Neuroimaging of neonatal nociception

Luke Baxter

St Cross College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

March 2020

Abstract

Newborn infants undergo painful procedures during standard clinical care. However, pain assessment in non-verbal infants is challenging. Improving understanding of infant pain neurophysiology, including anticipating pain responses, is central to improving pain management. Magnetic resonance imaging (MRI) is an ideal technique for examining the structure and function of the infant pain system, allowing whole-brain coverage at relatively high spatial resolutions. However, major differences between the adult and infant brain necessitate the use of infant-optimised data analysis methods.

The first study involved outlining and validating an optimised infant noxious-response functional MRI data analysis pipeline to improve data quality. Using a cohort of 15 healthy newborn infants, five key analysis steps were assessed. Adopting the Developing Human Connectome Project functional MRI preprocessing pipeline improved motion and distortion correction and spatial normalisation outputs, and including spatial ICA-based denoising and minimal spatial filtering furthered these improvements. Of three term infant haemodynamic response function parameterisations, the double gamma function was most suitable. Relative to a standard adult analysis pipeline adapted for infants, the infant-optimised pipeline significantly improved data quality.

The second study involved predicting infants' cerebral haemodynamic noxious-response amplitudes from resting-state activity demonstrating the potential for infant-unique brain activity to inform infant pain management. Using a cohort of 18 healthy newborn infants, nine resting-state networks were robustly identifiable, and the infants' resting-state network amplitudes predicted noxious-response amplitudes with statistically significant accuracy. Noxious-response amplitudes were related to white matter complexity using diffusion MRI, suggesting the response amplitudes were infant trait effects. Finally, the specific pattern of functional and structural neural correlates of the haemodynamic noxious-response amplitudes suggested the responses reflected brain maturity.

This work advances our understanding of the newborn infant's pain system, underscores the central importance of MRI in furthering this basic understanding, and helps form the foundations on which further infant pain research can be conducted.

Acknowledgements

Thank you to my supervisors Rebeccah, Eugene, and Saad. They provided continued support, encouragement, advice, and guidance on all aspects of my DPhil, while simultaneously giving me the freedom to pursue lines of enquiry that excite me. Thank you to all members of the Paediatric Neuroimaging Group who provided a warm, welcoming, and encouraging atmosphere, especially Sezgi Goksan who took me under her wing in the earliest rockiest days.

Thank you to my family who have been so positive and patient during the past three years of DPhil chaos. You will each be rewarded with a personal copy of my thesis, signed upon request. You are welcome.

Thank you to my friends who provided much needed escape from DPhil pressure. Special thanks to Malte, Michael, and Niall for much needed and much appreciated chats and beers. Theses for all of you lucky men!

Finally, thank you to my long-time partner and friend, the lovely Grace. While I can't imagine what it would have been like without you by my side, I can honestly say it has been wonderful with you there. I feel very lucky indeed.

Contents

List of Figures

List of Tables

List of Abbreviations

2DFT . . . . . Two-Dimensional Fourier Transform
ACC . . . . . Anterior Cingulate Cortex
AD . . . . . Axial Diffusivity
ANl . . . . . Left Auditory Network
ANr . . . . . Right Auditory Network
ANTs . . . . . Advanced Normalization Tools
AP . . . . . Anterior-Posterior
ASL . . . . . Arterial Spin Labeling
B0 . . . . . Main magnetic field
B1 . . . . . Radiofrequency excitation field
BIDS . . . . . Brain Imaging Data Structure
BEST . . . . . Biomarkers, EndpointS, and other Tools
BOLD . . . . . Blood Oxygen Level Dependent
BBR . . . . . Boundary-Based Registration
DAN . . . . . Dorsal Attention Network
DG . . . . . Double Gamma
dHCP . . . . . Developing Human Connectome Project
DIAPR . . . . . Development of Infant Acute Pain Responding
DKI . . . . . Diffusion Kurtosis Imaging
DMN . . . . . Default Mode Network
dMRI . . . . . Diffusion Magnetic Resonance Imaging
DOF . . . . . Degrees Of Freedom
Draw-EM . . . Developing brain Region Annotation With Expectation-Maximization
DTI . . . . . Diffusion Tensor Imaging
ECN . . . . . Executive Control Network
EEG . . . . . Electroencephalography
EPI . . . . . Echo Planar Imaging
FA . . . . . Fractional Anisotropy
FAST . . . . . FMRIB's Automated Segmentation Tool
FDR . . . . . False Discovery Rate

FEAT . . . . . FMRI Expert Analysis Tool
FILM . . . . . FMRIB's Improved Linear Model
FIX . . . . . FMRIB's ICA-based Xnoiseifier
FLIRT . . . . . FMRIB's Linear Image Registration Tool
FLOBS . . . . . FMRIB's Linear Optimal Basis Set
fMRI . . . . . Functional Magnetic Resonance Imaging
FSL . . . . . FMRIB Software Library
FUGUE . . . . FMRIB's Utility for Geometrically Unwarping EPIs
FWER . . . . . Family-Wise Error Rate
FWHM . . . . . Full Width at Half Maximum
GA . . . . . Gestational Age
GABA . . . . . Gamma-Aminobutyric Acid
GLM . . . . . General Linear Model
GRAPPA . . . GeneRalized Autocalibrating Partial Parallel Acquisition
GRE . . . . . Gradient Recalled Echo
HRF . . . . . Hemodynamic Response Function
IASP . . . . . International Association for the Study of Pain
ICA . . . . . Independent Component Analysis
MAD . . . . . Median Absolute Deviation
MCDC . . . . . Motion Correction and Distortion Correction
MCFLIRT . . . Motion Correction FLIRT
MD . . . . . Mean Diffusivity
MIRTK . . . . . Medical Image Registration ToolKit
MK . . . . . Mean Kurtosis
MLR . . . . . Multiple Linear Regression
mN . . . . . Millinewton
MRI . . . . . Magnetic Resonance Imaging
NIRS . . . . . Near Infrared Spectroscopy
NMI . . . . . Normalised Mutual Information
NODDI . . . . . Neurite Orientation Dispersion and Density Imaging
NPC . . . . . Non-Parametric Combination
ObsRO . . . . . Observer Reported Outcome
PA . . . . . Posterior-Anterior
PAG . . . . . Periaqueductal Grey
PALM . . . . . Permutation Analysis of Linear Models
PCA . . . . . Principle Component Analysis

PFM . . . . . Probabilistic Functional Mode
PIPP . . . . . Premature Infant Pain Profile
PIPP-R . . . . . Revised Premature Infant Pain Profile
PMA . . . . . Postmenstrual age
PNA . . . . . Postnatal age
PoCG . . . . . Postcentral Gyrus
PRO . . . . . Patient Reported Outcome
PROFUMO . . Probabilistic Functional Mode
RD . . . . . Radial Diffusivity
RF . . . . . Radiofrequency
rs-fMRI . . . . . Resting-state Functional Magnetic Resonance Imaging
RMSE . . . . . Root Mean Squared Error
ROI . . . . . Region Of Interest
RVM . . . . . Rostral Ventral Medial Medulla
SBref . . . . . Single-Band Reference
SCA . . . . . Seed-based Correlation Analysis
SCM . . . . . Social Communication Model
SE . . . . . Spin echo
SG . . . . . Single Gamma
SMN . . . . . Somatomotor network
SNR . . . . . Signal-to-Noise Ratio
SVR . . . . . Support Vector Regression
SyN . . . . . Symmetric image Normalization
TBSS . . . . . Tract-Based Spatial Statistics
T . . . . . Tesla
TE . . . . . Echo Time
TFCE . . . . . Threshold-Free Cluster Enhancement
TNR . . . . . True Negative Rate
TPR . . . . . True Positive Rate
TR . . . . . Repetition Time
TSE . . . . . Turbo Spin Echo
tSNR . . . . . Temporal Signal-to-Noise Ratio
VNm . . . . . Medial Visual Network
VNop . . . . . Occipital Pole Visual Network

1 Thesis overview

1.1 Thesis motivation, aims, and structure

Newborn infants undergo a multitude of necessary painful procedures shortly after birth as part of standard clinical care (Carbajal et al., 2008). Due to the subjective nature of pain, verbal report is the gold standard for assessing the presence and intensity of the pain experience (IASP, 2019). This causes serious challenges for pain management of newborn infants during their time in hospital. It takes several years' training to acquire the clinical acumen to competently decide upon appropriate pain management strategies for infants undergoing clinically required painful procedures, such as type and titration of analgesia and the required intensity of post-procedural monitoring. Lack of verbal communication, a limited behavioural repertoire, and the brief extra-uterine time window for establishing a medical history, ensure a considerable degree of uncertainty exists in clinical decision making for effective personalised analgesic care of the newborn infant. Current clinical pain assessment relies on ambiguous behavioural and physiological responses to noxious events, such as crying and elevated heart rate, as indicators of an infant's pain experience and sensitivity (Lee and Stevens, 2013). This uncertainty and ambiguity results in pain often going under-detected and under-treated in the infant population (Roofthooft et al., 2014). It would thus be of immense value to identify novel informative indicators of a newborn infant's unique sensitivity to nociceptive input that do not rely on prior observations of pain-related responses to clinical procedures, and that could reduce the uncertainty and ambiguity during clinical decision-making processes regarding infant pain management.

An infant's brain structure and function are unique, and may be a source of infant-specific indicators of how an infant will respond to painful events. In this thesis, a multimodal magnetic resonance imaging (MRI) approach was adopted to objectively quantify and explain an infant's cerebral haemodynamic response to nociceptive input. Resting-state activity was assessed for its capacity to predict an infant's noxious-response amplitude from spontaneous noxious-free brain activity. Resting-state activity has great potential to provide novel insight into the multiple dimensions of infant pain experience and expression, as it is unique to each individual (Finn et al., 2015; Tavor et al., 2016), originates in the organ from which pain perception emerges, and acquisition does not require pain or nociception exposure for the infant. In addition to generating resting-state-based predictions, specific structural and functional neural correlates of an infant's noxious-response amplitude were identified using white matter diffusion MRI (dMRI) and resting-state functional MRI (rs-fMRI). These novel neural correlates help to improve our understanding of the neurophysiological basis of individual variability in haemodynamic noxious-response amplitude.

MRI is a flexible neuroimaging technique with which both structural and functional data can be collected. Compared to electroencephalography (EEG) and near infrared spectroscopy (NIRS), the most commonly used neuroimaging techniques for studying infant pain, fMRI provides unparalleled insight into the spatial organisation of the brain, allowing whole-brain coverage and greatest spatial resolution. However, there are several limitations to using fMRI to study infant pain, which has restricted its uptake, with only three fMRI studies on infant pain processing published to date (Williams et al., 2015; Goksan et al., 2015; Goksan et al., 2018). General limitations of fMRI include the relatively poor temporal resolution and the indirect haemodynamic nature of the fMRI blood oxygen level dependent (BOLD) signal. Limitations that are particularly prominent in infant fMRI include the high degree of head motion signal contamination as well as the lack of data analysis pipelines that are optimised for the infant population, with researchers often relying on methods developed and optimised for the adult

brain. Relevant to data analysis optimisation, some major adult-infant differences include the reduced size and degree of gyrification of the infant brain, the infant's immature haemodynamic response to neural activity, fundamental differences in brain tissue composition, and physiological and motor outputs. Inappropriate data analysis methods will likely fail to account for the heavy motion contamination, smaller brain size, and immature haemodynamic response waveform, ultimately preventing the achievement of data quality high enough to convincingly assess infant brain activity at the individual level.

In order to successfully use fMRI to understand the neurophysiological basis of individual variability in noxious-response amplitudes, development of a bespoke newborn infant fMRI data analysis pipeline optimised for the nociception experimental paradigm was required. In this thesis, an analysis pipeline for data preprocessing and haemodynamic response modelling was outlined and validated. Five key stages of infant fMRI data analysis were tested: motion and distortion correction, spatial normalisation, independent component analysis (ICA)-based denoising, spatial filtering, and haemodynamic response function (HRF) modelling.

In summary, this thesis has three central aims. First, outlining and validating an optimised infant noxious-response fMRI data analysis pipeline to achieve high subject-level data quality. Second, assessing the potential to use an infant's unique resting-state brain activity to predict the size of the cerebral haemodynamic response to nociceptive input. Third, identifying novel functional and structural neural correlates of an infant's haemodynamic response to nociceptive input to improve our understanding of the neurophysiological basis of individual variability in response amplitude.

Four chapters follow this overview chapter. Chapter 2 provides the requisite background for contextualising the research performed in later chapters. Chapters 3 and 4 are the core research chapters, both of which are written in IMRaD format — Introduction, Methods, Results, and Discussion. Chapter 3 addresses the first thesis aim, outlining and validating an optimised analysis pipeline for infant noxious-response fMRI data. Chapter 4 addresses the second and third thesis aims, assessing

the predictive capacity of infant rs-fMRI data and identifying specific structural and functional neural correlates of noxious-response amplitude. Chapter 5 provides a general summary of the key findings of this thesis, as well as an outline of limitations, suggested improvements, extensions, and future work.

1.2 Novel contributions of this thesis

The novel contributions of the work in Chapter 3 are as follows:

- First publication to use EDDY adapted for fMRI data (Baxter et al., 2019)
- First publication with FIX applied to task fMRI data in infants (Baxter et al., 2019)
- Demonstration that the dHCP preprocessing pipeline is generalisable to an independent dataset with quite different acquisitions and generalisable to non-resting state data
- First demonstration that the advanced preprocessing pipeline, such as the use of EDDY, removes noise variance without diminishing signal-of-interest
- First direct comparison of a dHCP preprocessing pipeline with an existing preprocessing pipeline optimised for adults and subsequently adapted for infants
- First direct comparison of multiple infant HRF parameterisations

The novel contributions of the work in Chapter 4 are as follows:

- First demonstration that an infant's noxious-response brain activity is tightly coupled to ongoing resting-state activity
- First demonstration that cross-infant variability in noxious-response amplitude is related to underlying white matter microstructure

1.3 Publications, presentations, and awards

1.3.1 Publications

The study outlined in Chapter 3 has been published open-access under CC-BY licence with copyright maintained (Baxter et al., 2019). While the introduction, methods, and discussion sections of Chapter 3 have been restructured, expanded, and updated since publication, the core results have not changed.

Goksan, S., **Baxter, L.**, Moultrie, F., Duff, E., Hathway, G., Hartley, C., Tracey, I., Slater, R., 2018. *The influence of the descending pain modulatory system on infant pain-related brain activity.* eLife 7. `https://doi.org/10.7554/eLife.37125`

Personal contributions:

- Data analysis: performed all fMRI preprocessing using my optimised FEAT pipeline outlined in Chapter 3
- Writing, reviewing, and editing: contributed to early manuscript drafts, helped in results interpretation, and helped in responding to reviewer feedback

Baxter, L., Fitzgibbon, S., Moultrie, F., Goksan, S., Jenkinson, M., Smith, S., Andersson, J., Duff, E., Slater, R., 2019. *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. `https://doi.org/10.1016/j.neuroimage.2018.11.006`

Personal contributions:

- Data analysis: performed all analysis required to produce outputs in all tables and figures, reported requirements from our end to generalise the dHCP pipeline to non-dHCP data and non-resting state data (e.g. requesting the temporal filter not be hard coded, requesting the inclusion of spatial filtering and grand mean scaling), aiding in the updating of the version of FIX to

include in the preprocessing pipeline, general bug checking and reporting for the dHCP pipeline

- Writing, reviewing, and editing: wrote the core manuscript, applied edits at all stages, and responded to reviewer feedback

Fitzgibbon, S.P., Harrison, S.J., Jenkinson, M., **Baxter, L.**, Robinson, E.C., Bastiani, M., Bozek, J., Karolis, V., Grande, L.C., Price, A.N., Hughes, E., Makropoulos, A., Passerat-Palmbach, J., Schuh, A., Gao, J., Farahibozorg, S.-R., O'Muircheartaigh, J., Ciarrusta, J., O'Keeffe, C., Brandon, J., Arichi, T., Rueckert, D., Hajnal, J.V., Edwards, A.D., Smith, S.M., Duff, E., Andersson, J., 2019. *The developing Human Connectome Project (dHCP) automated resting-state functional processing framework for newborn infants.* bioRxiv 766030. `https://doi.org/10.1101/766030`

Personal contributions:

- Data analysis: used iterations of the pipeline in independent dataset as part of pipeline development and testing of usability and generalisability, reported requirements from our end to generalise the dHCP pipeline to non-dHCP data and non-resting state data (e.g. requesting the temporal filter not be hard coded, requesting the inclusion of spatial filtering and grand mean scaling), aiding in the updating of the version of FIX to include in the preprocessing pipeline, acted as one of the two secondary FIX ICA component raters to allow calculation of inter-rater reliability, general bug checking and reporting for the dHCP pipeline

- Writing, reviewing, and editing: helped review early and late versions of the manuscript

1.3.2 Presentations

Oral presentations

Predicting infants' pain sensitivity from spontaneous brain activity
Medical Sciences DPhil Day, Oxford, UK. 2018

Imaging Pain in the Developing Brain
Wellcome Centre for Integrative Neuroimaging Research Day, Oxford, UK. 2018

Variability in the Infant's Pain Response
11th International Conference on Brain Monitoring and Neuroprotection in the Newborn, Florida, USA. 2018

Individual variability in the infant pain response: Using MRI to understand underlying mechanisms
12th International Symposium on Paediatric Pain, Basel, Switzerland. 2019

Poster presentations

dHCP fMRI analysis pipeline enhances the detection of nociceptive brain activity in neonates
Organization for Human Brain Mapping Annual Meeting, Singapore. 2018

Functional magnetic resonance imaging reveals dependency between spontaneous brain activity and noxious stimulus response at the level of the individual infant
12th International Symposium on Paediatric Pain, Basel, Switzerland. 2019

1.3.3 Awards

Medical Research Council: £105,676
Medical Sciences Graduate School Scholarship (2016)

Guarantors of Brain, Travel Grant: £684
To attend the Organization for Human Brain Mapping Annual Meeting, Singapore (2018)

Reckitt Benckiser, Educational Grant: £4,874

Academic networking innovation grant (2019)

2 Background

2.1 The concept of pain

2.1.1 Defining pain and nociception

Pain is an unpleasant sensory experience that acts as a warning system to protect us from bodily harm. In order to be able to detect, measure, and alleviate pain, a clear definition is needed. The International Association for the Study of Pain (IASP) definition of pain, first put forward by Merskey and colleagues in 1979 (Merskey et al., 1979), has been widely adopted by the pain community. Pain is defined as "an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage" (IASP, 2019). This definition highlights the subjective multidimensional nature of pain, and addresses both the sensory and emotional dimensions. The sensory dimension encompasses localisation (where is the pain?), intensity (how intense is the pain?), and modality (is the pain burning, pricking, etc.?). And the emotional dimension encompasses the emotional valence (pain is unpleasant) and intensity (how unpleasant is the pain?). There are indeed other important dimensions to pain. Cognitive aspects include pain's attention-demanding nature, motor aspects include defensive and avoidance behaviours, and autonomic and endocrine aspects include cardiovascular, respiratory, and hormonal responses. While the IASP definition of pain is not the only one available, it is a useful and widely adopted operationalisation of pain that has stood the test of time (Treede, 2018). It will thus be adopted throughout this thesis to ensure a clear and consistent terminology.

Pain is typically classified into three categories: nociceptive, inflammatory, and pathological (Woolf, 2010). Nociceptive pain is the sensing of harmful, potential or actual tissue-damaging stimuli — caused by activation of high-threshold receptors. Inflammatory pain is pain due to heightened sensitivity and tenderness in bodily regions that are undergoing healing in response to tissue damage or infection — caused by activation of the immune system. Both nociceptive and inflammatory pain serve beneficial and protective functions: nociceptive pain motivates an individual to get out of immediate harm's way, and inflammatory pain discourages the individual from exposing damaged tissue to further harm until healing has occurred. Pathological pain is maladaptive pain due to abnormal functioning of the pain system. Causes of chronic pathological pain vary widely: chronic inflammation (e.g. rheumatoid arthritis), nervous system disease (e.g. post-herpetic neuralgia), or idiopathic pain (e.g. fibromyalgia). Pain can also be classified based on the organ of origin. Somatic pain originates from superficial organs, such as the skin i.e. cutaneous pain. And visceral pain originates from deep organs, such as the intestines.

