rigorous correction for intensity inhomogeneity (Mongerson et al. 2017; Makropoulos et al. 2018), which is substantial in infants (Prastawa et

al. 2005; please see Data Acquisition: Sequence Parameters subsection). N4ITK, an intensity inhomogeneity correction module implemented through ANTs (Tustison et al. 2010), is commonly used for infants (Wu and Avants 2012; Makropoulos et al. 2014; Turesky et al. 2019), but it is unclear whether optimal corrections for all developmental stages can be achieved by an identical set of parameters.

It is also unclear which templates are most appropriate for processing early developmental data. For infants, most groups use age-corresponding templates (Gao et al. 2015a,b; Turesky et al. 2019), and often with secondary transformations to or from adult templates to draw on stereotaxic standards (Graham et al. 2015a; Graham et al. 2016). On the other side of the early developmental age range, at five years, adult templates have been used (Raschle et al. 2012b; Raschle et al. 2014). However, brain structure from birth to age five changes so rapidly (Matsuzawa et al. 2001; Groeschel et al. 2010; Sanchez et al. 2011) that using an adult template or even a template corresponding to the mean or median of the developmental age range investigated can potentially favor registration accuracy for some developmental stages over others, as use of age-inappropriate templates can lead to misclassifications of tissue and brain regions (Wilke et al. 2003).

Use of age-appropriate templates (i.e., a 4D template with one target image per age group) can reduce this confound (Oishi et al. 2018), though there is still the concern that normalization in the youngest children is not as precise as in older children (Shi et al. 2011), potentially due to reduced im-

age contrast. Some have opted to use T1 structural images from one child with longitudinal data at early stages as a template (Gao et al. 2015a,b). Beyond this, a 4D template comprising both T1-weighted and T2-weighted images can reduce the bias introduced by image contrast as described earlier; however, this requires that laboratories collect both sequences (with sufficient image quality) at all stages along the developmental trajectory they wish to investigate. This can be costly and time-consuming. There are several 4D templates currently available (Oishi et al. 2018). However, none extend from birth to early childhood in a longitudinal sample, contain both T1 and T2 sequences, and have corresponding tissue priors. A close offering used by others (Graham et al. 2015a; Graham et al. 2016) was a template containing both T1- and T2-weighted images from infants ages 0 to 4.5 years (3D templates for 0-2, 2-5, 5-8, 8-11, 11-14, 14-17, 17-21, 21-27, 27-33, 33-44, 44-60 months; Almli et al. 2007; Fonov et al. 2009; <http://nist.mni.mcgill.ca/?p=1005>). However, this template was only partially longitudinal and did not have corresponding tissue priors, which are critical for accurate segmentation. More recently, Zhang and colleagues (2016) generated a longitudinal 4D template containing T1- and T2-weighted images and corresponding tissue priors. However, this template currently extends only to 12 months of age (Zhang et al. 2016). To our knowledge, no multi-year 4D templates for diffusion images have been released for the early developmental period, though a framework for constructing one has been proposed (Hart et al. 2010).

Suboptimal acquisition practices

(e.g., sequence parameters; please see Data Acquisition section above) for one developmental stage relative to another can also bias data processing of these images if resulting image quality (e.g., due to low contrast or motion artifacts) is lower for children of that stage. This is of particular concern when registering images. For instance, typically, preprocessing procedures for functional imaging involve (1) registering functional and structural volumes, (2) normalizing structural volumes to a standard template, and then (3) applying the transformation in step 2 to the functional image. However, poor T1 or T2 image quality could lead to misalignment of functional and structural volumes in native space and/or misalignment of structural images with standard templates. The resulting functional images might therefore exhibit more spatial variability in the developmental stage with poorer image quality, which can subsequently affect second-level, age-related comparisons. Conversely, some laboratories have opted to perform functional analyses in native space, which eliminates bias introduced by normalization (Ferradal et al. 2018), but also prevents brain regions of different participants from occupying the same stereotaxic space, encumbering inter-subject comparisons and interpretations.

A similar bias can emerge for functional and diffusion imaging modalities specifically. For instance, realignment of functional volumes (e.g., registering all volumes in a time series to the first or mean volume) to control for head movement during a functional run can be less accurate for one developmental stage if the constituent images are of lower quality. Less accurate registration can also be expected for normal-

ization if functional images are registered directly to a template (without first being registered with structural images for that participant, which are of higher contrast than functional images and therefore preferred for registration). In addition, a noisier signal can effect less precise estimation of task-evoked activation and functional connectivity. This process can be similarly disrupted in processing diffusion volumes, which may involve registration to subject-specific T1 or T2 images and to standard space templates for accurate fiber tract segmentation (Yeatman et al. 2012). Independent of misalignments and prior to tract segmentation, lower quality diffusion scans can also bias the quality of typical diffusion processing steps such as tensor fitting and fiber tractography, and optimal sequence parameters (e.g., b-values) may be different for infants and older cohorts (Bastiani et al. 2019).

4.3 Defining Regions-of-Interest and Parcellations

To examine effects in specific brain regions or voxels, an investigator might define a region-of-interest (ROI). Anatomical parcellations are useful as ROIs because parcellations reflect subdivisions of the entire brain according to structural or functional boundaries. However, the challenges described above for templates also propagate to corresponding anatomical parcellations (for the templates that have corresponding parcellations). Parcellations corresponding to templates for older children cannot accurately label the anatomically smaller, and possibly spatially shifted, cortical and subcortical structures in infants and toddlers.

Using age-specific parcellations can reduce these inaccuracies. Shi and colleagues (2011) have warped the AAL atlas from adults to infants (Shi et al. 2011), but this process still introduces bias as brain structures can shift in relative size and stereotaxic location to one another over the time course of brain development (Oishi et al. 2018), requiring visual inspection to confirm sufficient overlap between template and individual structures. It will be important for future anatomical parcellation methodologies to reflect the anatomy of the younger children, not simply warped from parcellation maps of older children or adults. Li and colleagues (2015) generated a 4D cortical surface atlas with volumes for chronological ages 1, 3, 6, 9, 12, 18, 24, 36, 48, 60, and 72 months using 339 longitudinally scanned MRIs from 50 healthy infants (Li et al. 2015b; numbers updated from original report). However, this template and corresponding parcellation map only include the cortical surface and do not contain subcortical brain areas nor corresponding tissue priors. For a list of available infant parcellations current to 2018, please see (Oishi et al. 2018). To the best of our knowledge, there are currently no longitudinal templates from birth to approximately five years of age (even with annual scanning intervals) with corresponding tissue priors and anatomical parcellations, which is a considerable challenge for developmental neuroimaging studies for this age range. Lastly, like segmentation, manual delineation of structures is a time-consuming and labor-intensive, but effective, alternative (Vanderauwera et al. 2018).

ROIs to examine specific brain regions or networks may be defined

in other ways, too. For instance, unlike anatomical parcellations, functional parcellations reflect brain regions subdivided according to their function (or functional connectivity) rather than structure (Arslan et al. 2018). Meanwhile, Brodmann areas represent brain regions according to cytoarchitecture. Longitudinal birth to two-year (Shi et al. 2018) and neonate (Scheinost et al. 2016) parcellations have been generated using resting-state, functional connectivity data; however, these parcellations did not include older children, limiting developmental comparisons with older children. Additionally, ROIs need not conform to a whole-brain subdivision, but can also be commonly constructed as representative spheres or cubes of voxels sampled from widespread parts of the brain. For instance, Smyser and colleagues (2016) warped infant resting-state data to an adult template and overlaid Power’s 264 ROIs, defined by homogeneity of BOLD signal fluctuations (Power et al. 2011). However, fifty ROIs did not overlap with infant gray matter (Smyser et al. 2016), demonstrating that functional network architecture (like anatomical architecture) can vary considerably in stereotaxic location by developmental stages and posing challenges for matching brain areas and networks from multiple stages. Overall, further research and tool development are needed to ensure unbiased parcellation or ROI selection across developmental stages.