Pain is a broad and complex topic. There are beneficial and harmful types of pain, that can work over acute or chronic time-scales, and in every case, each of the multiple dimensions of pain can play roles of varying degrees. As will be described in later chapters, the type of pain relevant to the work within this thesis is the pain experienced following acute medically-required nociceptive procedures, such as routine blood testing that is frequently performed in newborn infants using a heel lance, where a minor skin-breaking incision is made on the heel (Courtois et al., 2016). This type of pain is classified as acute somatic (cutaneous) nociceptive pain.

As MRI is the data-type used throughout this thesis, it is necessary to note that newborn infant pain is challenging to directly study in the MRI environment. There are several reasons for this, but the primary reason is ethics-based. In most research investigating infant pain, and that performed within my research group (Paediatric Neuroimaging Group), an infant's pain responses or pain-related brain activity are usually recorded while the infant is undergoing a clinically-required painful event e.g. blood test, vaccination, intubation, etc. Unlike adult studies, pain is not inflicted

purely for the purposes of research. Practical considerations and safety concerns make it unacceptable to perform painful clinical procedures in the MRI environment. Instead, very mild experimental stimuli are used in the MRI setting that do not distress the infant or cause clinical concern in any way. To mimic a heel lance blood test in the MRI scanner, a sharp-touch stimulus is applied to the infant's foot. The sharp nature of the stimulus activates the cutaneous receptors and subsequent pathways and brain regions involved in processing painful events, but at an intensity presumed to be below the pain threshold due to the lack of behavioural and physiological distress responses. These sharp-touch experimental stimuli facilitate probing the infant pain system without upsetting or distressing the newborn infant.

It is thus important to introduce the distinction between pain and nociception. IASP defines nociception as "the neural process of encoding noxious stimuli", where a noxious stimulus is "a stimulus that is damaging or threatens damage to normal tissues" (IASP, 2019). The sharp-touch stimulus used throughout this thesis does not cause tissue damage, yet its sharpness can be considered threatening of tissue damage. At the intensities used in experimental research, it activates noxious skin receptors that, if activated with greater intensity, would result in nociceptive pain. As is clear from these definitions, all nociceptive pain must involve nociception, but nociception does not always lead to pain.

2.1.2 Conceptual frameworks

In addition to a clear definition of pain, it is valuable to embed the definition in a conceptual framework. Conceptual frameworks allow the formulation of testable hypotheses, and facilitate the integration of theories of pain with theories of other perceptions and neural functions. Historically, there have been two major competing views on pain processing: the labelled lines theory and the pattern theory (Moayedi and Davis, 2013). Labelled lines theory (or specificity theory) posits that pain-specific receptors and pathways exist, transmitting noxious input along dedicated pathways to specific pain centres in the brain. In contrast, pattern theory posits that no pain-specific labelled lines exist, and it is the specific pattern of incoming

sensory information that encodes the sensation of pain. These theories remained in competition until Ronald Melzack and Patrick Wall published their influential gate control theory of pain in 1965 (Melzack and Wall, 1965). This theory embraced patterns over labelled lines, partly due to the influence of the recently discovered wide-dynamic-range neurons (Mendell and Wall, 1965), which respond to both noxious and non-noxious inputs, and partly due to failings of labelled lines theory (Wall, 1996). While the gate theory was tremendously fruitful conceptually, it is now clear that most of the finer details of the theory were incorrect (Mendell, 2014). There are specific mechanistic issues with the hypothesised gate theory spinal cord circuitry, and nociceptive-specific cells, in addition to wide-dynamic-range cells, exist at all points of pain pathways from the periphery to the brain (Price, Greenspan, and Dubner, 2003). It is now largely agreed that the peripheral sensory input aspect of the pain system (the primary afferent neuron) operates in a labelled line fashion. However, in both the spinal cord and brain, pain pathways feature both noxious-specific and non-noxious-specific neurons. The relative contribution and importance of patterns and labelled lines in processing pain in the central nervous system remains an interesting and open question (Basbaum, 2011).

More recently, pain has been incorporated into broader theoretical frameworks, such as the Bayesian learning model of pain (Clark, 2013; Tabor et al., 2017; Tabor and Burr, 2019). Bayesian learning models conceptualise pain as a perception, and thus as a sensory signal. Under this framework, pain is viewed as the brain's optimal inference of the nociceptive stimulus (Tabor et al., 2017; Tabor and Burr, 2019; Seymour, 2019). Pain is modulated by factors such as emotions and expectation, as these existing neural contextual factors (the Bayesian prior distribution) bias the noisy sensory input (the Bayesian likelihood function). Pain-related behaviours are then viewed as actions that minimise future unexpected pain (Tabor and Burr, 2019). Conceptual frameworks of pain other than the Bayesian learning model approach exist and will be discussed below. In brief, the conceptual frameworks most commonly adopted in the infant pain literature (biopsychosocial and biopsychomotor models) do not focus on or attempt to explain neural pain mechanisms, and so,

are of limited value to the work in this thesis. The Bayesian learning model of pain is adopted for three reasons. First, it is consistent with the IASP definition of pain as a "sensory and emotional experience". Second, it is an elegant model that offers a general computational approach encompassing all sensory perception modalities (Clark, 2013). Its general framework naturally incorporates cognitive and emotional modulatory influences on pain perception, and also provides a framework for understanding a person's pain-related motor outputs and behaviours. And third, a pain-specific literature is beginning to emerge demonstrating the value of conceptualising pain as a form of Bayesian statistical inference (Büchel et al., 2014; Anchisi and Zanon, 2015; Geuter et al., 2017; Fazeli and Büchel, 2018; Tabor et al., 2017; Tabor and Burr, 2019; Ongaro and Kaptchuk, 2019).

It is important to note that the development of conceptual frameworks for pain is an active area of research with ongoing debate. To appreciate what an alternative conceptualisation of pain may look like, Ben Seymour's recently detailed reinforcement learning model of pain (Seymour, 2019) will be briefly described. In contrast to the Bayesian learning model of pain, the reinforcement learning model conceptualises pain as a teaching and control signal that "drives the learning of values" (Seymour, 2019). Under this framework, pain is viewed as an optimal control signal to minimize current and future nociceptive stimuli. Pain is no longer viewed as a sensory signal and thus should not be viewed as 'modulated' by emotional or cognitive factors, as the concept of modulation assumes pain is reflecting a nociceptive sensory signal. Instead pain is simply tuned by factors in a precise manner to direct harm-minimizing behaviours. Whether the conscious perception of pain reflects "an optimal inference of a real or presumed nociceptive stimulus, or an optimal control signal to minimize current and future nociceptive stimuli" (Seymour, 2019) is an interesting contemporary debate.

Due to the difficulty in studying pain in infants, it is not surprising that definitions and conceptualisations of pain used in adult research are often not translated to infants. An early misconception that the IASP definition of pain is "strictly defined within the capabilities of adult interpretation" (Anand and Craig, 1996) was an

early and influential catalyst that shifted conceptualisations of infant pain from a conscious experience to a measurable behavioural or physiological response. In infant pain research, popular conceptual frameworks of pain include biopsychosocial and biopsychomotor models. Regarding biopsychosocial models, Craig's Social Communication Model (SCM) (Craig, 2015) is the most developed. This model has been adapted specifically for infants by Rebecca Pillai Riddell to form the Development of Infant Acute Pain Responding (DIAPR) model (Pillai Riddell et al., 2013). Similar to the Bayesian and Reinforcement learning models, these biopsychosocial models take a holistic view of pain, attempting to encompass pain's multidimensional nature. However, biopsychosocial models focus predominantly on infant behaviour and the infant's social environment, and do not attempt to provide mechanistic insight into infant neural pain processing – "Rather than pursuing the infinite challenge of operationalizing the non-verbal infant's subjective experience of the noxious stimulus, the [DIAPR] model focuses on overtly measurable pain behaviours" (Pillai Riddell et al., 2013). Due to the shift in focus from "pain experience" to "pain expression" and the lack a neural mechanistic basis, adopting the biopsychosocial conceptual framework would be inappropriate for the work in this thesis. However, biopsychosocial models are not in conflict with the Bayesian learning model of pain (unlike the reinforcement learning model in their current state), and indeed can straight-forwardly be accounted for by the broader Bayesian model. For example, biopsychosocial models highlight the importance of the social context, which is readily accounted for by the Bayesian approach (Krahé et al., 2013).

Another popular framework in infant pain research is the biopsychomotor model (Sullivan, 2008). Similar to the biopsychosocial models, this biopsychomotor model emphasises the importance of measurable pain behaviours. However, unlike the biopsychosocial models, there is an explicit attempt to redefine pain to incorporate overt motor behaviours. While behavioural output is central to pain, and is a major feature of the Bayesian and reinforcement learning models, it is not clear what the biopsychomotor definition of pain is, and what adopting it would entail, other than directing increased focus onto subject behaviour. And like the

biopsychosocial models, the biopsychomotor model does not attempt to provide a neural mechanistic account.

2.1.3 Summary

In this section, definitions of pain and nociception were outlined, as well as important conceptual frameworks in which pain is commonly embedded in both adult and infant pain research. Acute nociceptive somatic pain was outlined, as this subcategory of pain is the focus of this thesis. The important distinction between pain and nociception was introduced, as it is relevant to infant pain studies in the MRI setting, where clinically-required painful procedures cannot typically be performed. The Bayesian learning model was introduced to provide a brain-focused mechanism-based conceptual framework in which the IASP definition of pain can be embedded. Explicitly addressing these fundamental topics should ensure conceptual consistency and clarity throughout this thesis, and these concepts will be of central importance when critically evaluating currently available infant pain assessment measures (Section 2.3.2).

2.2 The Pain System

2.2.1 The adult pain system

Acute nociceptive somatic pain is initially transduced by first-order neurons in the periphery. This transduction of mechanical nociceptive stimulation is termed mechanotransduction, the mechanisms of which are still largely unknown (Lumpkin and Caterina, 2007). A subset of dorsal root ganglion cells is specifically sensitive to high-threshold mechanical input (i.e. mechanical nociceptive input) (Burgess and Perl, 1967; Perl, 1968), and are called nociceptors. Nociceptors encoding noxious touch are predominantly high-threshold Aδ-fibre and C-fibre mechanoreceptors. However, not all Aδ- and C-fibres are nociceptors. In hairy skin, low-threshold mechanoreceptor for both Aδ-fibres (Djouhri, 2016) and C-fibres (CT afferents) (McGlone, Wessberg, and Olausson, 2014) mediate gentle and pleasant touch. The high-threshold Aδ-fibre and C-fibre mechanoreceptors contain subpopulations

of both modality-specific nociceptors responding to mechanical stimulation alone (Cavanaugh et al., 2009; Scherrer et al., 2009), and polymodal nociceptors responding to both mechanical and thermal stimulation (Perl, 1996). Both Aδ-fibre and C-fibre nociceptors encode high-intensity mechanical stimuli (Slugg, Campbell, and Meyer, 2004), but are differentially tuned to the type of mechanical nociception encoded, with Aδ-fibres encoding sharp pricking pain and C-fibres encoding dull aching pain (Konietzny et al., 1981; Ochoa and Torebjörk, 1989). Both types of nociceptors project predominantly to the superficial layers of the spinal cord dorsal horn, laminae 1 and 2, but there are also significant projections to lamina 5 (Light and Perl, 1979a; Light and Perl, 1979b; Sugiura, Lee, and Perl, 1986; Sugiura, Terui, and Hosoya, 1989).

In the spinal cord dorsal horn, two major types of second-order nociception projection neurons exist: nociceptive-specific neurons (Christensen and Perl, 1970) and wide-dynamic-range neurons (Mendell, 1966). Both types of nociceptive neurons are found at all junctions of central pain pathways from the spinal cord to the cortex (Price, Greenspan, and Dubner, 2003). In addition to projection neurons, the dorsal horn also includes a range of modulatory neurons. Local inhibitory modulation in the dorsal horn is provided by GABA-ergic and enkephalinergic interneurons, both of which provide direct inhibition of the spinothalamic tract (Ruda, Coffield, and Dubner, 1984; Carlton et al., 1992; François et al., 2017). The brainstem also provides abundant external modulatory input to the dorsal horn. The rostral ventral medial medulla (RVM), a major output of the descending pain modulatory system, provides both inhibitory anti-nociceptive input from RVM off-cells and facilitatory pro-nociceptive input from RVM on-cells (Heinricher et al., 2009). Additionally, brainstem serotonergic and noradrenergic inputs provide anti-nociceptive modulation (Yoshimura and Furue, 2006), both indirectly via dorsal horn interneurons and directly to spinothalamic and other ascending tracts (Hylden et al., 1986; Westlund et al., 1991; Miller and Salvatierra, 1998). The serotonergic input comes from several regions including nucleus raphe magnus, while the noradrenergic input come from distinct regions including the locus coeruleus (Kwiat and Basbaum, 1992).

These pain modulatory regions are important sites for inducing analgesia: electrical stimulation of brainstem regions such as RVM and periaqueductal grey (PAG) can result in stimulation-induced analgesia (Hosobuchi, Adams, and Linchitz, 1977), and positive expectations can result in placebo analgesia also acting through activation of a network of brain regions involving the RVM and PAG (Eippert et al., 2009).

There is a multitude of routes for spinal cord nociceptive outputs. Nocifensive limb withdrawal reflexes occur when nociceptive input from high-threshold Aδ-fibres indirectly synapse, via spinal interneurons, on α-motoneuron (Sandrini et al., 2005). This rapid polysynaptic spinal withdrawal reflex provides a protective mechanism for the individual. There are also modulatory feedback loops via the brainstem. The spinoreticular tract includes projections to brainstem reticular formation regions including the RVM (Haber, Moore, and Willis, 1982; Mehler, Feferman, and Nauta, 1960), and the spinomesencephalic tract includes projections to the PAG (Yezierski, 1988; Mantyh, 1982). Both of these pathways are central to the influence of incoming nociceptive input on the output of the descending pain modulatory system. Beyond brainstem modulatory loops, the spinohypothalamic tract projects to the hypothalamus (Zhang et al., 1999), and the spinoparabrachial tract projects to parabrachial nuclei (Saper, 1995), which then project nociceptive input to both the hypothalamus (Bester et al., 1995) and amygdala (Bernard and Besson, 1990). These two tracts thus provide a route for nociceptive input to reach regions of the cerebrum involved in motivational and emotional processing as well as autonomic outputs.

Finally, the spinothalamic tract is the principal ascending spinal tract that carries nociceptive inputs to cortical regions, via several nuclei of the thalamus. The spinothalamic tract leaves predominantly from laminae 1 and 5 (Price, Greenspan, and Dubner, 2003) of the spinal cord dorsal horn, with 90% of the spinothalamic tract projecting to contralateral thalamus (Apkarian and Hodge, 1989). Spinothalamic inputs project predominantly to five subregions of the thalamus. The ventral posterior complex relays to primary somatosensory, secondary somatosensory, and insular cortices (Kenshalo et al., 1980; Kenshalo and Isensee, 1983; Stevens, London,

and Apkarian, 1993). The posterior complex relays to secondary somatosensory and insular cortices (Burton and Jones, 1976). The ventral lateral nucleus relays to primary motor cortex (Stepniewska et al., 2003). The intralaminar complex relays to the striatum and several cortical regions, including prefrontal and cingulate cortices (Jones, 2007). And finally, the medial dorsal nucleus relays to middle cingulate cortex (Dum, Levinthal, and Strick, 2009). In general, approximately 40% of spinothalamic inputs go to insular cortex via the posterior complex (posterior-suprageniculate nucleus), 30% go to secondary somatosensory cortex via the ventral posterior complex (ventral posterior inferior nucleus), 25% go to middle cingulate cortex via the mediodorsal nucleus, and less than 5% go to primary somatosensory cortex via the ventral posterior complex (ventral posterior lateral nucleus) (Dum, Levinthal, and Strick, 2009). Whole-brain human functional imaging studies corroborate the importance of the multitude of brain regions listed here in the processing of nociceptive input (Jones et al., 1991; Treede et al., 1999; Apkarian et al., 2005). In 2005, Apkarian and colleagues (Apkarian et al., 2005) conducted a thorough meta-analysis of 60 human neuroimaging studies, identifying the thalamus, primary and secondary somatosensory, insular, cingulate, and prefrontal cortices as key components of the human pain system.

2.2.2 The infant pain system

Note: Throughout this thesis, terminology used to describe infant age will be that outlined by the American Academy of Paediatrics (Engle and American Academy of Pediatrics Committee on Fetus and Newborn, 2004). Three terms will be commonly used throughout, and are explicitly defined here for clarity:

- Gestational age: time elapsed between the first day of the last menstrual period and the day of delivery
- Postnatal (or Chronological) age: time elapsed since birth
- Postmenstrual age: gestational age + postnatal age

Prior to term age, both Aδ- and C-fibres nociceptors have reached the spinal cord and are functional (Fitzgerald and Swett, 1983). In the dorsal horn of the spinal cord, both nociceptive-specific and wide-dynamic-range neurons project to the thalamus via the spinothalamic tract (Fitzgerald and Jennings, 1999; Davidson, Truong, and Giesler, 2010). Nociceptive signals in the dorsal horn are being modulated by both local interneurons and descending projection neurons from the brainstem. While the GABA neurotransmitter plays an excitatory role in earlier development (Ben-Ari, 2014), by term age, the local GABAergic interneurons provide inhibitory input to the spinothalamic projections (Bremner, Fitzgerald, and Baccei, 2006; Hathway et al., 2006). Unlike the adult, the descending modulatory input from brainstem regions appears to be solely facilitatory (Hathway et al., 2006; Hathway et al., 2009), with electrical stimulation of the PAG failing to result in stimulation-induced analgesia (vanPraag and Frenk, 1991). In addition to brainstem facilitatory input, the dorsal horn has increased excitatory peripheral input from Aβ-fibres (Jennings and Fitzgerald, 1996), resulting in the infant dorsal horn projection neurons being more excitable than adults. In addition to direct examination of the nervous system, the presence of functioning neural substrates can be inferred from typical responses to noxious stimuli. Term age infants display clear limb withdrawal reflexes, facial expression, crying, cardiovascular, respiratory, and hormonal responses to tissue-damaging stimuli, such as clinically required heel lance blood tests. These responses can only occur if the neural pathways underpinning nocifensive behavioural, autonomic, and endocrine outputs are functional. These outputs thus demonstrate the functional status of spinal motor outputs, brainstem motor and autonomic outputs, as well as hypothalamic endocrine outputs. However, for perception of the pain experience to be possible, the cerebral circuits underpinning the sensory and emotional dimensions of pain must also be functional.

Rodent studies as well as human histological studies confirm that thalamocortical and corticocortical connections have developed by term age (Kostović and Jovanov-Milosević, 2006). The functionality of these connections has been demonstrated using electroencephalographic (EEG) and imaging techniques in newborn infants.