Finally, coordinates are commonly used to characterize the location of brain effects. For children over five years of age and adults, individual brain images are often registered to templates in either Montreal Neurological Institute (MNI) and Ta-

lairach stereotaxic space, each of which come with standard coordinate systems. Any resulting brain effects may then be reported using MNI or Talairach coordinates; subsequent investigators can use these coordinates for ROI analyses after registering their own datasets to MNI or Talairach space. The lack of an accepted standardized coordinate system for infants is problematic (Akiyama et al. 2013), because if an investigator were to select ROIs for an infant study based on results of a previous infant study, then he or she would need to register the imaging data to the same template used by the previous study. Likewise, ROIs can be taken only from meta-analyses in older children and adults, but not infants, as quantitative meta-analytic techniques such as Activation Likelihood Estimation (Turkeltaub et al. 2002) are not possible without a standard coordinate system (or known conversions to such; Oishi et al. 2018).

4.4 Smoothing

Spatial smoothing, also considered blurring, is the process of averaging (weighted) intensities from nearby voxels. Smoothing increases SNR, but simultaneously decreases spatial specificity. The extent of the smoothing depends on the size of the smoothing kernel such that larger kernels increase SNR and decrease spatial specificity more than smaller kernels. As infant brains are considerably smaller than the brains of older children and adults (Groeschel et al. 2010), using kernels commonly used for older populations, such as 8 mm full width at half maximum (e.g., Yu et al. 2018), the default in SPM, could be problematic for localizing structures that are especially

small in infants, such as the amygdala (Graham et al. 2015b). A recent study in infants younger than 10 days observed that the optimal compromise between SNR and spatial specificity occurred when using a smoothing kernel size of 3 mm full width at half maximum (Baxter et al. 2019); many infant studies have used a smoothing kernel size of 4 mm (Ferradal et al. 2018; Howell et al. 2019), 5 mm (Gabard-Durnam et al. 2018), or 6 mm (Fransson et al. 2007; Pruett et al. 2015; Merhar et al. 2016; Smyser et al. 2016; Mitra et al. 2017; Turesky et al. 2019). For developmental studies, use of smaller kernels for earlier developmental stages could standardize spatial specificity relative to brain size across stages. For instance, Gao and colleagues (2015) used increasing kernel sizes for increasing ages and head sizes, ranging from 4 mm for 1-month-old to 8.2 mm for 12-month-old infants (Gao et al. 2015a). While this adaptive approach may increase SNR in later compared with earlier developmental stages, it can also reduce differences in relative spatial specificity across developmental stages (Graham et al. 2015b). Accordingly, an adaptive approach to spatial smoothing may represent the best compromise if one selects kernel sizes that are considered to optimally balance SNR and spatial specificity for each developmental stage.

4.5 Head Motion QC

Despite attempts during data acquisition to minimize participant motion, functional runs inevitably comprise head motion and the magnitude of this head motion can vary considerably, especially across developmen-

tal stages. Furthermore, for a given estimate of motion, relative to head size, this estimate would be greater for infants compared with later developmental stages. Head motion has been of major concern in developmental literature in children over five years of age, such that findings previously attributed to developmental changes (Fair et al. 2007, 2009; Dosenbach et al. 2010) were mitigated upon reanalysis with stricter motion criteria (Satterthwaite et al. 2012). The criteria for head motion have become extremely stringent in resting-state functional connectivity studies in infants, by comparison with early reports (Doria et al. 2010; Fransson et al. 2013). For instance, many studies now remove from functional time series data points in which the brain has moved from one volume to the next (i.e., framewise displacement) ≥ 0.5 mm (Gao et al. 2015a,b), ≥ 0.3 mm (Graham et al. 2015a; Graham et al. 2016; Rudolph et al. 2018), ≥ 0.25 mm (Smyser et al. 2016) or ≥ 0.2 mm (Pruett et al. 2015), sometimes along with additional data points that are temporally adjacent to this motion spike. A full run is often removed when the number of these artifactual volumes exceeds a particular threshold (e.g., 22.5%; Graham et al. 2015a) or does not provide sufficient remaining data (e.g., five minutes of continuous data; Smyser et al. 2016). Importantly, the thresholds used for examination of functional connectivity are generally considerably more strict than for task-evoked activation. Lastly, some studies performing group comparisons have additionally ensured that the numbers of artifactual volumes removed do not differ between groups; otherwise, differences due to head motion could be mistakenly at-

tributed to neuronal effects (Van Dijk et al. 2012). Although developmental neuroimaging practices in the 0-5 year age range might potentially benefit from this practice, further research is required to investigate the delicate trade-off between artifact removal and rejection of data that is extremely challenging to acquire in this population.

Additionally, in the calculation of framewise displacement (i.e., combined translation and rotation movement of the brain between contiguous volumes), an estimate of brain radius is often needed to convert rotational measures of degrees or radians to millimeters. Power and colleagues (2012) estimate 50 mm (Power et al. 2012); but this is presumably for the brain of an older child or adult, rather than an infant, which is considerably smaller (please see above and Table 1 for size comparisons). Apart from the question of consistent motion thresholds across ages, the field needs to establish whether an average head size estimate or different estimates for different ages should be used for the calculation of framewise displacement. Otherwise, framewise displacement could be assessed separately for translation and rotation, as is also fairly common in the literature (e.g., Yu et al. 2018), thereby obviating the need for conversion to millimeters using an estimate of head size. Though avoiding the concerns over head radius, this scenario would still be subject to the challenges described at the beginning of this subsection. Ultimately, to control for motion across the 0-5 year age range, the field of developmental cognitive neuroscience will need to determine whether head motion thresholds should be the same or differentially strict across developmental stages.

Head motion also hampers diffusion pipelines. Age-independent tools such as the FMRIB Software Library’s eddy tool and DTIPrep (Oguz et al. 2014) are often used in infants (Geng et al. 2012; Langer et al. 2017; Bastiani et al. 2019). Recently, a new motion correction tool called Spherical Harmonics and Radial Decomposition (SHARD; Christiaens et al. 2020; preprint) was introduced for multi-shell high angular resolution diffusion imaging (HARDI) data, which carries more fine-grained microstructural information compared with classical diffusion imaging. Using a data-driven representation of multi-shell HARDI data, SHARD avoids imposing a model with numerous assumptions, and germane to the present review, has demonstrated efficacy for the challenging motion correction in infant datasets. However, this tool can be applied only when HARDI data are available. Ultimately, to control for motion across the 0-5 year age range, the field of developmental cognitive neuroscience will need to determine whether head motion thresholds should be the same or differentially strict across developmental stages.

4.6 fMRI Specification

As described in the Data Acquisition section above (please see Experimental Conditions for fMRI subsection), fMRI captures brain responses over time, most frequently through the BOLD signal. The BOLD signal, like other types of signals, can be decomposed (via a Fourier transform) and characterized by the frequencies that constitute it, with some frequencies constituting more of the signal than others and the degree to which a particular

frequency constitutes a signal can be computed with power spectra analyses. Biases may also be introduced in the context of BOLD signal frequency. Power spectra analyses have revealed that the BOLD signal is dominated by different frequencies at different ages, with greatest representation for neonates at 0.0056 Hz compared with 0.028 Hz for one- and two-year-olds (Alcauter et al. 2015), likely with some variation depending on the number of points specified in the discrete Fourier transform. This has implications for resting-state, functional connectivity studies in infants, which often use band-pass filters with a lower threshold of roughly 0.008/9 Hz (Gao et al. 2015a; Graham et al. 2015a; Pruett et al. 2015; Rudolph et al. 2018; Turesky et al. 2019) to remove artifacts or non-neurophysiological signal thought to manifest at these lower frequencies in adults. As such, these studies effectively remove the dominant signal for the neonate datasets. In a rare deviation, Graham and colleagues (2016) used a low-pass filter ($< .1$ Hz), which presumably would have retained these low frequencies (Graham et al. 2016), but perhaps low frequency artifacts, too. Future work needs to determine whether low frequency artifacts are present in infants as they are in older children, and if they are, whether their removal can be performed without removing potentially useful information from the lower frequencies of the BOLD signal. Lastly, heart rate and respiration rate are higher at younger ages, which should be considered when deciding on low-pass filters for removal of non-neurophysiological noise functional connectivity studies (Mongerson et al. 2017).

Developmental changes in vasculature (please see above) may also introduce age-related differences in neurovascular coupling (Arichi et al. 2012; Hagmann et al. 2012; Cusack et al. 2015; Grayson and Fair 2017; Baxter et al. 2019), as has been investigated for children over five years of age and adults (Thomason et al. 2005). For instance, the hemodynamic response function (HRF), which characterizes changes in BOLD signal over time, was attenuated and exhibited a longer latency-to-peak in infants compared with adults for a passive somatosensory response (Arichi et al. 2012). Convolution of an age-appropriate HRF into the general linear model revealed activation in the somatosensory cortex that was not identified when convolving with the canonical adult HRF. Indeed, that HRF morphology and fit to stimulus-evoked responses affects the sensitivity of first-level modeling in infants (Baxter et al. 2019) and could generate false-positive effects in developmental comparisons of young children versus infants in task-evoked activation analyses. In other words, a poorly fit HRF would miss activation in the infants; then, when comparing to older children, it would appear as though only the older children exhibited activation in response to the stimulus, whereas the infants did not. Also, deactivation in infant primary sensory cortices has been observed in passive response to auditory (Anderson et al. 2001) and visual stimuli (Born et al. 1998; Martin et al. 1999), suggesting that neurovascular coupling may be quantitatively and qualitatively different in infants compared with older ages. In addition, scanner background noise may have a bearing on infants’ auditory

(and potentially visual) responses, as the neural processes underlying inhibition of scanner background noise may differ between infants and older children. Several approaches have been suggested to mitigate differences in hemodynamic responses between older children (age range: 7-12 years) and adults (Thomason et al. 2005), which broadly focus on ways to account for increased signal noise in children, and it is possible that some of these approaches can be applied to the early developmental age range as well.

In resting-state studies, the relationship between spontaneous fluctuations in hemodynamic signal and neural activity has also been investigated in adult mice (Mateo et al. 2017; Winder et al. 2017); but to our knowledge, this relationship remains to be characterized in young pups. Therefore, it is not entirely clear whether estimates of functional connectivity can similarly (as with estimates of fit in task-evoked designs) be affected by developmental differences in neurovascular coupling. Overall, neurovascular coupling across the 0-5 age range needs further elucidation (Hagmann et al. 2012) and quantification.

4.7 Controlling for Age Variance

When performing categorical comparisons between or among developmental stages, within-group variance in age also merits consideration. Specifically, the rapid neuroanatomical and neurophysiological changes early in development (Gilmore et al. 2018) can cause comparable age-related variances at early and late developmental stages to have markedly different consequences (e.g., \pm two months when the child is four months represents greater rel-

ative variance compared with \pm two months when the child is four years). Adjusting variables of interest (e.g., tissue volumes, functional connectivity

estimates) for within-group variance in age, therefore, may improve second-level models.

5 Decisions for the Field

Rapid changes to brain anatomy and physiology between birth and age five pose dozens of challenges for developmental MRI studies (Fig. 2). Acquisition and processing for children of different ages are often optimized using different tools, techniques, and parameters, which can lead to bias when conducting developmental comparisons. Broadly, this review has underscored and described the acquisition, processing, and analysis steps where these biases may emerge. For each step, the field must decide whether to optimize methods for the age-specific neurobiology, preserve methods across developmental stages, or employ an intermediate of these options.

Specific to data acquisition, we encourage the field to recognize that different states of consciousness can bias data interpretation. We urge the field to recommend whether voxel sizes should scale up with age and, for structural scans, whether infants should be scanned with T2 parameters while older children undergo T1 imaging, or whether we should use identical sequence parameters for all children. We urge the field to recommend how to account for differences in head coil fit and scanner and equipment upgrades during longitudinal studies, especially since upgrades can improve image quality. For fMRI, we must

address age-appropriateness of scanner tasks, including switches from passive to active tasks and difficulty levels. For determinations of sample size calculations, we need to take into account that higher attrition and data loss are likely for children 0-5 years of age, compared with developmental stages beyond five years.

Specific to data processing and analysis, the field needs to recommend whether to perform segmentation and image registration using age-appropriate tools, templates, tissue priors, and intensity inhomogeneity correction, or using age-independent practices. Notwithstanding, we encourage the field to acknowledge that tissue misclassifications and image misalignments will occur more frequently in earlier developmental stages and developmental stages scanned with age-inappropriate acquisition practices. The field also needs to recommend whether to use age-specific ROIs and parcellations, which capture age-specific structural and functional architecture, or a standard set of ROIs and parcellations that are subsequently registered to each developmental stage. Additionally, the field needs to recommend whether to smooth brain images with kernels that optimize SNR, spatial specificity, or the compromise between SNR and spa-

tial specificity for each age group, or to smooth with identically sized kernels regardless of the developmental stage. We also must decide whether to control for head motion using thresholds that are commensurate with the degree of head motion, head size, and therefore age of the child, or use identical thresholds across developmental stages. For fMRI, the field needs to recommend whether to employ age-appropriate frequency filters and HRFs. For statistical models, we should adjust variables of interest for within-group variance in age when performing comparisons across developmental stages.

Importantly, considerable care needs to be taken if employing age-specific approaches, to ensure that they are optimized for the anatomy and physiology unique to the age. We suggest two main approaches. The first is to rely on independently tested and reported methods that have been proven to be reliable for a particular age group. This approach could be useful for new investigators, who could follow the methodological approaches carefully developed, tested, and evaluated by another lab. If no “best practices” exist for the age of interest, then investigators need to adjust the parameters of their pipelines and carefully evaluate the fit and quality of the applied methods subject-by-subject. There are already a number of tools that can aid in determining

success; e.g., Qoala-T rates tissue labelling from FreeSurfer segmentations (Klapwijk et al. 2019), though investigators may need to train its algorithms using a local dataset. In the absence of tools, independent raters can serve as an effective alternative.

Overall, the field must decide on best practices to mitigate bias in developmental comparisons within individual studies. Further, standardizing these practices will also improve reproducibility across studies. Complex pipelines, such as those implemented in developmental studies, present investigators with acquisition, processing, analysis, and interpretation options at multiple stages, exponentially increasing the number of pipeline permutations (Fig. 3) and hindering reproducibility (Poldrack et al. 2017; Karayumak et al. 2019), which has become of major concern as part of the Open Science and Reproducibility movement (Poldrack et al. 2017). Importantly, these permutations may differ by study; therefore, the field must also clarify the contexts that would justify alternative pipelines (e.g., investigators may use smaller voxel sizes for younger compared with older children for structural, but not functional, pipelines). Ultimately, dissemination of guidelines should prune pipeline permutations and hopefully increase reproducibility.

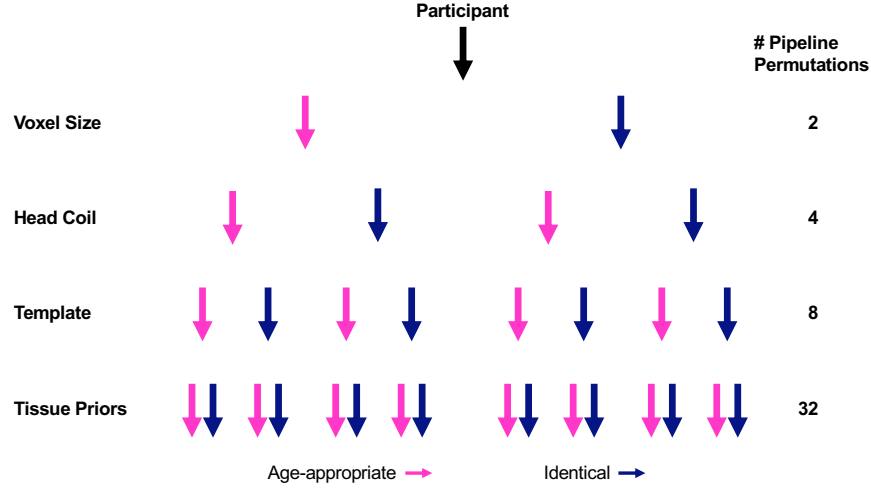


Figure 3. Abbreviated hypothetical decision tree. For multiple stages of a developmental study, investigators must choose between age-appropriate methods (pink) or methods that are identical across developmental stages (blue), exponentially increasing the number of pipeline permutations. Please note that we have limited pipeline permutations here to four stages and binary options.