Resting-state networks typically found in adults can be observed in the patterns of infant resting-state activity (Doria et al., 2010; Fransson et al., 2011; Fitzgibbon et al., 2019), demonstrating functional long-range intra- and inter-hemispheric connectivity. While the functional presence of the somatosensory system can be inferred from the presence of somatosensory resting-state networks, this has also been directly established in newborns using both EEG (Khazipov et al., 2004; Milh et al., 2007) and fMRI (Arichi et al., 2010) to detect somatosensory-evoked neural activity. More recently, there is evidence for the establishment of hierarchical organisation of the term-age infant's somatosensory system, at least in the earliest stages of stimulus processing (Whitehead et al., 2019). Comparing the infant cerebral responses of high-threshold nociceptive and low-threshold non-nociceptive cutaneous mechanical input, there is now clear evidence that nociceptive-specific information reaches the newborn infant's brain, detectable using both neural and haemodynamic methods (Slater et al., 2006; Hartley et al., 2017). Using EEG, the infant brain has been shown to differentially encode noxious and non-noxious touch from at least 35 weeks postmenstrual age (Fabrizi et al., 2011).

A landmark functional MRI (fMRI) study in 2015 by Goksan and colleagues (Goksan et al., 2015) compared regions of cerebral activation in response to mild noxious stimulation with a pinprick stimulator between adults and newborn infants. This study highlights several important aspects of infant nociception processing. First, the overall pattern of functional activation closely resembles that of adults. This establishes that the functional segregation division-of-labour organisation principle of the brain is already established in newborn infant nociception processing. Second, the list of similarly activated brain regions between adults and infants include both lower and higher levels of the cortical hierarchy. In infants, nociceptive inputs are processed in lower-level regions including the insula and primary and secondary somatosensory cortices, as well as higher-level behavioural output regions including the striatum and middle cingulate cortex, suggesting newborn infants can process the sensory-discriminative and nocifensive-behavioural dimensions of pain. Third, the major observed difference between adult and infant activity patterns was that infants

did not exhibit functional activation of either the amygdala or the orbitofrontal cortex in response to the 128 mN pinprick stimulus, while adults did in response to the 512 mN stimulus. The amygdala and orbitofrontal cortex are centrally involved in determining reward value and emotional valence of sensory input including pain (Rolls et al., 2003). Considering the IASP definition of pain requires an emotional response, lack of activity in these regions tentatively suggests the infants, unlike the adults, did not assign a strong emotional or reward value to the stimulus. While this lack of activity due to failing to reject the null hypothesis may simply indicate a lack of statistical power, the tentative non-activation hypothesis is consistent with the view of the 128 mN pinprick stimulus as mildly nociceptive, likely near or below the infants' pain threshold, not behaviourally or physiologically distressing, and not a cause for any clinical concern. Using EEG, the same experimental pinprick stimulus has been shown to evoke a noxious-specific event related potential (ERP) at a lower amplitude to that seen in response to the heel lance (Hartley et al., 2015; Hartley et al., 2017). Together, these EEG and fMRI studies help validate the use of the sharp-touch 128 mN pinprick stimulus to probe the infant nociceptive pain system in an ethically, clinically, and scientifically acceptable manner.

2.2.3 Summary

In this section, the core structural and functional neural components underpinning acute nociceptive somatic pain were outlined. Topics such as cell types, anatomical pathways, brain structures, and functional segregation in the brain were described for the healthy adult. Considering the population under study in this thesis is healthy newborn infants at peri-term age, the developmental status of the nociceptive somatic pain system for this population was outlined. This overview covered the commonalities and differences between the adult and infant systems, concluding that an immature system is structurally and functionally established by the peri-term period. EEG studies have demonstrated that both tissue-damaging clinical procedures as well as non-tissue-damaging sharp-touch experimental stimuli elicit noxious-specific responses at the level of the cerebral cortex, and that the infant

brain differentially encodes painful from non-painful touch from at least the age of 35 weeks PMA. None of the infants studied in this thesis were under this age. Using the same 128 mN pinprick stimulus that is used throughout this thesis, a previous fMRI study established the feasibility of newborn infant fMRI studies using a nociception paradigm. The same study also observed that the newborn infant brain response to nociceptive input involves a range of spatially distinct regions demonstrating functional segregation in the newborn infant brain, and also revealed that the infant pattern of evoked brain activity is similar to that of adults. The ability to study infant nociception using fMRI as well as the functional segregation observed in the newborn infant brain strongly motivate the need for further fMRI studies on the infant pain system to capitalise on the unique strengths of the fMRI modality i.e. whole-brain high spatial resolution imaging of brain function. The topic of MRI is introduced more systematically below (Section 2.4).

2.3 Infant Clinical Pain Assessment

2.3.1 Classifying infant clinical pain assessment tools

In Section 2.1, the IASP definition and Bayesian inference model of pain were introduced to provide clarity on the concept-of-interest i.e. pain. Due to the subjective and multidimensional nature of pain, assessment can be challenging in both adults and infants. "Many people report pain in the absence of tissue damage or any likely pathophysiological cause [...] If they regard their experience as pain, and if they report it in the same ways as pain caused by tissue damage, it should be accepted as pain" (IASP, 2019). This firmly establishes the verbal report as the gold standard for pain assessment. In adults, gold standard assessment of pain can be obtained, for example, using pain rating scales (e.g. visual analog scale) or pain questionnaires (e.g. McGill Pain Questionnaire). However, gold standard does not mean perfect measure, and several issues with the verbal report exist. First, verbal report involves introspection, the psychological demands of which will alter the pain experience to an unknown degree. Second, pain rating scales typically probe a single dimension of pain, such as the intensity of the sensory experience (e.g. the

visual analog scale), which does not take into account the intensity of the emotional experience. However, questionnaires can probe multiple pain dimensions. Third, for both pain rating scales and questionnaires, standardising intensities across subjects can be difficult e.g. "rate your pain from none to worst pain imaginable" depends on the person's imagination. There are several other flaws in verbal reports such as the subject being confused (a major issue in elderly subjects with declining cognitive capacity), or subjects lying (common in children who may exaggerate to get attention or deny they are in pain to avoid an injection, etc.). This list of verbal report limitations should not be considered an exhaustive list. In adults, there have been, and continue to be, many attempts to find alternative pain measures to use instead of the verbal report for use in challenging populations, such as non-verbal adults and infants. While great strides have been made towards developing alternatives (Cowen et al., 2015; Tracey, Woolf, and Andrews, 2019), there have been many observed disassociations between these alternatives and the verbal report (reviewed in Gracely, 2013), indicating that, currently, the presence or absence of pain perception can only be truly confirmed through the imperfect verbal report.

There are also situations in pain research and treatment where a person's current "pain experience" is not necessarily the endpoint-of-interest, and in these situations, the verbal report may thus not be the gold standard measure. For example, the development of the biopsychomotor model (Section 2.1.2) was motivated by the imperfect correlation between pain experience and pain-related behaviour: "treatments targeting pain sensation might not always yield the best outcomes" and "the behavioral dimensions of pain (eg, disability) may persist despite reductions in pain, and that disability can be reduced in the absence of reduction in pain" (Sullivan, 2008). Similarly, clinical trial drug development may wish to focus on long-term pain reduction and the reduction in chronic-pain-related brain activity, which may be dissociable from a person's current pain experience (Wanigasekera et al., 2018). In situations such as these, the endpoint-of-interest and thus the gold standard assessment tool might be an objectively quantifiable behavioural or neural feature, rather than the subjective verbal report of pain perception.

Much of infant pain research either explicitly or implicitly attempts to identify and reduce an infant's pain experience and suffering, making the pain experience the primary target and the verbal report the desired gold standard. Unfortunately, it simply is not possible to obtain an unambiguous measure of pain in newborns. All measures of infant pain are indirect, relying predominantly on pain-related responses such as behavioural and physiological changes. This introduces immense challenges into infant pain assessment. Before detailing the issues with infant pain assessment, a brief outline of a typical approach is provided. Currently, over 30 pain scales have been develop to assess pain in infants (Lee and Stevens, 2013). As they all suffer from similar issues, the Premature Infant Pain Profile (PIPP) (Stevens et al., 1996) including its revised version (PIPP-R) (Stevens et al., 2014) with minor updates, is selected as an exemplar, as this scale attempts to measure acute somatic nociceptive pain, is used regularly in the Paediatric Neuroimaging Group, is adopted and well-regarded internationally, and has been well validated in comparison to other infant pain scales (Stevens et al., 2010). The PIPP(-R) scale is multidimensional, scoring behaviours (brow bulge, eye squeeze, and nasolabial furrow), physiology (heart rate and oxygen saturation), and context (gestational age and behavioural state).

Considering the introduction of definitions and conceptual models of pain improved clarity and consistency of terminology, it is worth doing the same for pain assessment approaches. To achieve this, the Food and Drug Administration and National Institute of Health (FDA-NIH) Biomarkers, EndpointS, and other Tools (BEST) resource is adopted (FDA-NIH Biomarker Working Group, 2016). The BEST resource includes a glossary of important definitions used in translational science and medical product development, the purpose of which is to improve clarity in communication and remove ambiguities and inconsistencies to help improve evaluation and interpretation of scientific ideas and evidence. The BEST resource outlines two major categories of measurement: clinical outcome assessments and biomarkers. A clinical outcome is "An outcome that describes or reflects how an individual feels, functions or survives". Clinical outcomes must undergo assessment: "The interpretation or the evaluation of the measurement". "Assessment of a

clinical outcome can be made through report by a clinician, a patient, a non-clinician observer or through a performance-based assessment". There are thus four categories of clinical outcome assessments: clinician-reported outcome, observer-reported outcome, patient-reported outcome, and performance outcome. In contrast to clinical outcome assessments, a biomarker "is not an assessment of how an individual feels, functions, or survives". And so, the second category of measurement is the biomarker: "A defined characteristic that is measured as an indicator of normal biological processes, pathogenic processes, or responses to an exposure or intervention, including therapeutic interventions". In total, there are seven categories of biomarkers under the BEST resource terminology, but they will not be listed here. In summary, there are two major categories of outcomes — the clinical outcome assessments and the biomarkers.

Applying the BEST framework to adult and infant pain assessment approaches, the adult verbal report is classified as a patient-reported outcome (PRO):

> A type of clinical outcome assessment. A measurement based on a report that comes directly from the patient (i.e., study subject) about the status of a patient's health condition without amendment or interpretation of the patient's response by a clinician or anyone else. A PRO can be measured by self-report or by interview provided that the interviewer records only the patient's response. Symptoms or other unobservable concepts known only to the patient can only be measured by PRO measures (FDA-NIH Biomarker Working Group, 2016).

In contrast, the infant PIPP(-R) scale (and other clinical assessment tools for infant pain) is classified as an observer-reported outcome (ObsRO):

> A type of clinical outcome assessment. A measurement based on a report of observable signs, events or behaviors related to a patient's health condition by someone other than the patient or a health professional. Generally, ObsROs are reported by a parent, caregiver, or someone who observes the patient in daily life and are particularly useful for patients who cannot report for themselves (e.g., infants or individuals who are cognitively impaired) (FDA-NIH Biomarker Working Group, 2016).

2.3.2 Issues with infant clinical pain assessment tools

One of the major issues with all current infant clinical pain assessment tools is an under-appreciation for the lack of test validity. Due to variations in definitions

and frameworks for concepts of validity, the COSMIN Taxonomy of Measurement Properties (Mokkink et al., 2010; COSMIN, 2019) will be adopted in order to avoid ambiguity and any lack of clarity. The COSMIN initiative provides international consensus on terminology and definitions for measurement properties with the view to improving clarity in communication and reducing confusion and ambiguity when assessing measurement property concepts. It provides clear definitions regarding types of validity as well as a series of steps required to achieve validity. Under this framework, three categories of validity are outlined: content validity, criterion validity, and construct validity.

> Content validity: The degree to which the content of a measurement instrument is an adequate reflection of the construct to be measured.
>
> Criterion validity: The degree to which the scores of a measurement instrument are an adequate reflection of a gold standard.
>
> Construct validity: The degree to which the scores of a measurement instrument are consistent with hypotheses, e.g. with regard to internal relationships, relationships with scores of other instruments or differences between relevant groups.

The previous sections' discussion on the definition and conceptual frameworks for pain have established that the construct (concept-of-interest) is a subjective, multidimensional perception, best assessed through verbal report. Given the nature of infant clinical pain scales and the pre-verbal nature of infants, the establishment of content and criterion validity for any infant pain assessment tool cannot be achieved. However, having a clear definition of pain and an understanding of the conceptual framework in which pain fits will allow the development of hypotheses that can be tested in infants. Thus, construct validity can indeed be established for infant pain scales.

Another strength of the COSMIN framework is the clear outline of steps to establish validity (deVet et al., 2011):

1. Establish a good definition of the construct, preferably embedded in a conceptual model.

2. Establish content validity. Considered to be the most important measurement property, it is the start of the validation process.

3a. Establish criterion validity. If a gold standard exists, this is a powerful form of validity.

3b. If there is no gold standard, establish construct validity.

It is commonly acknowledged that no gold standard for pain assessment exists in infants, so claims of criterion validity (step 3a) are generally not made. However, issues regarding a lack of clarity when defining infant pain and conceptual models of infant pain (step 1) has resulted in claims of content validity (step 2) being asserted for the PIPP(-R) scale (Stevens et al., 1996; Stevens et al., 2010; Stevens et al., 2014) as well as many other infant pain scales (Lee and Stevens, 2013). This conceptual ambiguity is reflected in the clinical pain assessment tools. The measurement items constituting the PIPP(-R) scale and others rely on ambiguous cues such as facial grimacing and heart rate effects, which infants exhibit under many distressing situations other than pain.

The lack of content validity in infant clinical pain scales causes further confusion around the distinction between the multidimensional nature of pain and the multidimensional nature of infant pain scales. Pain includes sensory, emotional, cognitive, motor dimensions and more. And infant pain scales such as PIPP(-R) are designed to be multidimensional to increase sensitivity to multiple pain aspects. Unfortunately, the "pain expression" dimensions captured by infant pain scales are non-specific pain-related responses that do not guarantee the presence of a pain experience. The unidimensional or multidimensional nature of infant pain scales that rely on these pain expressions provide ambiguous non-specific insight into the sensory, emotional, or cognitive dimensions of pain, with variability in pain scores potentially being due to differences in perceived stimulus intensity, emotion intensity, salience, and so on. Thus, all current infant pain scales (unidimensional and multidimensional alike) provide an ambiguous unidimensional estimate of a potential pain experience.

Another issue relates to the multiplicity of pathways from nociceptive input to response outputs, as detailed in Section 2.2.2. The spinothalamic, spinoparabrachial, spinohypothalamic, and spinomesencephalic pathways allow for the presence of behavioural responses such as limb withdrawals and facial expression, autonomic responses such as changes to respiration and the cardiovasculature, and endocrine responses such as stress hormone release, without thalamocortical circuit involvement, which is required for pain experience. This issue highlights an important point about content validity in general. As discussed in deVet et al., 2011 and originally pointed out in Borsboom, Mellenbergh, and vanHeerden, 2004, "a test is valid for measuring a construct if and only if (a) the construct exists, and (b) variations in the construct causally produce variations in measurement outcomes. [...] tables of correlations between test scores and other measures provide only circumstantial evidence for validity" (deVet et al., 2011). Without the guarantee that items in infant pain scales are causally downstream of the pain experience (the way a verbal report of pain is), further calls into question the ability of these test items to establish content validity.

Due to the lack of content and criterion validity, there is a need to invest in developing and improving the construct validity of infant clinical pain scales. This requires the development and testing of hypotheses, which requires a clear and unambiguous definition of pain embedded in a reasonable conceptual model. Due to the lack of a gold standard in infant pain assessment, there is also a need to take a multidimensional approach to infant pain assessment. While behavioural cues are highly informative, they can be improved upon through optimised combination with physiological measures, neural measures, and others (Vaart et al., 2019). The identification of novel indicators of infant pain will also be key to this process.

2.3.3 Summary

In this section, the BEST and COSMIN frameworks were introduced to provide structure and clarity regarding the nature of infant pain assessment tools and the degree to which they can be validated. Having previously provided a clear definition of pain and embedding it in a reasonable conceptual model, it is clear that, using

the BEST classification system, the verbal report measure of adult pain qualifies as a patient-reported outcome, while the infant PIPP(-R) pain scale (as well as other infant clinical pain scales) qualifies as an observer-reported outcome. Additionally, the verbal report qualifies as a gold standard for assessing pain experience, but no gold standard exists for infant pain experience assessment. Following the COSMIN guidelines for establishing test validity, once one defines the infant pain experience as the construct-of-interest, it follows that no available infant clinical pain scale has established content or criterion validity. To improve infant pain assessment and pain relief, adopting optimal definitions and conceptual models of pain, and using these to formulate and test hypotheses in order to advance construct validity is key. Finally, the lack of a gold standard indicates that the development of improved multimodal approaches to infant pain assessment will be of central importance, with the identification of novel pain correlates being a core requirement.

These points are relevant to the current work for two reasons. First, adopting clear definitions, models, and frameworks is essential for critical evaluation of the field of infant pain research, and thus identifying the major gaps in knowledge that the work of this thesis attempts to narrow. Second, the core aim of this thesis is to advance our use of MRI in infant pain research to avail of the unique strengths of the MRI modality, as well as identify novel neural features of infant nociception that could be used to further the construct validity of infant pain assessment tools.

2.4 Measures of pain-related brain activity

2.4.1 Neural and haemodynamic correlates of neonatal pain

While the test items that constitute infant clinical pain scales are typically limited to behavioural and physiological measures, research into infant pain has also used more technologically advanced methods, which provide greater insight into infant pain processing at the cost of clinical feasibility. To date, it has been electroencephalography (EEG) and near-infrared spectroscopy (NIRS) that have dominated research into the infant's brain response to pain and nociceptive inputs. EEG and NIRS are complementary techniques, that can non-invasively detect the brain's neural

activity (neurodynamic) response and blood perfusion (haemodynamic) response to nociceptive input. Using the BEST resource terminology, EEG and NIRS measures of infant pain constitute potential pain biomarkers. Unlike the observer-reported outcomes (e.g. PIPP(-R) and other infant clinical pain scales), which attempt to describe how an infant feels and depend on the subjective interpretation of the observer, these brain-based biomarkers are objective measures that are indicators "of normal biological processes, pathogenic processes, or responses to an exposure or intervention" (FDA-NIH Biomarker Working Group, 2016). However, exactly what aspect, correlate, or subcategory of pain brain-based measures are biomarkers of remains the most challenging and controversial question to biomarker identification, development, and validation.

The first brain-based recording of the neonatal response to pain was performed using NIRS in 2006 (Bartocci et al., 2006; Slater et al., 2006), demonstrating for the first time a cortical response to nociceptive input in both term and pre-term infants. The first EEG study that detected the nociceptive-specific neural activity likely underlying the NIRS-observed haemodynamic response was subsequently reported in 2010 (Slater et al., 2010a). Since these landmark studies, there have been several EEG and NIRS studies revealing mechanistic insight into the newborn brain's ability to process nociceptive input. Using these techniques, it has been revealed that the amplitude of the neonatal brain neurodynamic and haemodynamic response to nociceptive input increases with age (Slater et al., 2006; Hartley et al., 2016). NIRS has revealed that the cortical haemodynamic response has a larger amplitude to noxious input compared to non-noxious input (Verriotis et al., 2016), however, no noxious-specific haemodynamic response pattern has been reported. Using EEG, a noxious-specific event-related potential (ERP) can be distinguished when comparing noxious and non-noxious tactile stimulation (Slater et al., 2010a). From around 35 weeks PMA onwards, this noxious-specific ERP response (strongest over the vertex) becomes the dominant noxious-response EEG feature, while non-noxious-specific neuronal burst activity responses to both noxious and non-noxious tactile stimulation dominates the EEG prior to 35 weeks PMA. While the current

inability to distinguish noxious-response from non-noxious-response activities using EEG recordings certainly does not demonstrate that infants under 35 weeks cannot distinguish the two modalities, it does mean that by at least 35 weeks PMA, it has been demonstrated that infants differentially encode these two tactile modalities. Finally, using the EEG-recorded noxious-specific ERP amplitude, it has been demonstrated that peri-term-aged neonates are able to intensity encode nociceptive inputs (Hartley et al., 2015).