One final consideration is that the large-scale, known neuroanatomical differences between early developmental stages (e.g., Matsuzawa et al. 2001; Knickmeyer et al. 2008; Groeschel et al. 2010; Sanchez et al. 2011; Geng et al. 2012; Gilmore et al. 2012; Li et al. 2013; Holland et al. 2014) underlie the utility of different methods to optimize images acquired in different developmental stages, and ultimately drive the trade-off between age-optimized methods or standardized methods across developmental stages.

However, the experiments that demonstrate these large-scale neurodevelopmental changes were inevitably conducted in a manner that did not—and perhaps could not—completely eliminate bias, and it is unclear to what extent these biases may have affected reported results. Accordingly, circular reasoning emerges as our ability to measure the brain accurately and unbiasedly has been based in part on results that were acquired with methods that also suffered from this bias.

6 Conclusions

The first years of life are marked by the most rapid and widespread changes in brain anatomy and physiology (Gilmore et al. 2018), with genetic or environmental experiences during this time shaping or altering typical developmental trajectories (Nelson and Gabard-durnam 2020). Several laboratories have begun to use MRI to elucidate these trajectories (e.g., Matsuzawa et al. 2001; Knickmeyer et al. 2008; Geng et al. 2012; Gilmore et al. 2012; Li et al. 2013; Holland et al. 2014), but the rapid growth during this period poses substantial challenges for future developmental studies with methodological bias emerging in several places throughout data acquisition, processing, and analysis. One consistent theme of these specific challenges relates to a trade-off between optimizing approaches for each developmental stage, which inherently requires that multiple methods be used, or maintaining the same approaches across stages, which would likely alter the accuracy and precision in the es-

timates of brain structure or function across developmental stages. While the present review focused on the evaluation of different challenges for the field of developmental neuroimaging, further research is required for the construction of guidelines and approaches that are well-suited for the population of children ages 0-5 years. In addition, the utility of MRI in elucidating developmental trajectories is limited; consequently, there is a need for analogous descriptions of challenges for other modalities such as EEG, MEG, and fNIRS. We also encourage others to consider similar “calls to action” and to develop recommendations for the inchoate field of prenatal neuroimaging (van den Heuvel and Thomason 2016) and for infants born preterm, whose brain structure has been shown to differ from infants born at term (Giménez et al. 2008). To conclude, this review aims to underscore the challenges of developmental MRI studies of early childhood and to inform subsequent resolutions to these challenges.

7 Acknowledgments

We thank our funding sources for enabling us to write this review: Eunice Kennedy Shriver National Institute of Child Health and Human Development R01HD65762, Jacobs Advanced Research Fellowship, and NIH/Harvard Medical School Catalyst 5UL1RR025758 (all awarded to NG). We also thank Ola Ozernov-Palchik, Xi Yu, Jennifer Zuk, Jade Dunstan, Carolyn King, Kathryn Garrisi, and Joseph Sanfilippo for conceptual feedback and additional edits.

8 References

- Akiyama LF, Richards TR, Imada T, Dager SR, Wroblewski L, Kuhl PK. 2013. Age-Specific Average Head Template for Typically Developing 6-Month-Old Infants. 8:1–9.
- Alcauter S, Lin W, Smith JK, Goldman BD, Reznick JS, Gilmore JH, Gao W. 2015. Frequency of spontaneous BOLD signal shifts during infancy and correlates with cognitive performance. *Dev Cogn Neurosci*. 12:40–50.
- Almli CR, Rivkin MJ, McKinstry RC. 2007. The NIH MRI study of normal brain development (Objective-2): Newborns , infants , toddlers , and preschoolers. *Hum Brain Mapp J*. 35:308–325.
- Altaie M, Holland SK, Wilke M, Gaser C. 2008. Infant brain probability templates for MRI segmentation and normalization. *Neuroimage*. 43:721–730.
- Anderson AW, Ph D, Marois R, Ph D, Colson ER, D M, Peterson BS, D M, Duncan CC, D M, Ehrenkranz RA, D M, Schneider KC, H MP, Gore JC, Ph D, Ment LR, D M. 2001. Neonatal auditory activation detected by functional magnetic resonance imaging. 19:1–5.
- Antonov NK, Ruzal-shapiro CB, Morel KD, Millar WS, Kashyap S, Lauren CT, Garzon MC. 2017. Feed and Wrap MRI Technique in Infants.
- Arichi T, Fagiolo G, Varela M, Melendez-Calderon A, Allievi A, Merchant N, Tusor N, Counsell SJ, Burdet E, Beckmann CF, Edwards AD. 2012. Development of BOLD signal hemodynamic responses in the human brain. *Neuroimage*. 63:663–673.
- Arslan S, Makropoulos A, Robinson EC, Rueckert D, Parisot S. 2018. NeuroImage Human brain mapping: A systematic comparison of parcellation methods for the human cerebral cortex. 170:5–30.
- Ashburner J, Friston KJ. 2005. Unified segmentation. *Neuroimage*. 26:839–851.
- Bastiani M, Andersson JLR, Cordero-grande L, Murgasova M, Hutter J, Price AN, Makropoulos A, Fitzgibbon SP, Hughes E, Rueckert D, Victor S, Rutherford M, Edwards AD, Smith SM, Tournier J, Hajnal J V, Jbabdi S, Sotiropoulos SN. 2019. NeuroImage Automated processing pipeline for neonatal diffusion MRI in the developing Human Connectome Project. *Neuroimage*. 185:750–763.
- Baxter L, Fitzgibbon S, Moultrie F, Goksan S, Jenkinson M, Smith S, Andersson J, Duff E, Slater R. 2019. NeuroImage Optimising neonatal fMRI data analysis: Design and validation of an extended dHCP preprocessing pipeline to characterise noxious-evoked brain activity in infants. *Neuroimage*. 186:286–300.
- Beare RJ, Chen J, Kelly CE, Alexopoulos D, Smyser CD, Rogers CE, Loh WY, Matthews LG, Cheong JLY, Spittle AJ, Anderson PJ, Doyle LW, Inder TE, Seal ML, Thompson DK. 2016. Neonatal Brain Tissue Classification with Morphological Adaptation and Unified Segmentation. *Front Neuroinform*. 10:1–17.
- Born A, Miranda M, Rostrup E, Toft P, Peitersen B, Larsson H, Lou H. 2000. Functional Magnetic Resonance Imaging of the Normal and Abnormal Visual

System in Early Life. *Neuropediatrics*. 31:24–32. Born P, Leth H, Miranda MJ, Rostrup E, Stensgaard A, Peitersen B, Larsson HBW, Lou HC. 1998. Visual Activation in Infants and Young Children Studied by Functional Magnetic Resonance Imaging. 44:578–583.

Buckner RL, Hrien FM, Yeo TBT. 2013. Opportunities and limitations of intrinsic functional connectivity MRI. *Nat Rev Neurosci*. 16:832–837.

Centers for Disease Control and Prevention. 2001. Data Table of Infant Head Circumference-for-Age Charts.

Christiaens D, Cordero-grande L, Pietsch M, Hutter J, Price AN, Hughes EJ, Vecchiato K, Deprez M, Edwards AD, Hajnal J V, Tournier J. 2020. Scattered slice SHARD reconstruction for motion correction in multi-shell diffusion MRI. *arXiv*.

Coelho-santos V. 2019. Postnatal development of cerebrovascular structure and the neurogliovascular unit. 1–20.

Coelho-Santos V, Shih AY. 2019. Postnatal development of cerebrovascular structure and the neurogliovascular unit. *WIREs Dev Biol*. e363:1–20.

Cole MW, Bassett DS, Power JD, Braver TS, Petersen SE. 2014. Intrinsic and task-evoked network architectures of the human brain. *Neuron*. 83:238–251.

Counsell SJ, Rutherford MA. 2002. Magnetic resonance imaging of the newborn brain.

Courchesne E, Chisum HJ, Townsend J, Cowles A, Covington J, Egaas B, Harwood M, Hinds S, Press GA, Address JT. 2000. Normal Brain Development and Aging: Quantitative Analysis at in Vivo MR. *Neuroradiology*. 216:672–682.

Cusack R, Wild C, Linke AC, Arichi T, Lee DSC, Han VK. 2015. Optimizing stimulation and analysis protocols for neonatal fMRI. *PLoS One*. 10:1–13.

Dai Y, Shi F, Wang L, Wu G. 2013. iBEAT: A Toolbox for Infant Brain Magnetic Resonance Image Processing. *Neuroinformatics*. 11:211–225.

de Macedo Rodrigues K, Ben-Avi E, Sliva DD, Choe M, Drott M, Wang R, Fischl B, Grant PE, ZÅ¶lle L. 2015. A FreeSurfer-compliant consistent manual segmentation of infant brains spanning the 0-2 year age range. *Front Hum Neurosci*. 9:1–12.

Dean DC, Dirks H, Muircheartaigh JO, Walker L, Jerskey BA, Lehman K, Han M, Waskiewicz N, Deoni SCL. 2014. Pediatric neuroimaging using magnetic resonance imaging during non-sedated sleep. *Pediatr Radiol*. 44:64–72.

Deen B, Richardson H, Dilks DD, Takahashi A, Keil B, Wald LL, Kanwisher N, Saxe R. 2017. Organization of high-level visual cortex in human infants. *Nat Commun*. 8:1–10.

Deoni SCL, Dean DC, Muircheartaigh JO, Dirks H, Jerskey BA. 2012. NeuroImage Investigating white matter development in infancy and early childhood using myelin water fraction and relaxation time mapping. 63:1038–1053.

Devi CN, Chandrasekharan A, Sundararaman VK, Alex ZC. 2015. Neonatal brain MRI segmentation: A review. *Comput Biol Med*. 64:163–178.

Dhankhar S, Tyagi S, Prasad T. 2010. Brain MRI Segmentation using K-means Algorithm. In: *National Conference on Advances in Knowledge*

Management. p. 1–5.

Doria V, Beckmann CF, Arichi T, Merchant N, Groppo M, Turkheimer FE, Counsell SJ, Murgasova M, Aljabar P, Nunes RG, Larkman DJ, Rees G, Edwards AD. 2010. Emergence of resting state networks in the preterm human brain. *Proc Natl Acad Sci.* 107:20015–20020.

Dosenbach NUF, Nardos B, Cohen AL, Fair DA, Power JD, Church JA, Nelson SM, Wig GS, Vogel AC, Lessov-schlaggar CN, Barnes KA, Dubis JW, Feczko E, Rebecca S, Jr JRP, Barch DM, Petersen SE, Bradley L. 2010. Prediction of Individual Brain Maturity Using fMRI Published by: American Association for the Advancement of Science Linked references are available on JSTOR for this article: Prediction of Individual Brain Maturity Using fMRI. *Science* (80-). 329:1358–1361.

Dubois J, Alison M, Counsell SJ, Benders MJNL. 2020. MRI of the Neonatal Brain: A Review of Methodological Challenges and Neuroscientific Advances. *J Magn Reson imaging.* 1–26.

Edelstein W, Glover G, Hardy C, Redington R. 1986. The Intrinsic Signal-to-Noise Ratio in NMR Imaging. *Magn Reson Med.* 3:604–618. Ellis C, Skalaban L, Yates T, Bejjanki V, Cordova N, Turk-Browne N. 2020. Re-imagining fMRI for awake behaving infants. *Nat Commun.* 11:1–12.

Fair DA, Dosenbach NUF, Church JA, Cohen AL, Brahmbhatt S, Miezin FM, Barch DM, Raichle ME, Petersen SE, Schlaggar BL. 2007. Development of distinct control networks through segregation and integration. *Proc Natl Acad Sci U S A.* 104:13507–13512. Fair DA, Cohen AL, Power JD, Dosenbach NUF, Church JA, Miezin FM, Schlaggar BL, Petersen SE. 2009. Functional Brain Networks Develop from a “Local to Distributed” Organization. *PLoS Comput Biol.* 5:1–16.

Ferradal SL, Gagoski B, Jaimes C, Yi F, Carruthers C, Vu C, Litt JS, Larsen R, Sutton B, Grant PE, Zöllei L. 2018. System-Specific Patterns of Thalamocortical Connectivity in Early Brain Development as Revealed by Structural and Functional MRI. *Cereb Cortex.* 29:1218–1229.

Fonov V, Evans A, McKinstry R, Almli C, Collins D. 2009. Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *Neuroimage.* 47:S102.

Fransson P, Åden U, Blennow M, Lagercrantz H. 2011. The functional architecture of the infant brain as revealed by resting-state fMRI. *Cereb Cortex.* 21:145–154.

Fransson P, Metsaranta M, Blennow M, Ulrika A, Lagercrantz H, Vanhatalo S. 2013. Early Development of Spatial Patterns of Power-Law Frequency Scaling in fMRI Resting-State and EEG Data in the Newborn Brain. *Cereb Cortex.* 23:638–646.

Fransson P, Skiold B, Engstro M, Hallberg B, Mosskin M, Åden U, Fransson P, Skio B, Lagercrantz H, Blennow M. 2009. Spontaneous Brain Activity in the Newborn Brain During Natural Sleep — An fMRI Study in Infants Born at Full Term. *Pediatr Res.* 66:301–305.

Fransson P, Skiold B, Horsch S, Nordell A, Blennow M, Lagercrantz H, Aden U. 2007. Resting-state networks in the infant brain. *Proc Natl Acad Sci USA.*

104:15531–15536.

Gaab N, Gabrieli JDE, Glover GH. 2007a. Assessing the Influence of Scanner Background Noise on Auditory Processing . I . An fMRI Study Comparing Three Experimental Designs with Varying Degrees of Scanner Noise. *Hum Brain Mapp.* 28:703–720.

Gaab N, Gabrieli JDE, Glover GH. 2007b. Assessing the Influence of Scanner Background Noise on Auditory Processing . II . An fMRI Study Comparing Auditory Processing in the Absence and Presence of Recorded Scanner Noise Using a Sparse Design. *Hum Brain Mapp.* 28:721–732.

Gaab N, Gabrieli JDE, Glover GH. 2008. Resting in Peace or Noise: Scanner Background Noise Suppresses Default-Mode Network. *Hum Brain Mapp.* 29:858–867.

Gaab N, Gaser C, Zaehle T, Jancke L, Schlaug G. 2003. Functional anatomy of pitch memory — an fMRI study with sparse temporal sampling. *Neuroimage.* 19:1417–1426.

Gabard-Durnam LJ, Muircheartaigh JO, Dirks H, Dean D, Tottenham N, Deoni S. 2018. Human amygdala functional network development: A cross-sectional study from 3 months to 5 years of age. *Dev Cogn Neurosci.* 34:63–74.

Gao W, Alcauter S, Elton A, Hernandez-Castillo CR, Smith JK, Ramirez J, Lin W. 2015. Functional network development during the first year: Relative sequence and socioeconomic correlations. *Cereb Cortex.* 25:2919–2928. Gao W, Alcauter S, Smith JK, Gilmore JH, Lin W. 2015. Development of human brain cortical network architecture during infancy. *Brain Struct Funct.* 220:1173–1186.

Gao W, Zhu H, Giovanello K, Smith JK, Shen D, Gilmore JH, Lin W. 2009. Evidence on the emergence of the brain ’ s default network from 2-week-old to 2-year-old healthy pediatric subjects. *Proc Natl Acad Sci.* 106:6790–6795.

Geng X, Gouttard S, Sharma A, Gu H, Styner M, Lin W, Gerig G, Gilmore JH. 2012. NeuroImage Quantitative tract-based white matter development from birth to age 2 years. *Neuroimage.* 61:542–557.