While EEG and NIRS are the dominant research modalities for studying infant brain responses to pain, fMRI is another rarely-used alternative. Similar to NIRS, fMRI records the brain's haemodynamic activity. However, while NIRS uses the different optical properties of oxy- and deoxy-haemoglobin to measure changes to both oxy- and deoxy-haemoglobin concentration, fMRI relies on the different magnetic properties of oxy- and deoxy-haemoglobin to measure relative changes to deoxy-haemoglobin concentrations. FMRI thus provides alternative haemodynamic-based insight into infant brain processing of pain. There are currently only three fMRI studies into infant nociception (Williams et al., 2015; Goksan et al., 2015; Goksan et al., 2018). Different aspects of these studies will be discussed throughout the thesis, as they become relevant. In brief, the Williams study revealed, in a similar fashion to EEG (Hartley et al., 2015), that the amplitude of the infant haemodynamic response reflects intensity encoding of nociceptive input (Williams et al., 2015). Additionally, this intensity encoding was found in both primary and secondary somatosensory cortices. The 2015 Goksan study compared the spatial pattern of haemodynamic response activity between adults and neonates and found overall similar patterns of activity, with some population differences too (Goksan et al., 2015). The 2018 Goksan study found that an infant's trial-to-trial variability in noxious-response amplitude was related to the prestimulus functional connectivity of a pain-relevant brain network (Goksan et al., 2018).

2.4.2 fMRI strengths and weaknesses

FMRI is the central modality used throughout this thesis. There are many strengths to fMRI that are lacking from other approaches (behaviours, physiology, EEG, and NIRS) to studying infant pain. There are also many challenges that have prevented large-scale adoption of fMRI for studying infant pain. In this section, the key strengths and challenges of using the fMRI technique will be outlined.

An advantage of fMRI (and both EEG and NIRS) over the behavioural and physiological measures that constitute infant clinical pain scales, is that the data originate from brain activity. Pain is an unpleasant sensory experience that emerges from thalamocortical brain networks, the activity of which is measured by fMRI, EEG, and NIRS. This feature can greatly help with the construct validity causality issue mentioned in Section 2.3.2. While behavioural and physiological changes typically associated with pain can be mediated without the involvement of thalamocortical pathways, the pain-related signals detected by fMRI, EEG, and NIRS require this circuitry. While the exact causal relationship between these brain-based measures and pain is not straight-forward to establish, the brain-based nature of the signals helps overcome many of the ambiguities inherent to behavioural and physiological measures.

There are three inter-related properties of fMRI that give it significant advantages over both EEG and NIRS: image reconstruction is a well-posed problem, whole-brain coverage, and relatively high spatial resolution. First, the fMRI signal (and all MRI signals in general) is generated by inputting radiofrequency (RF) energy into the brain tissue to perturb the system, then recording the re-emitted RF signal. This allows the brain to be actively probed in a highly controlled manner using the MRI gradients, to acquire hundreds of thousands of k-space points and thus hundreds of thousands of image space voxels. When converting the recorded MRI signal to an image (i.e. image reconstruction), there is typically a one-to-one mapping between recorded signal and voxel signal (in non-accelerated sequences), with the number of image space voxels determined by the acquisition sequence. Due to this ability to actively probe the brain using the MRI gradients to collect abundant

brain-related signal, the image reconstruction process (solving the inverse problem) is a well-posed problem. Thus, converting the recorded MRI signal to an image is a relatively straight-forward process, compared to EEG or NIRS. Unlike MRI, EEG and NIRS do not perturb the system, but simply passively record the signals emitted by the brain. The amount of spatial signal recorded is thus limited by the number of sensors (electrodes in EEG, optodes in NIRS), with 25 EEG electrodes or fewer being typical in infants. In these modalities, the source reconstruction issue is massively underdetermined, making the ability to project these data onto a brain surface very challenging. There are several advantages of having brain signals in source space, including increased biological interpretability (contrast the interpretation of signal localised to an occipital electrode to signal localised to primary visual cortex) and increased options for biologically meaningful analysis (contrast the interpretation of functional connectivity measures of correlating signal timeseries from left and right temporal electrodes to correlating signal timeseries from left and right primary auditory cortex).

Second, fMRI is sensitive to brain activity throughout the entire brain. In contrast, both EEG and NIRS are maximally sensitive to cortical regions that can be passively detected directly beneath the sensor locations on the scalp, with minimal sensitivity to deep structures (Hoshi, 2003; Burle et al., 2015; Verriotis et al., 2016). While high-density (256-electrode) EEG studies have the potential to record deep brain structures (Seeber et al., 2019), the limitations of infant EEG (and NIRS) mean sensitivity is limited to the activity of superficial structures. Key brain regions involved in pain processing, such as the thalamus, amygdala, and orbitofrontal cortex, are typically not accessible with techniques other than fMRI.

Third, fMRI typically works at the millimetre spatial resolution scale. This is controlled at the data acquisition stage determined by the degree to which k-space is sampled, with 2 mm voxels being typically achieved in infant fMRI. While the spatial resolution of both EEG and NIRS is again a function of the sensor density, the challenges of working with infants typically limits the spatial resolution of these methods to the scale of centimetres (Hoshi, 2003; Burle et al., 2015; Verriotis

et al., 2016). This issue of spatial resolution is more problematic for infants than adults, because of the relative difference in brain size. For example, a 1 cm^3 volume spans much more brain matter in an infant compared to an adult, which means that matched absolute spatial resolutions between adults and infants results in lower effective spatial resolution in infants. Ideally, a recorded signal could be assigned to a single functional brain region, but poor spatial resolution means signals from several functional regions will be blurred together making biologically meaningful interpretations of the data difficult.

There are several aspects of fMRI that make it a challenging modality to use when studying infant pain. The first challenge of fMRI is shared with NIRS, and this is the haemodynamic nature of the signal. Typically, the researcher is interested in the neural activity related to the processing of nociceptive input, and these neural processes are detectable using EEG. However, fMRI (and NIRS) rely on changes in cerebral perfusion of brain regions as a result to changing levels of brain activity i.e. the haemodynamic response. In fMRI, this haemodynamic response causes the blood oxygen level dependent (BOLD) signal that underlies fMRI, and this haemodynamic response is linked to the changing neural activity by a process called neurovascular coupling (Logothetis, 2003; Harris, Reynell, and Attwell, 2011). In infants, not only is the neural architecture of the brain immature, but so too is the cerebral vasculature and the neurovascular coupling mechanism, making the link between infant noxious-related neural and haemodynamic responses potentially more challenging to interpret (Kozberg et al., 2013; Kozberg and Hillman, 2016). (This topic is noted here and will be returned to in later chapters as it arises.) Additionally, the haemodynamic response is relatively sluggish in comparison to the underlying neural activity. The noxious-specific EEG ERP response to nociceptive input is over in under a second (Slater et al., 2010a; Hartley et al., 2017), while the associated haemodynamic response signal only peaks approximately 7 s after stimulus delivery (Arichi et al., 2012). This sluggishness of the BOLD signal confers a relatively poor temporal resolution on fMRI. Additionally, infant motion related to the delivery of a noxious stimulus will have a greater effect on the fMRI BOLD

signal than the EEG ERP signal, as infants typically exhibit motor responses such as facial grimacing and crying 1 – 2 s after stimulus delivery. Thus, the EEG ERP signal can often be collected without motion contamination, while the fMRI BOLD signal would be dramatically influenced by subject motion.

This motion issue is one practical reason why infant pain is typically not studied in the MRI environment. The other major issue is related to ethical concerns. It is typically not possible to safely perform clinically required painful procedures in the MRI environment, and ethical constraints ensure distress-inducing pain will not be inflicted on infants for research purposes alone. Thus, it is typically not possible to study infant pain in the MRI setting for ethical reasons. While it is possible that an infant may experience an experimental noxious stimulus (such as the 128 mN pinprick) as mildly painful without exhibiting behavioural or physiological signs of distress, this would not be a convincing or consistent description across infants, especially for infants who appear to sleep through the process. This inability to directly apply a painful stimulus when using fMRI to study infant pain is a major challenge. However, as outlined in Section 2.1.1, nociceptive pain is caused by nociceptive input, and low-threshold nociceptive input can result in activation of brain regions involved in processing pain without causing pain or behavioural distress to the infant. Substituting low-intensity nociception for pain allows the researcher to study an infant's cerebral pain system using fMRI, but conclusions drawn from such fMRI studies will require confirmation with alternative modalities (behavioural, physiological, EEG, NIRS, etc.) outside the MRI environment, under situations where directly studying pain is possible.

2.4.3 Summary

In this section, the three main modalities for measuring infant pain-related brain activity were introduced: EEG, NIRS, and fMRI. To date, EEG and NIRS are the most common modalities used to study infant pain, and these studies have established that newborn infants process nociceptive input at the level of thalamocortical circuits. These methods have also established some basic principles

of infant pain processing in the brain, such as an age-related increase in noxious-response activity, the ability to differentiate noxious from non-noxious tactile stimuli from at least 35 weeks PMA, as well as noxious stimulus intensity encoding. The much less frequently used fMRI modality has also been successfully used to establish haemodynamic stimulus intensity encoding as well as qualitative similarities between infants and adults in the spatial pattern of brain activity in response to nociceptive input. EEG, NIRS, and fMRI all provide the ability to measure brain activity, which is important to note considering pain is an emergent property of the brain's neural activity, and the test items that constitute clinical pain scales rely on behavioural and physiological activities that need not involve the thalamocortical circuits necessary for pain experience. These advantages of brain-based measures and previous EEG, NIRS, and fMRI findings highlight the importance of advancing the use of these methods to study infant pain, as well as form an important literature base that will help interpret the fMRI findings in this thesis.

Three advantages of using fMRI over EEG or NIRS were outlined, including fMRI's whole-brain coverage, relatively high spatial resolution, and well-posed image reconstruction problem. Finally, some drawbacks to using fMRI to study infant pain were outlined. The fMRI BOLD signal is a sluggish haemodynamic response with immature coupling to the underlying neural activity. The indirect nature of the BOLD signal makes interpretation more challenging than neural signals recorded using EEG. The sluggishness of the BOLD signal results in fMRI having relatively poor temporal resolution compared to EEG, and relatively high sensitivity to stimulus-correlated head motion compared to EEG analyses (if these EEG analyses are limited to the immediate $1 - 2$ s post-stimulus period, as is often the case). Throughout this thesis, attempts will be made to capitalise on the strengths of fMRI and to overcome or mitigate the weaknesses of fMRI. The importance of these strengths and weaknesses will be further addressed throughout the thesis, as they arise.

2.5 MRI: a brief overview

2.5.1 Introduction

The work in this thesis used two broad categories of MRI modalities: functional MRI (fMRI) and diffusion MRI (dMRI). FMRI data tracks the changing cerebral haemodynamics over time, which are driven primarily by fluctuations in the brain's neural activity. Thus, fMRI is an MRI modality for studying brain function. These fluctuations in neural activity can be due to spontaneous brain activity (i.e. not caused by researcher-specified stimulation or task), or be due to stimulus-response or task-response activity (i.e. directly caused by researcher-specified stimulation or task). The former is referred to as resting-state fMRI (rs-fMRI), and the latter as task fMRI. For this thesis, the "task" was the application of a mild noxious stimulus, so this type of fMRI data will be referred to as noxious-response fMRI. In contrast, dMRI data probes how freely water molecules diffuse throughout the brain, which is driven primarily by microstructural barriers to water diffusion, such as the presence or absence of cellular membranes, intracellular structures, myelin, and extracellular matrix. Thus, dMRI is an MRI modality for studying brain structure.

In this thesis, to improve understanding of the infant cerebral haemodynamic response to nociceptive input, infants' noxious-response fMRI data were studied in detail. An MRI-based multimodal approach was adopted, in which the noxious-response fMRI data were related to both the rs-fMRI and dMRI data.

2.5.2 Principles underpinning functional and diffusion MRI

The following background on MRI physics, covering topics relevant to the work of thesis, has been summarised primarily from Chappell, Okell, and Jenkinson, 2020 and Huettel, Song, and McCarthy, 2014.

Magnetisation and excitation

In basic terms, MRI uses a large magnetic field to interact with molecules in the body. The majority of MRI sequences are designed to interact with the hydrogen nuclei of water molecules, the most abundant molecule in the body. Each water

molecule can be thought of as a tiny bar magnet with its own magnetic field pointing in a particular orientation, which can be affected by other sources of magnetic fields in its environment. The MRI main magnetic field (B0) imposes a preferential orientation direction on the collection of water molecules in the brain, parallel to the B0 field. The MRI field strength is measured in units of tesla (T). The hydrogen nuclei of water molecules precess (rotate) around the axis of the B0 field with a rotational frequency that is proportional to the total field strength to which the nuclei are exposed. For the nuclei of hydrogen atoms in water molecules at 3T, this frequency is approximately 127.7 MHz, which lies in the radiofrequency (RF) range. The frequency at which these hydrogen nuclei precess in the magnetic field is their resonant frequency.

To detect the net magnetic signal from the hydrogen nuclei of water, the net magnetisation needs to be brought out of alignment with the B0 field. This is accomplished using a second magnetic field, the B1 field or RF field, which will be called the excitatory RF pulse. This excitatory RF pulse is tuned to vary at the resonant frequency of the molecules one wants to excite. To impose spatially varying field strengths across the brain, three gradient fields are added to the B0 field, each imposing a field gradient along the x-, y-, and z-axes. Typically, the brain is excited in a slice-wise manner along the z-axis. During image reconstruction, the signal emitted by the excited nuclei can be localised based on precessional frequency. This is due to the fact that the precessional frequency is directly proportional to the magnetic field strength, which varies according to space in a manner controlled by the user through varying the gradient fields.

Relaxation

Immediately after excitation, the component of the net magnetisation that is parallel to B0 is at its minimum due to being maximally out of alignment with B0, and the component of the net magnetisation that is perpendicular to B0 is at its maximum due to the constituent hydrogen nuclei precessing maximally in phase. Over time, the net magnetisation parallel to B0 increases as the constituent

hydrogen nuclei realign with B0, returning to the state pre-excitation. This return to alignment with B0 is called T1 relaxation and is quantified using the T1 relaxation time constant. Similarly, over time, the net magnetisation perpendicular to B0 decreases to zero as the constituent hydrogen nuclei dephase, returning to the state pre-excitation. This dephasing is called T2* relaxation and is quantified using the T2* relaxation time constant.

Both fMRI and dMRI contrast rely on relaxation due to dephasing, so further detail is given on this process. The T2* relaxation process can be broken down into two broad categories of causes. First, dephasing occurs due to random Brownian motion of molecules in the brain. A hydrogen nucleus precesses at a frequency determined by the field strength of its environment. In addition to the B0 and gradient fields, other molecules in the brain have their own magnetic field that contributes to a hydrogen nucleus's environment. The random molecular motion causes unpredictable non-static field inhomogeneities resulting in dephasing. Second, dephasing occurs due to relatively static sources of field inhomogeneities, such as deoxyhaemoglobin in blood vessels (important for fMRI image contrast, discussed more below) and air filled sinuses in the skull (relevant to signal loss called dropout). Dephasing due to random molecular interactions alone is called T2 relaxation, dephasing due to relatively static effects alone is called T2′ relaxation, and the combination of both is called T2*.

It is worth noting that, due to the relatively static nature of the causes of T2′, these dephasing effects can be mitigated using a second RF pulse. After the first excitatory RF pulse, another "refocusing" RF pulse can be used to flip the net magnetisation, which can bring the precessing hydrogen nuclei temporarily back into phase, undoing the T2′ dephasing effects only. MRI sequences that use a single excitatory RF pulse prior to readout (described below) are called Gradient Recalled echo (GRE) acquisitions, while those that follow the excitatory RF pulse with a refocusing RF pulse are called Spin Echo (SE) acquisitions.

Nature and acquisition of fMRI and dMRI data

The basis for the fMRI and dMRI signal differ in their dependency on dephasing effects (T2* vs T2) and thus the acquisition type used (GRE vs SE). FMRI depends on the T2′ effect caused by deoxyhaemoglobin in blood vessels, and consequently fMRI is T2*-dependent and uses a GRE acquisition. Oxyhaemoglobin and deoxyhaemoglobin have different magnetic properties, such that deoxyhaemoglobin is disruptive of the magnetic field in a way that oxyhaemoglobin is not. Deoxyhaemoglobin causes local field inhomogeneities and the resultant dephasing results in signal loss. This is called the Blood Oxygen Level Dependent (BOLD) effect. In adults, an increase in neuronal activity typically results in an overabundance of blood flow increase to the active region, resulting in a relative decrease in deoxyhaemoglobin, and a reduction in dephasing. Thus, increases in brain activity typically result in an increase in the BOLD signal.

DMRI depends on the T2 effect caused by random molecular motion, and consequently dMRI is T2-dependent and uses a SE acquisition to remove unwanted T2′ effects. The dMRI acquisition requires additional diffusion weighting gradients, one before and one after the refocusing RF pulse, that enhance the dephasing effect along a given diffusion weighting gradient direction. However, due to the refocusing RF pulse, the second diffusion weighting gradient undoes the enhanced dephasing of the first diffusion weighting gradient as long as the molecules have not moved. Molecules that do move along the direction of the diffusion weighting gradient are not refocused, and those that remain static are. Using this acquisition, water molecules that do not move along the diffusion weighting gradients exhibit dephasing due to T2 (i.e. interactions with surrounding molecules that are diffusing), while those that do diffuse along the diffusion weighting gradients exhibit dephasing due to both T2 (i.e. interactions with surrounding molecules that are diffusing) and diffusion (i.e. diffusion of the water molecule in question itself). The degree of diffusion is thus encoded in the degree of signal loss.

Readout

After exciting a slice of the brain along the z-axis, signal from the xy-plane is detected by the head coil, a process called readout. The MRI signal from the entire xy-plane is recorded simultaneously. Within each slice, the active manipulation of the xy gradients allows the user to impose patterns of spatial frequencies. The spatial frequency data can then be used to construct an image using the Fourier Transform. Low spatial frequencies encode broad image outlines, while high spatial frequencies encode rapidly changing spatial details such as edges.

A 2-dimensional coordinate system of spatial frequencies, or frequency space (k-space), is sampled during readout. For both functional and diffusion imaging, the most commonly used readout approach is Echo Planar Imaging (EPI). EPI captures all the data of a single slice in a single readout (single shot EPI), which means both the x and y gradient fields are used during readout to cover all of the pre-specified k-space. In EPI, the direction along which the gradient field is varying more slowly, and thus has a lower average bandwidth, is called the phase encoding direction; conversely, the more rapidly varying, higher bandwidth gradient defines the direction called the frequency encoding direction.

Acceleration

As mentioned in section 2.4.2, MRI is rich in spatial information. There can be as many data points collected (in k-space) during readout that are required for the final image (in image space), making the image reconstruction issue well defined. There is also extra spatial information encoded in the head coil. The different head coil channels are distributed around the head, which results in channels being more sensitive to signals coming from closer brain regions than regions at a distance. This additional head coil spatial information allows us to acquire signal from the brain in an accelerated fashion that would typically result in spatial ambiguity when reconstructing the imaging. The spatial ambiguity can be resolved using the head coil information, and the accelerated data acquisition can be used to attain higher spatial and/or temporal resolutions. Two broad approaches to acceleration

include in-plane acceleration and multislice (or multiband) acceleration. In-plane acceleration involves acquiring a particular slice more quickly by acquiring less of k-space, while multiband acceleration involves acquiring multiple slices simultaneously.