Gilmore J, Knickmeyer RC, Gao W. 2018. Imaging structural and functional brain development in early childhood. *Nat Rev Neurosci.* 19:123–137.

Gilmore JH, Shi F, Woolson SL, Knickmeyer RC, Short SJ, Lin W, Zhu H, Hamer RM, Styner M, Shen D. 2012. Longitudinal development of cortical and subcortical gray matter from birth to 2 years. *Cereb Cortex.* 22:2478–2485.

Giménez M, Miranda MJ, Born AP, Nagy Z, Rostrup E, Jernigan TL. 2008. Authors Accelerated cerebral white matter development in preterm infants: A voxel-based morphometry study with diffusion tensor MR imaging. *Neuroimage.* 41:728–734.

Glover GH, Mueller BA, Turner JA, Erp TGM Van, Liu TT, Greve DN, Voyvodic JT, Rasmussen J, Brown GG, Keator DB, Calhoun VD, Lee HJ, Ford JM, Mathalon DH, Diaz M, Leary DSO, Gadde S, Preda A, Lim KO, Wible CG, Stern HS, Belger A. 2012. Function Biomedical Informatics Research Network Recommendations for Prospective Multicenter Functional MRI Studies. 54:39–54.

Graham AM, Buss C, Rasmussen JM, Rudolph MD, Demeter D V., Gilmore

JH, Styner M, Entringer S, Wadhwa PD, Fair DA. 2016. Implications of newborn amygdala connectivity for fear and cognitive development at 6-months-of-age. *Dev Cogn Neurosci.* 18:12–25.

Graham AM, Pfeifer JH, Fisher PA, Carpenter S, Fair DA. 2015. Early life stress is associated with default system integrity and emotionality during infancy. *J Child Psychol Psychiatry Allied Discip.* 56:1212–1222.

Graham AM, Pfeifer JH, Fisher PA, Lin W, Gao W, Fair DA. 2015. The potential of infant fMRI research and the study of early life stress as a promising exemplar. *Dev Cogn Neurosci.* 12:12–39.

Gratton C, Laumann TO, Nielsen AN, Schlaggar BL, Dosenbach NUF, Petersen SE, Gratton C, Laumann TO, Nielsen AN, Greene DJ, Gordon EM. 2018. Functional Brain Networks Are Dominated by Stable Group and Individual Factors , Not Cognitive or Daily Article Functional Brain Networks Are Dominated by Stable Group and Individual Factors , Not Cognitive or Daily Variation. 439–452.

Grayson DS, Fair DA. 2017. Development of large-scale functional networks from birth to adulthood: A guide to the neuroimaging literature. *Neuroimage.* 160:15–31.

Groeschel S, Vollmer B, King MD, Connelly A. 2010. International Journal of Developmental Neuroscience Developmental changes in cerebral grey and white matter volume from infancy to adulthood. *Int J Dev Neurosci.* 28:481–489.

Hagmann P, Grant PE, Fair DA, Knösche TR, Planck M. 2012. MR connectomics: a conceptual framework for studying the developing brain. 6:1–17.

Harb R, Whiteus C, Freitas C, Grutzendler J. 2013. In vivo imaging of cerebral microvascular plasticity from birth to death. 106815:146–156.

Hart G, Shi Y, Zhu H, Sanchez M, Styner M, Niethammer M. 2010. DTI Longitudinal Atlas Construction as an Average of Growth Models. PrePrint.

Heuvel MI Van Den, Thomason ME. 2016. Functional Connectivity of the Human Brain in Utero. *Trends Cogn Sci.* 20:931–939.

Holland D, Chang L, Ernst TM, Curran M, Buchthal SD, Alicata D, Skranes J, Johansen H, Hernandez A, Yamakawa R, Kuperman JM, Dale AM. 2014. Structural growth trajectories and rates of change in the first 3 months of infant brain development. *JAMA Neurol.* 71:1266–1274.

Horovitz SG, Braun AR, Carr WS, Picchioni D, Balkin TJ, Fukunaga M, Duyn JH. 2009. Decoupling of the brain’s default mode network during deep sleep. *Proc Natl Acad Sci.* 106:11376–11381.

Howell BR, Styner MA, Gao W, Yap P, Wang L, Baluyot K, Yacoub E, Chen G, Potts T, Salzwedel A, Li G, Gilmore JH, Piven J, Smith JK, Shen D, Ugurbil K, Zhu H, Lin W, Ellison JT. 2019. The UNC / UMN Baby Connectome Project (BCP): An overview of the study design and protocol development. *Neuroimage.* 185:891–905.

Hughes EJ, Winchman T, Padormo F, Teixeira R, Wurie J, Sharma M, Fox M, Hutter J, Cordero-grande L, Price AN, Allsop J, Bueno-conde J, Tusor N, Arichi T, Edwards AD, Rutherford MA, Counsell SJ, Hajnal J V. 2017. A Dedicated Neonatal Brain Imaging System. 804:794–804.

Ibrahim T, Few K, Greenwood R, Smith C, Malcolm P, Johnson G, Lally P, Thayyil S, Clarke P. 2015. ‘Feed and wrap’ or sedate and immobilise for neonatal brain MRI? *Arch Dis Child Fetal Neonatal* Ed. 100:465–466.

Karayumak SC, Bouix S, Ning L, James A, Crow T, Shenton M, Kubicki M, Rath Y. 2019. NeuroImage Retrospective harmonization of multi-site diffusion MRI data acquired with different acquisition parameters. *Neuroimage*. 184:180–200.

Keil B, Alagappan V, Mareyam A, McNab JA, Fujimoto K, Tountcheva V, Triantafyllou C, Dilks DD, Kanwisher N, Lin W, Grant PE, Wald LL. 2011. Size-optimized 32-Channel Brain Arrays for 3 T Pediatric Imaging. *Magn Reson Med*. 66:1777–1787.

Klapwijk ET, Kamp F Van De, Meulen M Van Der, Peters S, Wierenga LM. 2019. Qoala-T: A supervised-learning tool for quality control of FreeSurfer segmented MRI data. *Neuroimage*. 189:116–129.

Knickmeyer RC, Gouttard S, Kang C, Evans D, Wilber K, Smith JK, Hamer RM, Lin W, Gerig G, Gilmore JH. 2008. A Structural MRI Study of Human Brain Development from Birth to 2 Years. *J Neurosci*. 28:12176–12182.

Langer N, Peysakhovich B, Zuk J, Drott M, Sliva DD, Smith S, Becker BLC, Grant PE, Gaab N. 2017. White Matter Alterations in Infants at Risk for Developmental Dyslexia. *Cereb Cortex*. 27:1027–1036.

Li G, Nie J, Wang L, Shi F, Lin W, Gilmore JH, Shen D. 2013. Mapping region-specific longitudinal cortical surface expansion from birth to 2 years of age. *Cereb Cortex*. 23:2724–2733.

Li G, Wang L, Shi F, Gilmore JH, Lin W, Shen D. 2015. Construction of 4D high-definition cortical surface atlases of infants: Methods and applications. 25:22–36.

Li Z, Park B, Liu W, Zhang J, Reed MP, Rupp JD, Hoff CN, Hu J. 2015. A Statistical Skull Geometry Model for Children 0-3 Years Old. *PLoS*. 1–13.

Liu WC, Flax JF, Guise KG, Sukul V, Benasich AA. 2008. Functional connectivity of the sensorimotor area in naturally sleeping infants. *Brain Res*. 1223:42–49.

Lyll AE, Shi F, Geng X, Woolson S, Li G, Wang L, Hamer RM, Shen D, Gilmore JH. 2015. Dynamic Development of Regional Cortical Thickness and Surface Area in Early Childhood. 2204–2212.

Madhyastha T, Peverill M, Koh N, McCabe C, Flourney J, Mills K, King K, Pfeifer J, McLaughlin KA. 2018. Developmental Cognitive Neuroscience Current methods and limitations for longitudinal fMRI analysis across development. *Dev Cogn Neurosci*. 33:118–128.

Makropoulos A, Counsell SJ, Rueckert D. 2018. NeuroImage A review on automatic fetal and neonatal brain MRI segmentation. *Neuroimage*. 170:231–248.

Makropoulos A, Gousias IS, Ledig C, Aljabar P, Serag A, Hajnal J V, Edwards AD, Counsell SJ, Rueckert D. 2014. Automatic Whole Brain MRI Segmentation of the Developing Neonatal Brain. *IEEE Trans Med Imaging*. 33:1818–1831.

Martin E, Joeri P, Loenneker T, Ekatodramis D, Vitacco D, Hennig J. 1999.

Visual Processing in Infants and Children Studied Using Functional MRI. 46:135–140.

Mateo C, Knutsen PM, Tsai PS, Shih AY, Kleinfeld D. 2017. Entrainment of Arteriole Vasomotor Fluctuations by Neural Activity Is a Basis of Blood-Oxygenation-Level-Dependent “Resting-State” Connectivity. *Neuron*. 96:936–948.

Matsuzawa J, Matsui M, Konishi T. 2001. Age-related Volumetric Changes of Brain Gray and White Matter in Healthy Infants and Children. *Cereb Cortex*. 11:335–342.

Merhar SL, Gozdas E, Tkach JA, Harpster KL, Schwartz TL, Yuan W, Kline-Fath BM, Leach JL, Altaye M, Holland SK. 2016. Functional and structural connectivity of the visual system in infants with perinatal brain injury. *Pediatr Res*. 80:43–48.

Mitra A, Snyder AZ, Tagliazucchi E, Laufs H, Elison J, Emerson RW, Shen MD, Wolff JJ, Botteron KN, Dager S, Estes AM, Evans A, Gerig G, Hazlett HC, Paterson SJ, Schultz RT, Styner MA, Zwaigenbaum L, Network TI. 2017. Resting-state fMRI in sleeping infants more closely resembles adult sleep than adult wakefulness. 1–19.

Mongerson CRL, Jennings RW, Borsook D, Becerra L, Bajic D. 2017. Resting-State Functional Connectivity in the Infant Brain: Methods, Pitfalls, and Potentiality. *Front Pediatr*. 5:1–26.

Nelson CA, Gabard-durnam LJ. 2020. Early Adversity and Critical Periods: Neurodevelopmental Consequences of Violating the Expectable Environment. *Trends Neurosci*. 43:133–143.

Norman M, O’Kusky J. 1986. The Growth and Development of Microvasculature in Human Cerebral Cortex. *J Neuropathol Exp Neurol*. 45:222–232.

Ogawa S, Lee T-M. 1990. Magnetic Resonance Imaging of Blood Vessels at High Fields: In Vivo and in Vitro Measurements and Image Simulation. *Magn Reson Med*. 16:9–18.

Oguz I, Farzinfar M, Matsui J, Budin F, Liu Z, Gerig G, Johnson HJ, Styner M. 2014. DTIPrep: quality control of diffusion-weighted images. *Front Neuroinform*. 8:1–11.

Oishi K, Chang L, Huang H. 2018. Baby brain atlases. *Neuroimage*. 1–16.

Paus T, Collins DL, Evans AC, Leonard G, Pike B, Zijdenbos A. 2001. Maturation of white matter in the human brain: A review of magnetic resonance studies. 54:255–266.

Perrachione TK, Ghosh SS. 2013. Optimized design and analysis of sparse-sampling fMRI experiments. *Front Neurosci*. 7:1–18.

Pfefferbaum A, Mathalon DH, Sullivan E V, Rawles JM, Zipursky RB, Lim KO. 1994. A quantitative magnetic resonance imaging study of changes in brain morphology from infancy to late adulthood. *Arch Neurol*. 51:874–887.

Picchioni D, Duyn JH, Horovitz SG. 2013. Sleep and the functional connectome. *Neuroimage*. 80:387–396.

Poldrack RA, Baker CI, Durnez J, Gorgolewski KJ, Matthews PM, Munaf MR, Nichols TE, Poline J, Vul E, Yarkoni T. 2017. Scanning the horizon:

towards. *Nat Rev Neurosci.* 18:115–126.

Poldrack RA, Baker CI, Durnez J, Gorgolewski KJ, Matthews PM, Munafò MR, Nichols TE, Poline JB, Vul E, Yarkoni T. 2017. Scanning the horizon: Towards transparent and reproducible neuroimaging research. *Nat Rev Neurosci.* 18:115–126.

Pomponio R, Erus G, Habes M, Doshi J, Srinivasan D, Mamourian E, Bashyam V, Nasrallah IM, Satterthwaite TD, Fan Y, Launer LJ, Masters CL, Maruff P, Zhuo C, Johnson SC, Fripp J, Koutsouleris N, Wolf DH, Gur R, Gur R, Morris J, Albert MS, Grabe HJ, Resnick SM, Bryan RN, Wolk DA, Shinohara RT, Shou H, Davatzikos C. 2020. NeuroImage Harmonization of large MRI datasets for the analysis of brain imaging patterns throughout the lifespan. *Neuroimage.* 208:1–15.

Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE. 2012. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage.* 59:2142–2154.

Power JD, Cohen AL, Nelson SM, Wig GS, Barnes KA, Church JA, Vogel AC, Laumann TO, Miezin FM, Schlaggar BL, Petersen SE. 2011. Functional Network Organization of the Human Brain. *Neuron.* 72:665–678.

Prastawa M, Gilmore JH, Lin W, Gerig G. 2005. Automatic segmentation of MR images of the developing newborn brain. *Med Image Anal.* 9:457–466.

Pruett JR, Kandala S, Hoertel S, Snyder AZ, Elison JT, Nishino T, Feczko E, Dosenbach NUF, Nardos B, Power JD, Adeyemo B, Botteron KN, McKinstry RC, Evans AC, Hazlett HC, Dager SR, Paterson S, Schultz RT, Collins DL, Fonov VS, Styner M, Gerig G, Das S, Kostopoulos P, Constantino JN, Estes AM, Petersen SE, Schlaggar BL, Piven J. 2015. Accurate age classification of 6 and 12 month-old infants based on resting-state functional connectivity magnetic resonance imaging data. *Dev Cogn Neurosci.* 12:123–133.

Raichle ME. 1998. Behind the scenes of functional brain imaging: A historical and physiological perspective. *Proc Natl Acad Sci.* 95:765–772.

Raininko R, Autti T, Vanhanen SL, Ylikoski A, Erkinjuntti T, Santavuori P. 1994. The normal brain stem from infancy to old age: A morphometric MRI study. *Neuroradiology.* 36:364–368.

Raschle N, Stering PL, Meissner SN, Gaab N. 2014. Altered neuronal response during rapid auditory processing and its relation to phonological processing in prereading children at familial risk for dyslexia. *Cereb Cortex.* 24:2489–2501.

Raschle N, Zuk J, Gaab N. 2012. Functional characteristics of developmental dyslexia in left-hemispheric posterior brain regions predate reading onset. *Proc Natl Acad Sci.* 109:2156–2161.

Raschle N, Zuk J, Ortiz-Mantilla S, Sliva DD, Franceschi A, Grant PE, Benasich AA, Gaab N. 2012. Pediatric neuroimaging in early childhood and infancy: Challenges and practical guidelines. *Ann N Y Acad Sci.* 1252:43–50.

Raschle NM, Chang M, Gaab N. 2011. Structural brain alterations associated with dyslexia predate reading onset. *Neuroimage.* 57:742–749.

Redpath TW. 1998. Signal-to-noise ratio in MRI. *Br J Radiol.* 71:704–707.

Reynolds JE, Grohs MN, Dewey D, Lebel C. 2019. NeuroImage Global and regional white matter development in early childhood. *Neuroimage.* 196:49–58.

Rudolph MD, Graham AM, Feczko E, Miranda-Dominguez O, Rasmussen JM, Nardos R, Entringer S, Wadhwa PD, Buss C, Fair DA. 2018. Maternal IL-6 during pregnancy can be estimated from newborn brain connectivity and predicts future working memory in offspring. *Nat Neurosci.* 21:765–772.