2.5.3 Artefacts and data clean-up for fMRI and dMRI

Head motion and distortion

Two of the largest sources of fMRI and dMRI data contamination that are relevant for the work in this thesis are subject head motion and geometric distortion. Head motion results in changes in location of the brain within the imaging field of view, which manifests in several ways in the data (detailed below). Distortion is due to field inhomogeneities in the net magnetic field (B0 plus gradient fields). These inhomogeneities mean a particular brain region will be experiencing a magnetic field stronger or weaker than expected, will thus be precessing at a faster or slower speed than expected, and will consequently be mislocalise during image reconstruction because localisation of the MRI signal is based on resonance frequency. Distortion means images are not anatomically faithful, and there are several causes of these field inhomogeneities.

For both fMRI and dMRI data, the entire brain is scanned in a slice-wise manner, with all brain slices constituting a volume. If the subject moves their head, each subsequent volume will be out of alignment. Thus for a given point in an image, signal at that point will come from different brain regions from volume to volume. This will be called volume misalignment. If a subject moves during the collection of the slices of a volume, then subsets of slices within a volume will also be out of alignment. This will be called slice misalignment. In addition to misalignment, head motion can result in slice hyper- and hypo-intensities due to spin history effects. In brief, the signal intensity emitted by a brain region depends on the time between excitations. When the subject is still, this delay between excitations is constant for all brain regions. However, when the subject moves, brain regions can move to spatial locations that are either excited earlier or later than desired which leads to intensity variance within the volume. Head motion-induced volume

misalignment, slice misalignment, and spin history effects affect both fMRI and dMRI data. One other effect worth mentioning is slice hypointensities in dMRI data due to motion-induced signal dropout. This occurs if the subject moves during diffusion encoding, as the diffusion encoding gradients are designed to generate signal loss for non-static water molecules.

Comparing healthy adults to healthy infants, some of these motion-related artefacts are worse for infants. The degree of slice misalignment is approximately proportional to the rate of change of head motion (Andersson et al., 2017). In general, infants perform more sudden and jerky head motions in the scanner, and this higher rate of change of head motion can cause substantial slice misalignment (for examples, see Bastiani et al., 2019, Fitzgibbon et al., 2019, and Baxter et al., 2019). If we hypothetically matched an adult and infant for head motion profiles, the spin history effects in infants would also be more substantial. This is a consequence of the longer relaxation times for infant brain tissue due to the higher water content, and the dependency of spin history effects on relaxation time constants.

The two most relevant causes of distortion are susceptibility-induced distortion and eddy current-induced distortion. Susceptibility-induced distortion affects both fMRI and dMRI data, and is due to B0 inhomogeneities cause by the subject themselves. The 3T B0 field is relatively homogeneous until an object is placed in the bore. Once in the bore, the subject is magnetised, and the degree of subject-induced inhomogeneity depends on the exact geometrical properties of the subject. However in general, regions where brain tissue and air are particularly proximal, such as the skull's sinuses and ear canals, are the most strongly distorted. Using an EPI readout, susceptibility-induced distortion occur along the phase encoding direction. As mentioned previously, the frequency encoding direction has high bandwidth so the relative effect of susceptibility-induced off resonance within this high gradient direction is negligible, while in contrast, the phase encoding direction has low bandwidth so the relative effect of susceptibility-induced off resonance within this low gradient direction can be quite substantial. As long as the subject remains static, the susceptibility-induced distortions are also static. However, when

the subject moves, there is an interaction between subject susceptibility and head motion, termed susceptibility-by-movement distortions. Specifically, susceptibility-by-movement distortions occur when the head is rotated around an axis other than that of the B0 field i.e. rotations around the x- and y-axes cause this artefact. Finally, eddy current-induced distortions affects dMRI data only (the effects are negligible in fMRI data). The diffusion weighting gradients are gradient intensive, with strong gradients being rapidly ramped up and down. These changing gradients induce eddy currents in the scanner hardware creating magnetic fields that counteract the gradients, preventing them from changing as rapidly as desired. The decay of eddy currents typically persist into the EPI readout period causing additional eddy current magnetic field-related inhomogeneities.

Again, comparing healthy adults to healthy infants, susceptibility-by-movement distortion artefacts are typically worse for infants. The degree of susceptibility-by-movement distortion is proportional to the amount of head motion rotation around the x- and y-axes (Andersson et al., 2018). In addition to infant head motions typically being more sudden than in adults, infants typically have an overall larger average degree of head motion. This higher degree of head motion (rotations around x- and y-axes as well as other rotations and translations) means susceptibility-by-movement distortion effect is more problematic in infants (for examples, see Fitzgibbon et al., 2019, and Baxter et al., 2019).

EDDY and FIX

Due to the larger influence of slice misalignment, spin history effects, and susceptibility-by-movement distortion in infant populations, these artefacts should not be ignored during fMRI and dMRI data preprocessing if possible. This means simple rigid body data realignment approaches are not sufficient, as they will not account for local mislocalisation of signal. Two preprocessing tools that address these artefacts, EDDY and FIX, will now be described.

EDDY is a motion and distortion correction tool initially developed for diffusion data and recently expanded for use in fMRI data (Andersson and Sotiropoulos,

2015; Andersson and Sotiropoulos, 2016; Andersson et al., 2016; Andersson et al., 2017; Andersson et al., 2018; Fitzgibbon et al., 2019). Before introducing EDDY, it is important to note that EDDY takes as an input a fieldmap image. In this thesis, we collect a separate fieldmap acquisition which is essentially a static snapshot of the actual magnetic field including all static inhomogeneities. EDDY uses this fieldmap to correct for susceptibility-induced distortions. EDDY uses a generative prediction model to simultaneously perform both motion and distortion correction. As used in this thesis for dMRI data preprocessing, EDDY corrects for six artefact effects: volume misalignment, slice misalignment, slice dropout, susceptibility-induced distortion, susceptibility-by-movement distortion, and eddy current-induced distortion. For dMRI data, predictions are generated using a weighted mean of all volumes. For a particular volume with a given diffusion encoding direction, the largest weights are given to volumes with diffusion encoding gradient directions most similar to the volume in question. This is because the diffusion weighted contrast will be most similar to volumes with closely matched diffusion encoding directions, but averaging over volumes with both close to parallel and anti-parallel directions (as diffusion contrast should be identical for +x and -x directions) should average out eddy current-induced distortions resulting in a less distorted prediction than the volume in question. The volume in question is nudged closer towards the prediction on each iteration by more accurately modelling the sources of signal corruption i.e. motion, both susceptibility-induced and eddy current-induced distortions, and outlier slices due to bulk motion-induced slice dropout. Due to better modelling of these generative processes, correction of each volume improves with each iteration leading to better volume-wise prediction models toward which each volume is nudged. In summary, EDDY simultaneously performs both motion and distortion correction in an integrated manner using an iterative generative model approach.

As used in this thesis for fMRI data preprocessing, EDDY corrects for four artefact effects: volume misalignment, slice misalignment, susceptibility-induced distortion, and susceptibility-by-movement distortion. For fMRI data, predictions are generated using the unweighted temporal mean of all volumes. Weighting is

only required for dMRI data due to the dramatic changes between volumes in tissue contrast as the diffusion encoding direction changes. Eddy current-induced distortions are not modelled, because these are negligible in fMRI data due to the absence of the diffusion encoding gradients. Outlier slices due to bulk motion-induced dropout are also currently not identified and replaced, but are dealt with at a later preprocessing stage using the FIX tool (see below). Apart from these differences, correction for bulk head motion (volume and slice misalignment) and susceptibility-induced distortion (both static and susceptibility-by-movement) occurs in a similar iterative generative model approach as described above, essentially treating each fMRI volume as a diffusion B0 volume i.e. a spin echo volume with no diffusion encoding.

In fMRI data preprocessing, the post-EDDY residual effects of motion and distortion can be further dramatically reduced using the denoising tool FIX (FMRIB's ICA-based Xnoiseifier). FIX is a denoising tool that can remove variance related to many unwanted structured sources, and is used to clean-up fMRI data after motion and distortion correction (Salimi-Khorshidi et al., 2014; Griffanti et al., 2014; Griffanti et al., 2017). FIX denoising consists of five steps: spatial independent component analysis (ICA), feature extraction, classifier training, application of classifier to new datasets, and denoising i.e. removal of noise components from the data. ICA is a type of blind source separation technique, which decomposes a multivariate signal into its component signals in an unsupervised manner. Thus, the first step of FIX denoising involves unmixing the fMRI data into a linear combination of components. Each component consists of a timeseries and a spatial map outlining the spatially varying timeseries amplitudes. Due to the use of spatial ICA, it is the spatial maps that are spatially independent. The fMRI spatiotemporal signal is a mixture of BOLD-related variance and both motion- and distortion-related variance, and spatial ICA (as implemented in FIX) can separate these different contributing sources. The fMRI signal is also composed of other structured signals such as cardiovascular and image acquisition artefact signals, which can feature as independent components, as well as unstructured signal such as thermal noise.

This unstructured signal (and lower SNR structured signal) does not feature as independent components due to a principle component analysis (PCA) step prior to ICA that selects a subset of the data for ICA analysis. Once the data are unmixed into independent sources of structured signal, FIX performs feature extraction. FIX extracts 178 features across both temporal and spatial domains. The independent components must be manually labelled as either signal or noise in order for the FIX classifier (an ensemble machine learning classifier i.e. a classifier that combine multiple classifiers) to be trained. After training using the spatial/temporal features and manual signal/noise labels, the FIX classifier can be applied to new data that are reasonably matched to the training set. Finally, FIX cleans the data by removing the noise component timeseries from the fMRI data using multiple regression. The regression residuals are the cleaned data.

Relevant to the previous discussion of head motion and distortion corrections in EDDY, FIX denoising can dramatically reduce the residual influence of volume and slice misalignment caused by head motion, slice hyper- and hypo-intensities due to spin history effects caused by head motion, and susceptibility-by-movement distortions caused by the interaction between head motion and susceptibility-induced distortion. As will be expanded upon in Chapter 3, FIX denoising can also help remove noise variance from fMRI data due to a range of other sources, such as subject physiology and MRI artefacts.

The use of EDDY in dMRI data preprocessing and the combined use of EDDY followed by FIX in fMRI data preprocessing should dramatically improve the infant MRI data clean-up process. Classical motion and distortion correction approaches that sequentially perform rigid body motion correction followed by static distortion correction ignore slice misalignment and susceptibility-by-movement effects. And omitting FIX denoising from fMRI data preprocessing does not remove the abundant residual motion, distortion, and spin history effects. Due to the high degree of motion contamination coupled with the unique high water content brain composition, infant fMRI and dMRI data require advanced data clean-up methods in order to allow meaningful outputs from subsequent analysis methods.

2.5.4 Data analysis approaches

A brief outline of the analysis methods that were adopted in this thesis, as well as common analysis alternatives, are provided in this section. The outline aims to help contextualise the specific methodology employed in subsequent chapters.

Noxious-response fMRI

For the noxious-response fMRI data, the stimulus was a sharp touch pinprick applied to the infant's foot. The sharp nature of the stimulus is known to activate Aδ-fibres, but the stimulus used was of low-enough intensity (128 mN) to not cause clinical concern or behavioural distress (Hartley et al., 2015). The delivery of a sharp touch pinprick to the infant's foot was designed to mimic the clinical heel lance blood test procedure (Courtois et al., 2016). In task fMRI, there are two main experimental designs: block and event-related. In a block design, the stimulus (or task) is continuously applied for a multi-second block typically followed by a "rest" block with no stimulus, and these on-off blocks are repeated. In an event-related design, the stimulus is applied as multiple short individual trials. Block designs are inappropriate for experimental study of neonatal nociception in the MRI environment, so an event-related design with 1 s stimulus application trials was used (Goksan et al., 2015; Goksan et al., 2018).

The most widely adopted statistical analysis approach to study cerebral haemodynamic task and stimulus responses is to generate a predicted BOLD timeseries by convolving the stimulus-application timeseries with a haemodynamic response function (HRF). This predicted BOLD timeseries is then fitted to each voxel's fMRI recorded timeseries independently using a general linear model (GLM) approach (Monti, 2011). The GLM regression parameter is the predicted BOLD timeseries scaling factor, and is thus a measure of the haemodynamic response amplitude measured relative to baseline. This GLM approach generates a voxelwise map of regression parameters i.e. a haemodynamic response amplitude map. These response amplitudes can be assessed at several spatial scales, from individual voxels, to region-of-interest (ROI) and whole-brain measures. In general, this GLM

approach is very flexible facilitating a wide range of analysis questions, and is not limited to generating a simple amplitude map relative to baseline. However, the infant nociception paradigm used throughout this thesis involved a 10-trial event-related design with a single stimulus intensity (128 mN pinprick), so assessing response amplitude relative to baseline was the central noxious-response fMRI data analysis methodology.

Other than localised measures of amplitudes, measures of statistical association between brain regions are also possible for noxious-response fMRI data. Using a simple Pearson correlation coefficient to quantify the association between two brain regions' timeseries is a type of functional connectivity measure. Psychophysiological interaction (PPI) analysis is a type of functional connectivity analysis that can be used with noxious-response fMRI data (Friston et al., 1997), as an alternative to amplitude analysis. PPI and other functional connectivity analyses typically require the pre-specification of specific brain ROIs. Functional connectivity analysis is a valuable analysis alternative, but measures of functional connectivity were not assessed throughout this thesis. In chapter 3, both voxelwise and ROI-wise measures of response size were assessed; while in chapter 4, both voxelwise and whole-brain measures of response size were assessed.

Resting-state fMRI

Similar to noxious-response fMRI, rs-fMRI tracks the brain's functional haemodynamics. However, unlike noxious-response fMRI, the researcher does not apply any stimulus to the subject or request the subject to perform an explicit task. In adults, the resting-state is a behavioural state where the subject typically lies with eyes open, looking at a fixation cross, and remains awake; although, variations exist such as no fixation cross, or eyes closed. In infants, the ability to control behaviour and wakefulness is limited. Infants are typically fed, swaddled, and made comfortable in the scanner (identical to noxious-response fMRI setup). During the rs-fMRI run, the infant would ideally remain quietly asleep or quietly awake. Due to the lack of experimenter control over the subject's brain activity (in both adults and infants),

resting-state brain activity is typically described as spontaneous or unconstrained. The spontaneous fluctuations in cerebral haemodynamics typically occur in spatially regular patterns that closely match brain networks engaged in explicit stimulus or task processing (Smith et al., 2009). These spontaneous patterns of brain activity are called resting-state networks, and were first identified in adult fMRI in 1995 (Biswal et al., 1995). Similar resting-state networks were later identified in newborns in 2007 (Fransson et al., 2007). Resting-state networks have stereotypical spatial and temporal properties: they are haemodynamic fluctuations localised to grey matter, the timeseries of these networks slowly oscillate, and the frequency spectrum of these timeseries typically has the vast majority of spectral power (>90%) lying between 0.01 and 0.1 Hz (Niazy et al., 2015).

There are several methods for the identification of resting-state networks. Unlike task fMRI, there is no predicted BOLD response timeseries, so timeseries must typically be derived from the data. While ROI-based analyses, also known as seed-based correlation analyses (SCA), are quite common, the use of independent component analysis (ICA) is one of the most popular methods for resting-state network identification (Beckmann and Smith, 2004). The SCA-derived or ICA-derived timeseries for network identification can be considered analogous to the predicted BOLD timeseries of noxious-response fMRI analysis, only that the resting-state timeseries is data-derived. Similar to noxious-response fMRI analysis, the resting-state SCA- or ICA-derived timeseries can be fit in a whole-brain voxelwise manner using a GLM approach to generate a network map. Analogous to the noxious-response PPI analysis, these resting-state approaches to deriving network maps allow the assessment of functional connectivity, with Dual Regression being a specific example of this mapping approach to functional connectivity analysis (Beckmann et al., 2009; Nickerson et al., 2017). Alternatively, a collection of SCA- or ICA-derived timeseries can be used to generate a matrix of correlation coefficients, commonly referred to as a connectivity matrix or network matrix. These connectivity matrices are one of the most common rs-fMRI analysis approaches to studying functional connectivity (Smith et al., 2011).

Amplitude measures can be extracted from resting-state data, for example, by quantifying the amplitude of the SCA- or ICA-derived timeseries. Similar to the noxious-response fMRI analysis, the rs-fMRI analyses used throughout this thesis focused primarily on measures of activity amplitude rather than measures of functional connectivity. Three points motivate this methodological decision. First, there is a strong dependency of functional connectivity measures on signal amplitude, underscoring the primacy of amplitude (Cole et al., 2016; Duff et al., 2018; Friston, 2011), yet amplitude effects are currently under-explored in infant fMRI studies. Second, there is now increased recognition of the functional importance of fMRI activity amplitude (Bijsterbosch et al., 2017; Duff et al., 2008; Mennes et al., 2013; Xu et al., 2014; Zou et al., 2013), with resting-state activity amplitudes shown to associate with task activity in adults (Mennes et al., 2010; Mennes et al., 2011), a property of central importance in Chapter 4. Third, the challenges of studying infant pain with MRI typically constrains study sample sizes to modest numbers ($n < 20$) (Williams et al., 2015; Goksan et al., 2015; Goksan et al., 2018), so studies attempting to identify novel indicators of infant pain and nociception should avoid issues due to combinatorial explosion (N networks produce ${}^{N}C_{2}$ functional connectivity measures, but only N amplitude measures).

White matter dMRI

While the noxious-response fMRI and rs-fMRI modalities probe infants' brain function, dMRI is a modality for analysing primarily white matter microstructure. There are several dMRI models that can be adopted, and the available options can be limited by the nature of the data. The dMRI data used in this thesis was multishell with 140 directions, meaning most common dMRI models were applicable. The most commonly used dMRI model is the diffusion tensor imaging (DTI) model, which is typically used to quantify diffusivity-based model parameters. The two most commonly studied diffusivity-based DTI parameters are mean diffusivity (MD) which quantifies the overall degree to which water can freely diffuse in a voxel, and fractional anisotropy (FA) which quantifies the directional variability

of this diffusivity. However, the DTI model assumes Gaussian water diffusion, which typically is not the case in grey and white matter. The more recently described diffusion kurtosis imaging (DKI) model (Jensen et al., 2005; Jensen and Helpern, 2010) expands the DTI model to account for non-Gaussian diffusion, thus more accurately describing the diffusion environment. Without the assumption of Gaussian diffusion, the DKI model is able to quantify diffusional kurtosis, which is a measure of heterogeneity in the diffusional environment. Analogously to MD quantifying diffusivity across all diffusion directions, the mean kurtosis (MK) parameter quantifies the heterogeneity in diffusivities averaged across all diffusion directions. For each infant, a whole brain map of these different diffusion parameter estimates can be generated. Whole-brain voxelwise comparisons of local white matter diffusion properties such as MD, FA, or MK, can be assessed using a GLM analysis approach (Smith, Kindlmann, and Jbabdi, 2014).

The DTI and DKI models' parameter maps estimate local diffusion properties independently at each brain voxel. Alternatively, measures of white matter connectivity strength between two grey matter regions could be assessed. Thus, fMRI data (both resting-state and noxious-response) can be used to measure functional connectivity, and dMRI data can be used to measure structural connectivity. There exist many available methods for quantifying structural connectivity from dMRI data using tractography analysis (Behrens, Sotiropoulos, and Jbabdi, 2014). Deterministic tractography using DTI or DKI parameter maps is limited due to inability of these models to capture uncertainty in diffusion direction estimation. Using probabilistic tractography models, such as FSL's ball-and-stick models (Behrens et al., 2003; Behrens et al., 2007), more accurate measures of structural connectivity strength are possible. Similar to functional connectivity analyses that required the extraction of ROI or network timeseries by specifying specific brain regions or networks, structural connectivity analyses also typically require the specification of ROIs.

As will be outlined in Chapter 4, a whole-brain analysis of local white matter microstructural properties was a suitable choice for this thesis, negating the need to

specify or define specific ROIs or limit analysis to subregions of the brain. Due to the multishell nature of the data, the DKI model, which accounts for non-Gaussian water diffusion, was selected over the DTI model. Finally, the primary DKI measure selected for whole-brain voxelwise structural analysis was MK. The rationale for selecting MK is outlined in detail in Chapter 4.