Runge VM, Nitz WR, Trelles M, Goerner FL. 2013. *Physics of Clinical MR Taught Through Images*. 3rd ed. Thieme Medical Publishers Inc.

Sanchez CE, Richards JE, Almli CR. 2011. *Neurodevelopmental MRI Brain Templates for Children From 2 weeks to 4 Years of Age*.

Satterthwaite TD, Elliott MA, Gerraty RT, Ruparel K, Loughead J, Calkins ME, Eickhoff SB, Hakonarson H, Gur RC, Gur RE, Wolf DH. 2013. An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. *Neuroimage.* 64:240–256.

Satterthwaite TD, Wolf DH, Loughead J, Ruparel K, Elliott MA, Hakonarson H, Gur RC, Gur RE. 2012. Impact of in-scanner head motion on multiple measures of functional connectivity: Relevance for studies of neurodevelopment in youth. *Neuroimage.* 60:623–632.

Saunders DE, Thompson C, Gunny R, Jones R, Cox T. 2007. Magnetic resonance imaging protocols for paediatric neuroradiology. *Pediatr Radiol.* 37:789–797.

Scheinost D, Hyun S, Xilin K, Cheryl S. 2016. Preterm birth alters neonatal, functional rich club organization. *Brain Struct Funct.* 221:3211–3222.

Shen D, Davatzikos C. 2002. HAMMER: Hierarchical Attribute Matching Mechanism for Elastic Registration. *IEEE Trans Med Imaging.* 21:1421–1439.

Shi F, Salzwedel AP, Lin W, Gilmore JH. 2018. Functional Brain Parcellations of the Infant Brain and the Associated Developmental Trends. 1358–1368.

Shi F, Yap PT, Wu G, Jia H, Gilmore JH, Lin W, Shen D. 2011. Infant brain atlases from neonates to 1- and 2-year-olds. *PLoS One.* 6.

Smyser CD, Dosenbach NUF, Smyser TA, Snyder AZ, Rogers CE, Inder TE, Schlaggar BL, Neil JJ. 2016. Prediction of brain maturity in infants using machine-learning algorithms. *Neuroimage.* 136:1–9.

Stange JP, Jenkins LM, Bessette KL, Kling LR, Bark JS, Shepard R, Hamlat EJ, Deldonno S, Phan KL, Passarotti AM, Ajilore O, Langenecker SA. 2018. Predictors of Attrition in Longitudinal Neuroimaging Research: Brain Connect. 8:527–536.

Teague S, Youssef GJ, Macdonald JA, Sciberras E, Shatte A, Fuller-tyszkiewicz M, Greenwood C, McIntosh J, Olsson CA. 2018. Retention strategies in longitudinal cohort studies: a systematic review and meta- analysis. *BMC Med Res Methodol.* 18:1–22.

Thieba C, Frayne A, Walton M, Mah A, Benischeck A, Dewey D, Lebel C. 2018. Factors Associated With Successful MRI Scanning in Unsedated Young Children. *Front Pediatr.* 6:1–5.

Thomason ME, Burrows BE, Gabrieli JDE, Glover GH. 2005. Breath holding reveals differences in fMRI BOLD signal in children and adults. *Neuroimage.* 25:824–837.

Turesky T, Jensen S, Yu X, Kumar S, Wang Y, Sliva D, Gagoski B, Sanfilippo J, Zöllei L, Boyd E, Haque R, Kakon S, Islam N, Petri W, Nelson C, N G.

2019. The relationship between biological and psychosocial risk factors and resting-state functional connectivity in 2-month-old Bangladeshi infants: a feasibility and pilot study. *Dev Sci.* 22:e12841.

Turesky T, Xie W, Kumar S, Sliva DD, Gagoski B, Vaughn J, Lilla Z, Petri WA, Nelson CA, Gaab N. 2020. Relating anthropometric indicators to brain structure in 2-month-old Bangladeshi infants growing up in poverty: A pilot study. *Neuroimage.* 210:1–10.

Turkeltaub PE, Eden GF, Jones KM, Zeffiro TA. 2002. Meta-Analysis of the Functional Neuroanatomy of Single-Word Reading: Method and Validation. *Neuroimage.* 16:765–780.

Tustison NJ, Avants BB, Cook PA, Zheng Y, Egan A, Yushkevich PA, Gee JC. 2010. N4ITK: Improved N3 bias correction. *IEEE Trans Med Imaging.* 29:1310–1320.

Van Dijk KRA, Hedden T, Venkataraman A, Evans KC, Lazar SW, Buckner RL. 2010. Intrinsic Functional Connectivity As a Tool For Human Connectomics: Theory , Properties , and Optimization. 02138:297–321.

Van Dijk KRA, Sabuncu MR, Buckner RL. 2012. The influence of head motion on intrinsic functional connectivity MRI. *Neuroimage.* 59:431–438.

Vanderauwera J, Altarelli I, Vandermosten M, Vos A De, Wouters J, Ghesquière P. 2018. Atypical Structural Asymmetry of the Planum Temporale is Related to Family History of Dyslexia. *Cereb Cortex.* 28:63–72.

Vanderauwera J, Wouters J, Vandermosten M. 2017. Developmental Cognitive Neuroscience Early dynamics of white matter deficits in children developing dyslexia. *Dev Cogn Neurosci.* 27:69–77.

Vasung L, Turk EA, Ferradal SL, Sutin J, Stout JN, Ahtam B, Lin P, Grant PE. 2019. Exploring early human brain development with structural and physiological neuroimaging. *Neuroimage.* 187:226–254.

Vijayakumar N, Mills KL, Alexander-bloch A, Tamnes CK, Whittle S. 2018. Developmental Cognitive Neuroscience Structural brain development: A review of methodological approaches and best practices. *Dev Cogn Neurosci.* 33:129–148.

Wilke M, Schmithorst VJ, Holland SK. 2003. Normative Pediatric Brain Data for Spatial Normalization and Segmentation Differs From Standard Adult Data. *Magn Reson Med.* 50:749–757.

Winder AT, Echagarruga C, Zhang Q, Drew PJ. 2017. Weak correlations between hemodynamic signals and ongoing neural activity during the resting state. *Nat Neurosci.* 20:1761–1769.

Wu J, Avants B. 2012. Automatic Registration-Based Segmentation for Neonatal Brains Using ANTs and Atropos. *MICCAI Gd Chall Neonatal Brain Segmentation.* 36. Yeatman JD, Dougherty RF, Myall NJ, Wandell BA, Feldman HM. 2012. Tract Profiles of White Matter Properties: Automating Fiber-Tract Quantification. *PLoS One.* 7.

Yin W, Li T, Hung S, Zhang H, Wang L, Shen D, Zhu H. 2020. The emergence of a functionally flexible brain during early infancy. *Proc Natl Acad Sci.* 117:23904–23913.

Yu H, Schummers J, Sur M. 2010. The influence of astrocyte activation on hemodynamic signals for functional brain imaging. In: Roe A, editor. *Imaging*

the Brain with Optical Methods. Springer Science+Business Media.

Yu X, Raney T, Perdue M V, Zuk J, Ozernov-palchik O, Becker BLC, Raschle NM, Gaab N. 2018. Emergence of the neural network underlying phonological processing from the prereading to the emergent reading stage: A longitudinal study. *Hum Brain Mapp.* 39:2047–2063.

Zahr L, de Traversay J. 1995. Premature infant responses to noise reduction by earmuffs: effects on behavioral and physiologic measures. *J Perinatol.* 15:448–455.

Zhang H, Shen D, Lin W. 2019. Resting-state Functional MRI Studies on Infant Brains: a Decade of Gap-Filling Efforts *Hum. Neuroimage.* 185:664–684.

Zhang Y, Shi F, Wu G, Wang L, Yap P, Shen D. 2016. Consistent Spatial-Temporal Longitudinal Atlas Construction for Developing Infant Brains. *IEEE Trans Med Imaging.* 35:2568–2577.

Zöllei L, Iglesias JE, Ou Y, Grant PE, Fischl B. 2020. Infant FreeSurfer: An automated segmentation and surface extraction pipeline for T1-weighted neuroimaging data of infants 0-2 years. *Neuroimage.*