2.5.5 Summary

In this section, three broad topic of MRI that are relevant to the work in this thesis were introduced: principles of fMRI and dMRI imaging, artefacts and data clean-up for fMRI and dMRI data, and common analysis methodologies for fMRI and dMRI data.

Both the fMRI and dMRI data used in this thesis were acquired with a B0 field of 3T, the excitatory RF pulse was used to excite hydrogen nuclei of water molecules, the gradient fields were used to sequentially excite slices of the brain along the z-axis, the readout approach used was EPI, and the y-axis gradient field direction (which runs along the Anterior-Posterior, or AP, direction) was the phase encoding direction. The fMRI data were acquired using GRE, which included an excitatory RF pulse prior to readout, and using multiband acceleration acquiring four slices simultaneously i.e. MB4. The dMRI data were acquired using SE, which included both excitatory and refocusing RF pulses as well as diffusion encoding gradients prior to readout. The dMRI readout also included multiband acceleration acquiring three slices simultaneously i.e. MB3.

Head motion and distortion are major sources of data contamination for both fMRI and dMRI data. These effects manifest in the data in several forms. Head motion can result in both volume and slice misalignments as well as slice hyper- and hypo-intensities. Distortion means the data are not anatomically faithful, and results from intrinsic magnetisation of the structure being imaged (susceptibility-induced), eddy currents induced within the MRI hardware itself (eddy current-induced), and interactions between subject motion and the magnetic field (susceptibility-by-movement). These motion and distortion effects are amplified in infants due

to increased average motion, more rapid and sudden motion, and longer tissue relaxation times due to higher brain water content. These challenges in the infant population require sophisticated approaches to data clean-up. The tools EDDY and FIX were introduced due to their central role in dMRI and fMRI data preprocessing throughout this thesis, and the substantial improvements in data quality they allow us to achieve even with heavily contaminated data.

Finally, the three MRI modalities used throughout this thesis were outlined: noxious-response fMRI, resting-state fMRI, and white matter dMRI. A description of the sharp touch noxious stimulus was provided, and the rationale for its use as a low-intensity non-skin-breaking substitute for the clinical heel lance procedure was given. Next, an overview of the brain properties being probed by each MRI modality was described, as well as a general description of common analysis approaches. In both noxious-response fMRI and rs-fMRI, measures of activity amplitude and functional connectivity can be extracted; and in white matter dMRI, measures of local diffusion properties and structural connectivity can be extracted. In this section, it was briefly motivated why measures of activity amplitudes were the primary focus of both the noxious-response fMRI and rs-fMRI data analyses, while local measures of diffusion properties were the primary focus of the white matter dMRI data analyses. The rationale for these choices will be further expanded upon in the relevant chapters, as they arise.

3

Optimising fMRI data analysis for the infant haemodynamic response to nociceptive input

3.1 Introduction

As was outlined in chapter 1, a major aim of this thesis is to further our understanding of the neurophysiological basis for individual variability in newborn infants' cerebral haemodynamic responses to nociceptive input. A major obstacle to achieving this goal is the lack of optimisation in the earliest fMRI data analysis stages, including data preprocessing and noxious-response signal modelling. Relevant to infant fMRI data analysis, some major differences between newborn infants and adults include the reduced size and degree of gyrification of the infant brain, the infant's immature haemodynamic response to neural activity, fundamental differences in brain tissue composition, and physiological and motor outputs. Despite these adult-infant differences, infant fMRI studies often rely on analysis tools that have been developed and optimised for the adult brain. The lack of data analysis optimisations for this challenging infant nociception fMRI paradigm prevents the achievement of data quality high enough to meaningfully assess infant brain activity at the individual level. The goal of the current chapter is thus to outline a data analysis pipeline optimised for the infant nociception fMRI paradigm.

A major obstacle to high quality infant fMRI data is subject motion. The high degree of infant motion during MRI scans, especially in noxious-response fMRI

experimental paradigms, results in data that is heavily contaminated with motion-related artefacts. Due to the significant impact motion has on MRI data quality (Power et al., 2012; Reuter et al., 2015; Satterthwaite et al., 2012; Siegel et al., 2017; Yendiki et al., 2014), removal of these motion-related signals is necessary. However, simply using data censoring and rejection based on recommendations from the adult literature leads to high levels of data loss. Applying spike regression and scrubbing to resting-state fMRI data to remove volumes with motion greater than 0.25 mm (Satterthwaite et al., 2013) and not including data with less than 4 minutes of continuous uncorrupted data (Power et al., 2014; Satterthwaite et al., 2013; Parkes et al., 2018), data loss of over 95% has been reported (Fitzgibbon et al., 2019). Considering how challenging infant MRI data are to acquire, improvements to dealing with motion in fMRI data are required. Advanced motion and distortion correction methods were tested in this study, as well as the value of incorporating and training a semi-automated spatial ICA-based denoising classifier (Salimi-Khorshidi et al., 2014), which, at the time of publishing this study, had only been successfully trained in one infant resting-state fMRI study (Ball et al., 2016).

In addition to motion, the rapidly changing brain size, gyrification, and tissue contrast (Paus et al., 2001; Dubois et al., 2014; Dubois and Dehaene-Lambertz, 2015), make the alignment of individual infant's MRI data for group analysis challenging. While neonatal structural template maps have been developed (Shi et al., 2011; Shi et al., 2018; Makropoulos et al., 2016; Schuh et al., 2018) and non-linear spatial normalisation tools have been adapted for use with infant data (Goksan et al., 2015), an optimised registration pipeline for infant functional data has yet to be developed and tested.

Spatial filtering is often performed to improve SNR by removing high frequency noise and reducing the effect of post-spatial normalisation misalignment (Lowe and Sorenson, 1997). Unfortunately, many infant fMRI studies, and most adult fMRI studies (over 80%), use spatial filtering kernels equal to or greater than 5 mm FWHM, which appears to be primarily due to the default settings of the two major MRI software packages (FSL's default is 5 mm FWHM, and SPM's default is 8

mm FWHM) (Carp, 2012). Given the small volume of the infant brain, a 5 mm kernel is likely too large; yet, some spatial filtering may be beneficial in datasets of modest sample size, assuming it is appropriate for subsequent analysis purposes. As such, the effect of spatial filtering was assessed.

Finally, the cerebral haemodynamic response to neuronal activation, the source of the fMRI BOLD signal, is fundamentally different between adults and infants. The immaturity of the neuronal tissue, the cerebral vasculature, and the neurovascular coupling mechanism results in an immature haemodynamic response (Harris, Reynell, and Attwell, 2011), which gradually changes morphology throughout development until adulthood (Kozberg et al., 2013; Kozberg and Hillman, 2016). A previous study (Arichi et al., 2012) into the neonatal BOLD response identified a bimodal double gamma response function for term neonates. While the canonical adult double gamma HRF (as implemented in the FSL software package) has a latency to peak of 5.5 s and an overshoot-to-undershoot ratio of 11:1 (Glover, 1999), the term-age newborn infant double gamma HRF has a latency to peak of 7 s and an overshoot-to-undershoot ratio of 1:1 (Arichi et al., 2012). However, to date, no independent study has compared the quality of fit of different HRF parameterisations based on the study by Arichi and colleagues, and so, HRF modelling was assessed in this study.

In total, five key stages of infant fMRI data analysis were tested in this study: motion and distortion correction, spatial normalisation, FIX denoising, spatial filtering, and HRF modelling. To assess the motion and distortion correction and spatial normalisation stages, equivalent outputs from two pipelines were compared. The first pipeline was based on FSL's FEAT (Jenkinson et al., 2002), an adult-optimised pipeline that has been adapted for use in infant data and used in previous published studies (Goksan et al., 2015; Goksan et al., 2018). These outputs were directly compared to the equivalent outputs from the infant-specific pipeline developed as part of the Developing Human Connectome Project (dHCP) (Fitzgibbon et al., 2019), that has yet to be released or tested on an independent non-dHCP non-resting-state fMRI dataset. The value of incorporating FIX denoising into the preprocessing pipeline was tested by comparing outputs with and without

the adoption of this analysis step. To test effects of spatial filtering on data quality, data outputs were compared for unfiltered data, and data filtered with 3 mm and 5 mm FWHM smoothing kernels. Finally, HRF modelling was based on the neonate-specific HRF outline by Arichi and colleagues (2012). Three parameterisations of varying complexity were compared: a single gamma function model, a double gamma function model, and a three basis function model.

This work provided the necessary foundations on which the work in the following thesis chapter was built. Additionally, the analysis pipeline outlined in this chapter provides a basis from which further infant fMRI studies into infant pain and nociception can be based. The successful adoption and implementation of the dHCP fMRI preprocessing pipeline also aided in the generalisation of the pipeline for non-dHCP datasets and non-resting-state fMRI modalities.

3.2 Methods

3.2.1 Infant recruitment and demographics

A clinical investigator of the Paediatric Neuroimaging Group recruited healthy neonates from the postnatal ward at the John Radcliffe Hospital (Oxford University Hospital NHS Trust) for an MRI scan. Infants were considered healthy if they were inpatients on the postnatal ward that never required admission to the neonatal unit, had no history of congenital conditions or neurological problems, and were clinically stable at the time of study. Written informed consent was obtained from parents prior to the study. Ethical approval was obtained from an NHS Research Ethics Committee (National Research Ethics Service, REC reference: 12/SC/0447), and research was conducted in accordance with standards set by Good Clinical Practice guidelines and the Declaration of Helsinki.

Fifteen neonates were included in this study. At the time of study, the mean postmenstrual age (PMA) was 39.3 weeks (range: 37.1 – 42.7 weeks), mean postnatal age (PNA) was 2.3 days (range: 1 – 7 days). The mean birth weight was 3,408 g (range: 2,235 – 4,570 g), and the sample included eight males and seven females.

3.2.2 Experimental setup and design

Neonates were transported to the Wellcome Centre for Integrative Neuroimaging (Oxford, UK), then screened for metal, fed, and swaddled prior to scanning. Infants were fitted with silicone ear plugs (Hush Plugz, Robinson Healthcare, Worksop, UK; 23 dB attenuation), ear-muffs (Minimuffs, Natus Medical Inc., Galway, Ireland; 7 dB attenuation), and ear-defenders (Em's 4 Bubs Baby Earmuffs, Em's 4 Kids, Brisbane, Australia; 22 dB attenuation), and placed on a vacuum-positioning mattress with additional soft padding around the head to restrict motion. Heart rate and blood oxygen saturation were monitored throughout scanning (Fibre Optic Pulse Oximeter, Nonin Medical, Plymouth, Minnesota).

An event-related experimental design was used, and the mild non-skin-breaking nociceptive stimulus was a 128 mN sharp-touch pinprick (PinPrick Stimulator, MRC Systems). Ten trials of the stimulus were delivered to the dorsum of the left foot, each trial was 1 s, and the minimum inter-stimulus interval was 25 s. This long inter-stimulus interval was used to minimise the influence of motion at the time of stimulus delivery. The stimuli were applied when the infants were naturally still, and were time-locked to the fMRI data using Neurobehavioural Systems software (Presentation, `https://www.neurobs.com`). Due to the size of newborn infants, the stimuli must be delivered with the researcher lying in the bore. To perform stimulus delivery in a time-locked manner with minimal researcher movement (to minimize field disruption), two researchers were required. The first researcher lay in the bore and delivered the stimulus once the infant was lying still and not moving the foot to be stimulated (i.e. naturally still). The second researcher, outside the scanner bore, indicated to the first researcher when to deliver the stimulus by gently tapping the first researcher. The second researcher recorded the time of stimulus delivery using the a button press that was visually time-locked to the time of contact of the pinprick with the infant's skin. Visual assessment by the second researcher of stimulus contact was always possible, and while the visual time-locking may introduce timing inaccuracies, these would be on the order of approximately ± 0.5 s. Relative to the slugishness of the HRF (in particular, the infant HRF),

this inaccuracy is negligible. For all other scan types, infants lay passively in the scanner. No sedatives were used at any stage of this study.

3.2.3 MRI data acquisition

All data were collected on a 3T Siemens Prisma with an adult 32 channel receive coil.

The structural data acquisition was: T2-weighted, TSE (factor 11), 150° flip angle, TE = 89 ms, TR = 14,740 ms, parallel imaging GRAPPA 3, 192×192 in-plane matrix size, 126 slices, 1 mm isotropic voxels, and 2 mins 13 s acquisition time.

The fieldmap data acquisition was: gradient echo, 2DFT readout, dual echo TE1/TE2 = 4.92/7.38 ms, TR = 550 ms, 46° flip angle, 90 × 90 in-plane matrix size, 56 slices, 2 mm isotropic voxels, and 1 min 40 s acquisition time.

The noxious-response fMRI data acquisition was: T2* BOLD-weighted, gradient echo, EPI readout, 70° flip angle, TE = 50 ms (Goksan et al., 2017), TR = 1,300 ms, multiband 4 (Moeller et al., 2010; Xu et al., 2013), 90 × 90 in-plane matrix size, 56 slices, 2 mm isotropic voxels, AP phase encode direction, a single-band reference (SBref) image was acquired at the start, and the mean acquisition time was approximately 6 min (approximately 277 volumes).

3.2.4 Overview of preprocessing pipelines

In this study, key aspects of two fMRI preprocessing pipelines were compared. The first pipeline was based primarily on the FMRIB Software Library (FSL) FMRI Expert Analysis Tool (FEAT) (Jenkinson et al., 2012), closely matching the FEAT-based pipeline used in a previous infant noxious-response fMRI study (Goksan et al., 2015). This pipeline closely corresponds to what can be called a 'typical' adult preprocessing pipeline with certain adaptations for neonatal data. In this study, this FEAT-based neonatal pipeline is referred to as 'the FEAT pipeline'. The second pipeline was based primarily on the dHCP fMRI preprocessing pipeline (Fitzgibbon et al., 2019) with extensions required for small sample sizes and the modelling of noxious-response fMRI data (as opposed to the much larger resting-state dHCP dataset with which the dHCP preprocessing pipeline was being developed). In this study, this dHCP-based neonatal pipeline is referred to as 'the dHCP pipeline'. The

general workflow of these pipelines is detailed in Figure 3.1. The preprocessing steps have been conceptually split into 'artefact removal' and 'spatial normalisation' subsections, and do not match the actual temporal sequence of events. Details of these pipelines are given next, and details of comparative statistical analyses and non-pipeline-comparison analyses follow below.

FEAT fMRI preprocessing pipelines	
'Typical' adult pipeline	**Modified infant pipeline**
Artefact removal:	
Motion correction (MCFLIRT)	Motion correction (MCFLIRT)
Distortion correction (FLIRT-BBR and FUGUE)	Distortion correction (FUGUE)
Brain extraction (BET)	Brain extraction (BET)
High-pass temporal filtering	High-pass temporal filtering
	Denoising (MELODIC and FIX) *
Low-pass spatial filtering (SUSAN)	Low-pass spatial filtering (SUSAN)
Grand mean scaling	Grand mean scaling
Spatial normalisation:	
Func to SBref using 6 DOF (FLIRT)	Func to Struct using 6 DOF (FLIRT)
SBref to Struct using BBR (FLIRT)	
Struct to MNI152 standard template (FNIRT)	Struct to age-matched standard template (FNIRT)
	Age-matched standard template to 40 week standard template (FNIRT) *

dHCP fMRI preprocessing pipelines	
Extended infant pipeline	**'Typical' infant pipeline**
Artefact removal:	
Motion and distortion correction (EDDY) Distortion correction uses Gradient Echo Dual Echo derived field maps (fsl_prepare_fieldmap)	Motion and distortion correction (EDDY) Distortion correction uses Spin Echo Blip-up/Blip-down derived field maps (TOPUP)
High-pass temporal filtering 100 s filter for stimulus-based data	High-pass temporal filtering 150 s filter for resting-state data
Denoising (MELODIC and FIX)	Denoising (MELODIC and FIX)
Brain extraction (BET)	Brain extraction (BET)
Low-pass spatial filtering (SUSAN)	
Grand mean scaling	
Spatial normalisation:	
Func to SBref using 6 DOF (FLIRT)	Func to SBref using 6 DOF (FLIRT)
SBref to Struct using BBR (FLIRT)	SBref to Struct using BBR (FLIRT)
Struct to age-matched standard template (SyN)	Struct to age-matched standard template (SyN)
Age-matched standard template to 40 week standard template (SyN)	Age-matched standard template to 40 week standard template (SyN)

GLM analysis
Subject-level:
Pre-whitening, infant double gamma HRF fit (FEAT)
Group-level:
Group mean activity with five nuisance EVs; 10,000 permutations; variance smoothing, voxel/cluster/TFCE thresholding, 5% FDR/FWER/FWER correction (Randomise)

Figure 3.1: FMRI analysis pipeline workflow and comparison highlighting key pipeline differences. "FEAT fMRI preprocessing pipelines" box: the "'Typical' adult pipeline" column is the hypothetical FEAT fMRI preprocessing pipeline using analysis tools and steps currently found in FEAT, appropriate for analysis of an adult multiband fMRI dataset; the "Modified infant pipeline" is the modification of this 'typical' FEAT pipeline for use in our infant multiband fMRI dataset (* = preprocessing steps performed partially or fully external to FEAT). "dHCP fMRI preprocessing pipelines" box: the "'Typical' infant pipeline" column is a condensed description of some key analysis steps in the dHCP fMRI preprocessing pipeline; the "Extended infant pipeline" extends the dHCP pipeline for use with the current noxious-response fMRI dataset. "GLM analysis" box: this is a condensed description of the key GLM analysis steps performed at the subject- and group-level in this chapter, and is common to the outputs of both the modified FEAT and extended dHCP fMRI preprocessing pipelines. See main text for expansion of abbreviations.

The FEAT preprocessing pipeline

The structural data were brain extracted using FSL's BET (Brain Extraction Tool) (Smith, 2002), and bias field corrected within FEAT using FSL's FAST (FMRIB's Automated Segmentation Tool) (Zhang, Brady, and Smith, 2001). Optimisation of BET parameters (optimal fractional intensity threshold and its vertical gradient) required manual specification per subject. The fieldmap data were processed using fsl_prepare_fieldmap.

For functional data preprocessing, the artefact removal stage consisted of seven core processes (see Figure 3.1). Motion correction was performed using FSL's MCFLIRT (Motion Correction FMRIB's Linear Image Registration Tool) (Jenkinson et al., 2002), which realigns each volume using linear registration, correcting for inter-volume head motion i.e. volume-to-volume motion correction. Distortion correction was performed using FSL's FUGUE (FMRIB's Utility for Geometrically Unwarping EPIs) (Smith et al., 2004), which corrects for susceptibility-induced distortions, with the fieldmap aligned to the reference functional volume selected during motion correction i.e. static distortion correction. The corrected functional data are then brain extracted using BET, and temporally filtered using a 0.01 Hz high-pass filter (100 s period). External to FEAT, the data were then decomposed into spatially independent components, consisting of a spatial map and timeseries, using probabilistic spatial independent component analysis (ICA) using FSL's MELODIC (Multivariate Exploratory Linear Optimised Decomposition into Independent Components) (Beckmann and Smith, 2004). Components were manually labelled as signal or noise using recommended guidelines (Griffanti et al., 2017). The noise component timeseries, and the 24 head motion parameter timeseries estimated during motion correction, were regressed from the data using FSL's FIX (FMRIB's ICA-based Xnoiseifier) (Griffanti et al., 2014; Salimi-Khorshidi et al., 2014). The combination of ICA decomposition and noise timeseries regression is referred to as FIX denoising. This step is not an automated component of FEAT, as the ICA component labelling requires manual intervention, and so, FIX denoising is not a step of a 'typical' adult FEAT preprocessing pipeline. However,

it was included in our infant FEAT pipeline, as it has previously been used in our group (Goksan et al., 2015) to deal with heavily contaminated neonatal fMRI data. Additionally, it is a core component of the dHCP pipeline, and to allow fair comparison of the two pipelines, its inclusion in the FEAT pipeline was required. The final two preprocessing stages are typical of a FEAT pipeline and included low-pass spatial filtering and grand mean scaling. Three spatial filters were tested, as will be detailed below, using FSL's SUSAN (Smoothing over Univalue Segment Assimilating Nucleus) (Smith and Brady, 1997). Finally, the data were grand mean scaled to have a spatiotemporal median value of 10,000, as is default in FEAT.

The spatial normalisation stage consisted of three core steps (see Figure 3.1). An infant's functional data were aligned with the infant's high-resolution structural image using a 6-DOF (degrees of freedom) linear rigid-body registration using FSL's FLIRT (FMRIB's Linear Image Registration Tool) (Jenkinson and Smith, 2001; Jenkinson et al., 2002). In a 'typical' adult FEAT pipeline, this global linear registration between functional and structural images is subsequently refined by aligning the boundary of the cerebral cortical grey matter and white matter using BBR (Boundary Based Registration) (Greve and Fischl, 2009). BBR was omitted from our infant FEAT pipeline for two reasons. First, it requires an accurate delineation of the appropriate tissue boundary, which is not possible using standard FSL tools external to those developed for the dHCP structural analysis pipeline. Second, the BBR parameterisation is hard-coded in FEAT for the cross-boundary intensity gradient of the adult brain, which is the reverse of the intensity gradient in the infant brain. After aligning to structural space, the structural image was aligned to a standard space template (Makropoulos et al., 2016; Serag et al., 2012) that was age-matched to the particular infant e.g. if the infant was 38 weeks PMA, that infant's structural image was aligned to the 38-week infant standard template. This alignment was done using FSL's FNIRT (FMRIB's Non-linear Image Registration Tool) (Andersson, Jenkinson, and Smith, 2007), with minor alterations to the default FNIRT configuration required for infant data (Goksan et al., 2015). The age-matched standard template was finally aligned to the global 40-week PMA standard

template space (Makropoulos et al., 2016; Serag et al., 2012) using FNIRT. All registration transformations between functional and standard space are applied in a single step to minimise data quality degradation due to multiple interpolation steps.

The dHCP preprocessing pipeline

The structural data were processed using the MIRTK Draw-EM (Developing brain Region Annotation With Expectation-Maximization) neonatal pipeline (Makropoulos et al., 2014), which is the tool forming the basis of the dHCP structural preprocessing pipeline (Makropoulos et al., 2018). The 'typical' dHCP fMRI preprocessing pipeline was extended for this study to include the required structural image preprocessing steps (brain extraction, bias field correction, and tissue segmentation) for analysis of functional MRI data. The fieldmap data were processed using a modified version of fsl_prepare_fieldmap.

For functional data preprocessing, the artefact removal stage consisted of six core processes (see Figure 3.1). Motion and distortion correction were performed as a single step using FSL's EDDY (Andersson et al., 2016; Andersson and Sotiropoulos, 2016). This a key dHCP innovation, as EDDY was originally designed for the correction of diffusion MRI data, but has been adapted for use with functional data due to its advanced method of correction. Motion correction realigned each volume and then each slice within a volume, which corrected for both inter- and intra-volume head motion i.e. volume-to-volume and slice-to-volume motion correction (Andersson et al., 2017). Distortion correction adapted the fieldmap image according to the infant's head position, estimating the dynamic changes to the field induced by subject motion. This estimated dynamic distortion correction corrected for standard susceptibility-induced distortions and additional susceptibility-by-movement distortions (Andersson et al., 2018). Unlike the sequential use of MCFLIRT and FUGUE in the FEAT pipeline, EDDY simultaneously corrects for motion and distortion, which is more accurate due to the interactions between the two processes. While these are the theoretical advantages, this was the first study published using EDDY to correct functional data, and thus was directly

compared to the more traditional FEAT pipeline corrections. High-pass temporal filtering was performed using a 0.01 Hz filter followed by FIX denoising as described in the FEAT pipeline. However, due to the dHCP structural preprocessing pipeline, accurate tissue segmentation of an infant's data was possible allowing training of the FIX classifier. FIX was thus trained and the classification accuracy reported. Brain extraction, spatial filtering, and grand mean scaling were performed in the dHCP pipeline in a similar manner to the FEAT pipeline. For the purposes of this study (and overall generalisation of the dHCP pipeline), the 'typical' minimal preprocessing pipeline was extended to include spatial filtering and grand mean scaling.

The spatial normalisation stage consisted of four core steps (see Figure 3.1). An infant's functional data were aligned with the infant's SBref image using a 6-DOF linear rigid-body registration using FSL's FLIRT. The SBref image was aligned with the high-resolution structural image initially using a 6-DOF linear rigid-body registration using FLIRT, refined with BBR using the cerebral grey-white matter cortical boundary. BBR was possible due to the structural image segmentation. The alignment between structural and the 40-week standard space (Makropoulos et al., 2016; Serag et al., 2012) was a two-step process with registration going via the age-appropriate template, similar to the FEAT pipeline. However, these non-linear registrations were performed using ANTs's SyN (Avants et al., 2008).

3.2.5 GLM, ROI, and region-constrained analyses

GLM analyses

The preprocessed functional data output from both the FEAT and dHCP pipelines underwent identical subsequent general linear model (GLM) analyses. The subject-level GLM analysis was performed using FEAT. The infant's experimenter-controlled stimulus application timeseries was convolved with an infant-appropriate haemodynamic response function (HRF) (Arichi et al., 2012). Three HRFs were tested, detailed below. The GLM model was filtered using the 0.01 Hz high-pass temporal filter, and prewhitening correction for temporal autocorrelations was performed using FSL's FILM (FMRIB's Improved Linear Model) (Woolrich et al., 2001).

Two group-level GLM analyses were performed. The first analysis involved getting the group average activation map (specifically, the group t-statistic map), derived from each infants GLM regression parameter map, using a standard whole-brain mass-univariate group analysis. These group t-statistic maps were generated for each analysis pipeline variation to allow a cross-pipeline comparisons. The second analysis involved taking each infant's GLM t-statistic maps generated from the variable iterations of the pipelines, and getting a 'difference-in-t-statistic map', or t-statistic difference maps, per infant by subtracting two t-statistic maps within an infant. These t-statistic difference maps were then entered into a group analyses to test cross-pipeline differences.

For the first GLM analysis, each infant's GLM regression parameter map was entered into a whole-brain group-average GLM analysis. Statistically significant regions of activation were identified using permutation analysis using FSL's Randomise (Winkler et al., 2014). Adjustment for five nuisance variables was applied: PMA, PNA, sex, brain volume (derived from structural segmentation data), and mean head motion (DVARS). 10,000 permutations were run, and a 10 mm FWHM variance smoothing (Holmes et al., 1996) kernel was applied to account for the low degrees of freedom. Due to variable strengths and weaknesses of thresholding methods, three thresholding methods were reported to facilitate interpretation. First, voxel-based thresholding with a 5% false discovery rate (FDR) correction for multiple testing. Using a familywise error rate (FWER) correction for voxel-based whole-brain analysis was too conservative, resulting in many 'empty' thresholded maps. Second, cluster-mass-based thresholding (Bullmore et al., 1999) with a cluster-defining threshold of 2.3 and a 5% FWER correction. Third, threshold-free cluster enhancement (TFCE)-based thresholding (Smith and Nichols, 2009) with default parameters and a 5% FWER correction.

For the second GLM analysis, each infant's GLM t-statistic difference map was entered into a region-constrained (see below) group-average GLM analysis. Identical permutation analysis as above was performed, including adjustment for nuisance variables. However, due to the spatially discontinuous non-whole-brain analysis,

spatial statistic thresholding methods (cluster-mass and TFCE) were not used. Voxel-based thresholding was used with a 5% FWER correction for multiple testing.

ROI and region-constrained analyses

Four grey matter and one white matter regions-of-interest (ROIs) were defined for the ROI and region-constrained analyses. The grey matter regions were selected to assess analysis outputs in brain regions functionally relevant to the stimulus paradigm, while the white matter region was used to identify noxious-response activity incorrectly localised to white matter. These ROIs were defined using a combination of functionally and anatomically defined regions. To define functionally relevant grey matter regions, the FEAT pipeline was used on an independent dataset of 15 neonates (Goksan et al., 2018) to generate the group average activation map in response to the same 128 mN noxious stimulus. Statistically significantly activated regions were identified by thresholding using TFCE and a 5% FWER correction for multiple testing. This network of functionally-defined regions was then masked with four anatomically-defined regions to extract distinct well-defined ROIs. Using the standard brain atlas (Makropoulos et al., 2016), the four bilateral ROIs were the thalamus, insula, postcentral gyrus (PoCG), and anterior cingulate cortex (ACC), which are brain regions known to be involved in processing nociceptive input in both adults and infants (Apkarian et al., 2005; Williams et al., 2015; Goksan et al., 2015). The white matter ROI was defined using the atlas alone, and was used to identify noxious-response activity incorrectly localised to white matter during the spatial normalisation comparison, spatial filtering comparison, and overall pipeline comparison analyses.

Using these ROIs, variability in group GLM t-statistics were compared for the numerous pipeline comparisons. The rationale for comparing t-statistics follows from Smith and colleagues that, at the group-level, it is reasonable to assume that larger t-statistics are better due to reduced noise variance originating from imperfect analysis methods (Smith et al., 2005). To compare the GLM t-statistic results between pipelines, mean and peak t-statistics extracted from the grey matter ROIs were

compared, and proportions of mislocalised active voxels extracted from the white matter ROI were compared. To test for statistical significance, the collection of grey matter ROIs were used to define the constrained region for the region-constrained analyses, when performing group analysis on the t-statistic difference maps.

3.2.6 Motion correction and distortion correction assessment

To compare the motion correction and distortion correction (MCDC) outputs of the FEAT and dHCP pipelines, the post-MCDC correction mean DVARS motion metric and whole-brain temporal SNR (tSNR) measures were compared using a Wilcoxon signed-rank test with a significance level $\alpha = 0.05$. To assess the effects of the two MCDC pipelines on sensitivity to signal within specific ROIs, the mean and peak t-statistics within the four individual grey matter ROIs were compared, and whole-brain active voxel counts were compared for each of the three thresholding methods. To compare the pipeline outputs from the GLM region-constrained analysis, group analysis on the t-statistic difference maps were performed. The GLM ROI and region-constrained analyses required registration to standard space. In order to disambiguate the MCDC effects from the spatial normalisation effects, both the FEAT and dHCP outputs were registered to standard space using the FEAT-style registrations. Finally, TFCE thresholded group average activation maps were also presented for qualitative assessment.

3.2.7 Spatial normalisation assessment

To compare the spatial normalisation outputs of the FEAT and dHCP pipelines, alignment of the functional data with the standard space template was assessed using normalised mutual information (NMI). Using the cerebral cortical boundary between grey and white matter of the standard template, the mean intensity gradient across this boundary, the BBR cost function, was also compared. This metric is referred to as the boundary intensity difference. While BBR was used to align the functional and structural images, BBR is not used to align the structural and standard template images, which would otherwise make the boundary intensity

difference metric biased. The NMI and boundary intensity difference measures were compared using a Wilcoxon signed-rank test with a significance level $\alpha = 0.05$. To assess the effects of the two spatial normalisation pipelines on sensitivity to signal within specific ROIs, the mean and peak t-statistics within the four individual grey matter ROIs were compared, and whole-brain active voxel counts were compared for each of the three thresholding methods. To assess the pipelines' effects on spatial specificity (correctly localising active voxels to grey matter), the proportions of significantly activated voxels incorrectly localised to white matter were compared. Due to the different spatial normalisation pipelines resulting in different alignment of data to standard space, no GLM region-constrained t-statistic difference map analysis was performed. In order to disambiguate spatial normalisation effects from MCDC effects, the dHCP MCDC outputs were used as the inputs to both spatial normalisation pipelines. Finally, TFCE thresholded group average activation maps were also presented for qualitative assessment.

3.2.8 FIX denoising assessment

To ensure equally accurate FIX denoising between pipelines, the following approach was adopted. The trained FIX model was used in both pipelines initially only to label components. All components from both pipelines were then manually inspected, and misclassified components relabelled where necessary. Thus, FIX was used in both pipelines to support semi-automatic classification of ICA components. Notably, the FIX model, which was trained on the dHCP data, worked equally well at classifying components in the FEAT pipeline; the small number of misclassified components was comparable between pipelines, and no obvious difference in the nature of the components was discernible to the researchers. Once manual inspection and correction of ICA components was complete, FIX was then used for the final clean-up step (using the corrected manual classification labellings) to remove these noise component timeseries and extended head motion parameter timeseries (24 motion timeseries) from the data.

To assess the influence of FIX denoising in neonatal fMRI data preprocessing, GLM outputs of data with and without FIX denoising were compared. The dHCP pipeline MCDC and spatial normalisation pipelines were used for this FIX denoising analysis, to reduce the influence of motion, distortion, and registration misalignments on this FIX denoising analysis. This was due to the observed superiority of the dHCP pipeline over the FEAT pipeline in these aspects of preprocessing (from our MCDC and spatial normalisation comparisons above). From the GLM ROI analyses, the mean and peak t-statistics within the four individual grey matter ROIs were compared, and whole-brain active voxel counts were compared for each of the three thresholding methods. From the GLM region-constrained analysis, group analysis on the t-statistic difference maps were performed. Finally, TFCE thresholded group average activation maps were also presented for qualitative assessment.

3.2.9 Spatial filtering assessment

Spatial filtering has the benefit of increasing SNR coupled with the reduction of spatial resolution. As such, minimal spatial filter extents are preferable, especially for the small brain volumes of newborn infants. In both the FEAT and dHCP pipelines, FSL's SUSAN was used to perform filtering. In order for SUSAN to work effectively, the spatial filter extent (specified as FWHM) must be larger than the data voxel size. To perform minimal spatial filtering, a spatial filter extent equal to 1.5 times the voxel size was selected as default for all pipeline comparison analyses, excluding this spatial filtering comparison analysis. For the specific voxel size of this dataset, this default filter was 3 mm FWHM (1.5×2 mm isotropic voxels). For the spatial filtering comparison analysis, the 3 mm filter was compared to a 0 mm filter (no filtering) and 5 mm filter. This larger 5 mm filter was selected for three reasons. First, it is the default filter in FEAT, and according to a 2012 review (Carp, 2012), over 80% of adult fMRI studies use a filter of 5 mm or greater. Second, a 5 mm kernel has been previously used in preprocessing infant noxious-response fMRI data (Goksan et al., 2015). Third, it is currently not uncommon in the infant

fMRI literature at large to see filters of 5 mm or larger in both resting-state (Mitra et al., 2017) and stimulus-response (Scheef et al., 2017) studies.

To compare the effect of spatial filter extent on sensitivity to signal in specific ROIs, the mean and peak t-statistics within the four individual grey matter ROIs were compared, and whole-brain active voxel counts were compared for each of the three thresholding methods. To assess the effect of spatial filtering on spatial specificity, the proportions of significantly activated voxels incorrectly localised to the white matter ROI were compared. From the GLM region-constrained analysis, group analysis on the t-statistic difference maps were performed between the 0 mm and 3 mm filters as well as between the 3 mm and 5 mm filters. Finally, TFCE thresholded group average activation maps were also presented for qualitative assessment.

3.2.10 HRF modelling assessment

Due to the differences between the adult and infant cerebral haemodynamic responses, HRF models were based on neonate-specific parameterisations (Arichi et al., 2012). Three HRF models were tested. The first was a single gamma (SG) HRF, with a baseline-to-peak latency of 7 s and no undershoot, defined within FEAT, which has been used in a previous publication (Goksan et al., 2015). The second was a double gamma (DG) HRF, with a baseline-to-peak latency of 7 s and an overshoot-to-undershoot ratio of 1:1, derived from data kindly provided by Dr Tomoki Arichi (Arichi et al., 2012). The third was a three basis function model developed from neonatal fMRI data by Arichi and colleagues using FSL's FLOBS (FMRIB's Linear Optimal Basis Set) (Woolrich, Behrens, and Smith, 2004). Due to the method of deriving the HRF basis set, the first basis function very closely corresponds to the DG HRF: it has a baseline-to-peak latency of 7 s and an overshoot-to-undershoot ratio of 1:1, but is a principle component analysis (PCA) timeseries as opposed to a double-gamma function. To generate noxious-response activation maps output from subject-level GLM analysis, a HRF-fit size statistic map must be defined. The SG and DG HRF models involve only a single function, and so, the size statistic map is simply the regression parameter map output

automatically by FEAT. For the FLOBS HRF, the size statistic was defined as the 2-norm of the linear combination of the three basis functions. The sign (positive or negative) of this size statistic was taken from that of the first basis function i.e. the function closely resembling the neonatal DG HRF waveform.

To compare HRF model fits, peristimulus timeseries plots were derived for each model in two grey matter ROIs: the PoCG due to its robust signal in all three model fits, and the thalamus due to its high variability across model fits. The peristimulus timeseries plots were derived as the average timeseries across voxels and across all 10 trials. The timeseries were 26 s in length, with t = 0 s being the point of stimulus delivery. The minimum inter-stimulus interval was 25 s, and this 26 s timeseries spanned 20 time-points (20 volumes). Individual subject timeseries for the raw data and HRF fits are provided. Additionally, timeseries for the volumewise temporal cross-subject mean and temporal cross-subject standard deviations are provided. These temporal mean and standard deviation timeseries comparisons allow assessment of group-average HRF amplitude and variability independently of the choice of HRF model size statistic. Baseline-to-peak latency and upshoot-to-undershoot ratio of the group-mean HRF timeseries were compared to that of the raw data. To get robust baseline-to-peak latency and upshoot-to-undershoot ratio estimates from the raw data, the Woody average of the subject timeseries was used to account for minor temporal jitter. The Woody average (Woody, 1967) improves alignment by maximizing the cross-correlation between individual timeseries and the average timeseries in an iterative manner. Jitter was restricted to ± 2 volumes, was restricted to integer shifts, and a maximum of 100 iterations, although convergence occurred after two iterations. Subsequently, a DG function was closely fit to this mean timeseries, with DG fit performed using MATLAB fit function optimisation.

To assess HRF model sensitivity to signal, the mean and peak t-statistics within the four individual grey matter ROIs were compared, and whole-brain active voxel counts were compared for each of the three thresholding methods. From the GLM region-constrained analysis, group analysis on the t-statistic difference maps were

performed. Finally, TFCE thresholded group average activation maps were also presented for qualitative assessment.

3.2.11 Overall comparison of preprocessing pipelines

Based on the results of the spatial filtering and HRF modelling analysis, the complete FEAT and dHCP pipelines were compared using a 3 mm FWHM spatial filter and the DG HRF model. To assess the pipelines' sensitivities to signal, the mean and peak t-statistics within the four individual grey matter ROIs were compared, and whole-brain active voxel counts were compared for each of the three thresholding methods. Additionally, the pipelines' spatial specificities were compared by assessing the proportions of significantly activated voxels incorrectly localised to the white matter ROI. From the GLM region-constrained analysis, group analysis on the t-statistic difference maps were performed. Finally, TFCE thresholded group average activation maps were also presented for qualitative assessment.

3.3 Results

3.3.1 Motion correction and distortion correction assessment

The dHCP pipeline corrected slice-to-volume and susceptibility-by-movement artefacts that were not corrected by the FEAT pipeline (Figure 3.2). The dHCP estimated dynamic fieldmap approach provided greatest noise reduction in frontal and occipital polar regions. Motion-related noise remaining after MCDC was assessed using DVARS and tSNR (Figure 3.3). At the subject-level, the dHCP MCDC resulted in a greater reduction in motion-related signal variance, especially during large movements (Figure 3.3A). At the group-level, the dHCP MCDC pipeline significantly reduced the mean DVARS values ($p = 1.8 \times 10^{-4}$) and increased the mean tSNR values ($p = 6.1 \times 10^{-5}$) across subjects (Figure 3.3B and C).

The dHCP MCDC resulted in a modest increase in the mean group-level t-statistic in all grey matter ROIs (Table 3.1). Comparing the number of significantly active voxels, all three thresholding approaches demonstrated increased number of

active voxels using the dHCP pipeline (Table 3.1). Using the thresholded activity maps to qualitatively compare pipelines, increased significant activity was seen in both cortical and subcortical regions, with greatest improvements in the thalamus (Figure 3.4). Together, these results suggest the dHCP MCDC pipeline increased sensitivity to signal by reducing noise variance. Comparing the subject-level t-statistic difference maps, changes were not statistically significant.

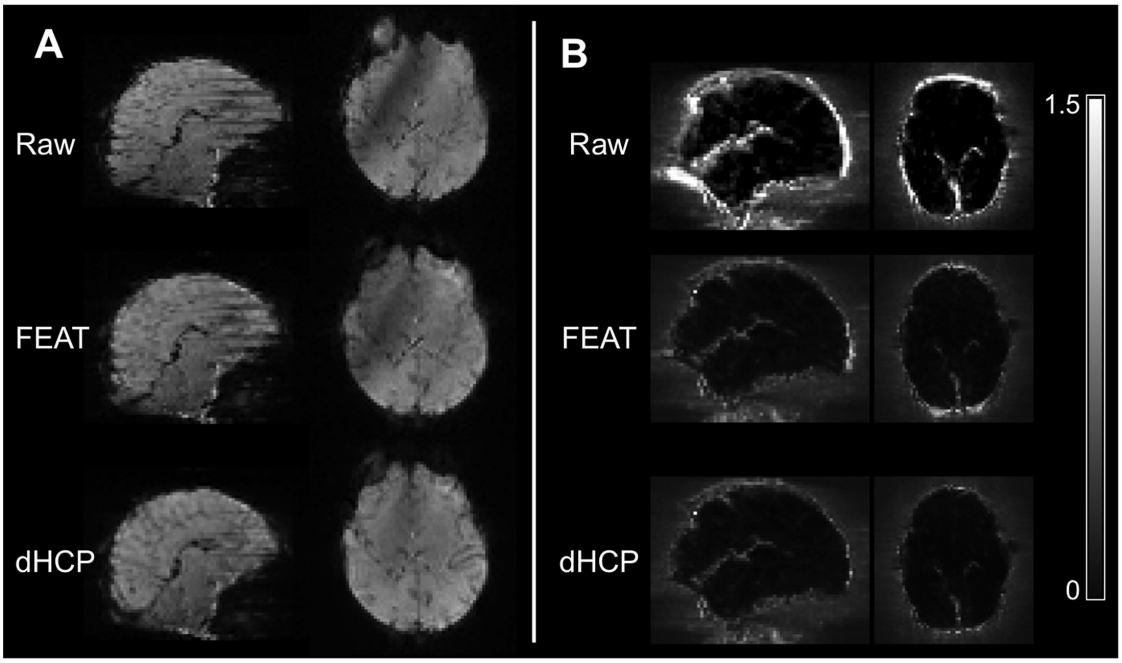

Figure 3.2: Comparison of the effects of MCDC between the FEAT and dHCP pipelines in representative subjects. (A) Comparison of volume-to-volume and slice-to-volume motion correction. Each row is the same volume from one subject before correction (Raw), after volume-to-volume motion correction used in the FEAT pipeline (FEAT), and after both volume-to-volume and slice-to-volume motion correction used in the dHCP pipeline (dHCP). (B) Comparison of static distortion correction and estimated dynamic distortion correction. Each row is a coefficient of variation map (i.e. ratio of temporal standard deviation to temporal mean) from one subject before correction (Raw), after static distortion correction used in the FEAT pipeline (FEAT), and after estimated dynamic distortion correction used in the dHCP pipeline (dHCP). Improvements are seen predominantly at the brain surface perpendicular to the phase encode directions, especially in frontal and occipital polar regions.

3.3.2 Spatial normalisation assessment

Improvements in spatial normalisation were apparent in individual subjects when using the dHCP pipeline (Figure 3.5). Greatest improvements were visible at the

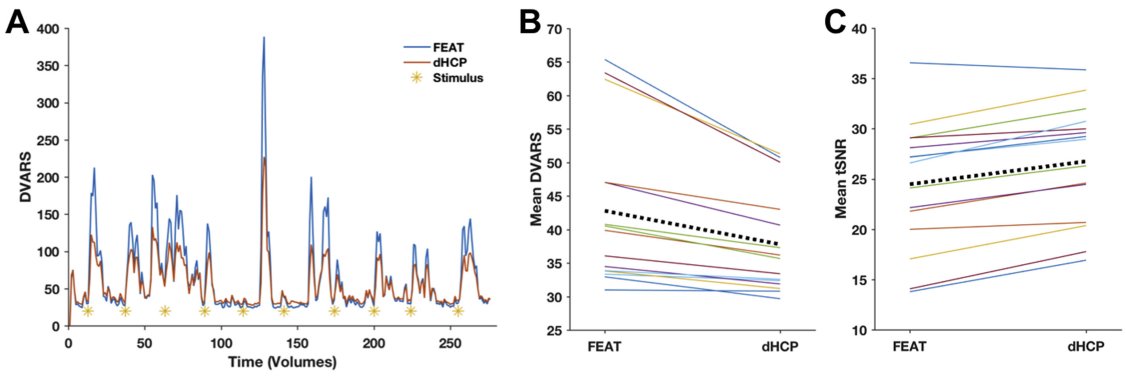

Figure 3.3: Comparison of effects of MCDC between FEAT and dHCP pipelines using DVARS and tSNR. (A) Comparison of DVARS plots post-correction for a representative subject. The dHCP pipeline corrections result in lower mean DVARS metric across the entire session with greatest effects seen during large head motions. The yellow asterisks indicate the time of stimulus delivery, demonstrating the presence of stimulus-correlated motion in this subject. (B–C) Comparison of MCDC effects between FEAT and dHCP pipelines across all 15 subjects using DVARS and tSNR. For each plot, solid coloured lines are individual subjects and the dotted black line is the group average. The dHCP MCDC results in statistically significantly lower DVARS and higher tSNR, indicating better artefact correction.

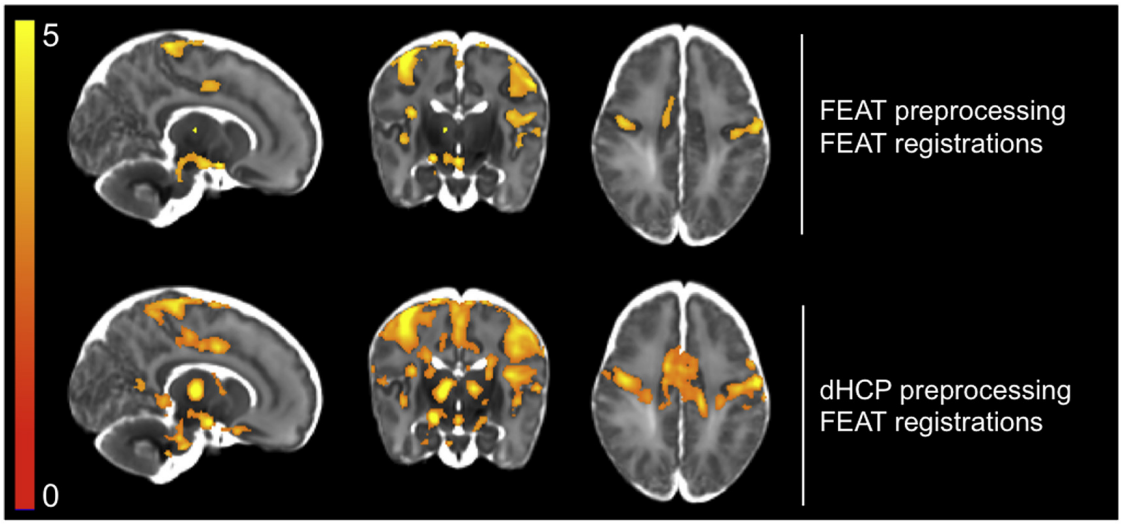

Figure 3.4: Comparison of MCDC between FEAT and dHCP pipelines using the group-level GLM output for all 15 subjects. After thresholding the group t-statistic maps (TFCE, default parameters, 5% FWER corrected), the dHCP pipeline resulted in greater sensitivity to signal in both cortical and subcortical regions. Using the dHCP pipeline, bilateral thalamic activity was detected. Using the FEAT pipeline, only contralateral thalamic activity was detected.

Table 3.1: Comparison of MCDC between FEAT and dHCP pipelines using t-statistics from the GLM output and significantly activated voxel counts for all 15 subjects. T-statistics: using the dHCP pipeline, all grey matter ROIs had an increase in mean t-statistic. The maximum t-statistic (in parentheses) increased in all regions using the dHCP pipeline, except the PoCG. # active voxels: the dHCP pipeline also resulted in an increase in the number of significantly active voxels using all three thresholding approaches. Voxel is FDR corrected; Cluster and TFCE are FWER corrected. ACC = anterior cingulate cortex; PoCG = postcentral gyrus; TFCE = threshold free cluster enhancement.

		FEAT	**dHCP**
T-statistics	ACC	1.529 (3.518)	2.012 (4.581)
	Insula	2.651 (5.244)	2.682 (5.422)
	PoCG	2.822 (6.568)	3.131 (6.123)
	Thalamus	2.168 (6.197)	2.542 (6.368)
# active voxels	Voxel	0	41,492
	Cluster	60,175	62,224
	TFCE	16,823	53,926

brain/non-brain surface. The FEAT pipeline registrations frequently incorrectly aligned the functional image cortical GM-CSF boundary with the template image cortical GM-WM boundary, unlike the dHCP pipeline. The dHCP registrations also produced marked improvements in cerebellum and brainstem alignment. Other non-surface improvements were visible but less consistent across subjects. To quantify these differences in spatial normalisation, we compared the alignment between the functional image in standard space and the standard template image using NMI and the boundary intensity difference (Figure 3.6). Using both metrics, the dHCP registrations resulted in a statistically significant improvement in spatial normalisation ($p = 1.2 \times 10^{-4}$ for NMI, $p = 6.1 \times 10^{-5}$ for boundary intensity difference).

The dHCP spatial normalisation resulted in a modest increase in the mean group-level t-statistic in all grey matter ROIs, except the PoCG (Table 3.2). Comparing the number of significantly active voxels and the proportion of significant activity incorrectly localised to white matter, all three thresholding approaches demonstrated increased number of active voxels and decreased proportion of mislocalised activity using the dHCP pipeline (Table 3.2). Using the thresholded activity maps to qualitatively compare pipelines, the dHCP pipeline provided greater

spatial specificity, limiting the regions of significant activity to the grey matter more accurately than the FEAT pipeline (Figure 3.7). Together, these results suggest the dHCP spatial normalisation pipeline increased spatial specificity and sensitivity to signal by reducing subject-template and subject-subject misalignment errors. The subject-level t-statistics between pipelines were however not statistically significantly different.

Table 3.2: Comparison of spatial normalisation between FEAT and dHCP pipelines using t-statistics from the GLM output and significantly activated voxel counts for all 15 subjects. T-statistics: using the dHCP pipeline, there was an increase in the mean t-statistic in all grey matter ROIs, except the PoCG (maximum t-statistic in parentheses). # active voxels: the dHCP pipeline resulted in greater sensitivity to signal, measured as increased number of active voxels using all three thresholding approaches. % white matter: the dHCP pipeline resulted in greater spatial specificity, measured as decreased percent of active voxels mislocalised to white matter using all three thresholding approaches. Voxel is FDR corrected; Cluster and TFCE are FWER corrected. ACC = anterior cingulate cortex; PoCG = postcentral gyrus; TFCE = threshold free cluster enhancement.

		FEAT	**dHCP**
T-statistics	ACC	2.012 (4.581)	2.104 (4.030)
	Insula	2.682 (5.422)	2.977 (5.224)
	PoCG	3.131 (6.123)	2.959 (6.987)
	Thalamus	2.542 (6.368)	2.565 (6.797)
# active voxels	Voxel	41,492	66,984
	Cluster	62,224	82,837
	TFCE	53,926	66,756
% white matter	Voxel	20.835	17.598
	Cluster	9.882	5.400
	TFCE	8.504	4.687

3.3.3 FIX denoising assessment

FIX was manually trained using the 15 subjects' data from the dHCP pipeline. Using the inbuilt leave-one-out cross-validation, accuracy was assessed for automatic denoising of new subjects to have a median True Positive Rate (TPR; percent of signal components correctly identified as signal) of 100%, a median True Negative Rate (TNR; percent of noise components correctly identified as noise) of 95%,

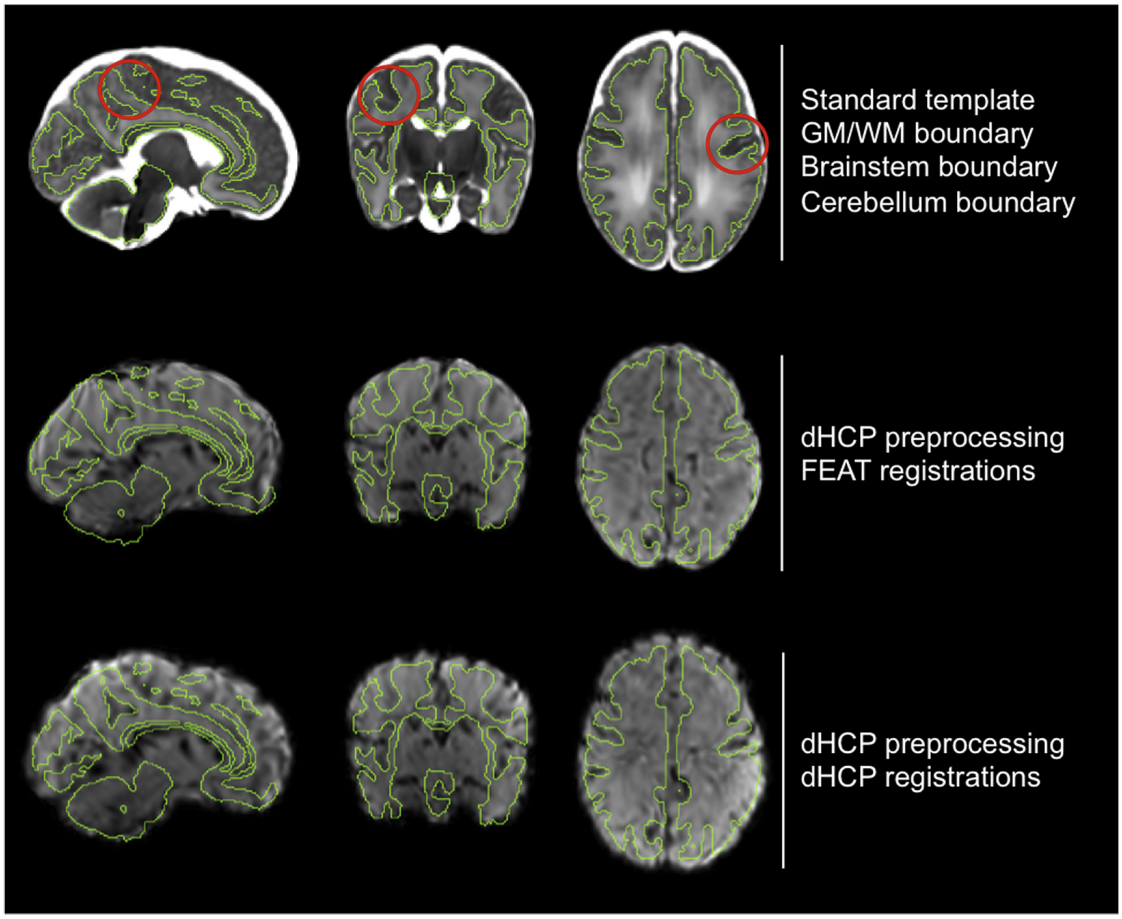

Figure 3.5: Comparison of the differences in spatial normalisation between the FEAT and dHCP pipelines in a representative subject. Top row: the week 40 PMA standard template in grey-scale, with a tissue boundary overlaid in green to aid assessment of registrations. The boundary is taken from the same atlas as the standard template, and includes the GM-WM boundary in the cerebrum, as well as the outer boundary of the brainstem and cerebellum. Second row: the mean functional image registered to standard space using the FEAT pipeline registrations (FSL's FLIRT and FNIRT). Third row: the same mean functional image registered to standard space using the dHCP pipeline registrations (FSL's FLIRT-BBR and ANTs's SyN). In general, the FEAT registrations tend to incorrectly register the GM-CSF boundary of the functional image to the GM-WM boundary of the template, likely due to lack of BBR, and this is corrected in the dHCP result. Also, the cerebellum and brainstem are more accurately aligned with the template in the dHCP result. In this specific subject, several other improvements are visible in the dHCP results, with example regions in each view highlighted with a red circle.

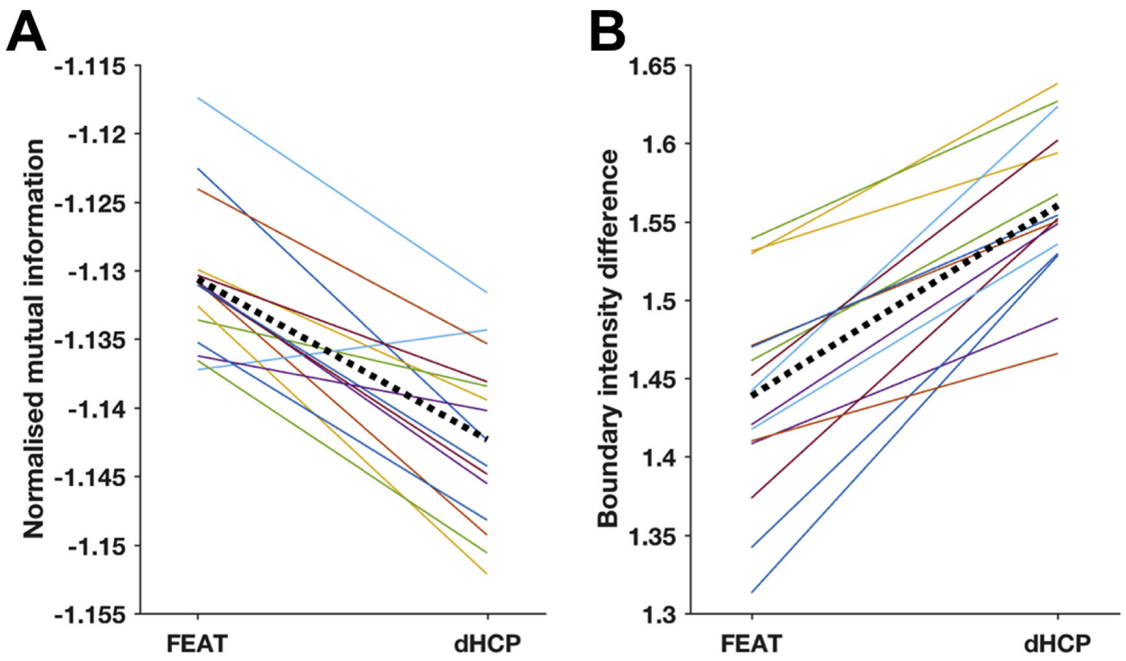

Figure 3.6: Comparison of spatial normalisation between FEAT and dHCP pipelines for all 15 subjects using (A) normalised mutual information and (B) boundary intensity difference. For each plot, solid coloured lines are individual subjects and the dotted black line is the group average. The dHCP spatial normalisation results in statistically significantly larger magnitudes for both alignment metrics.

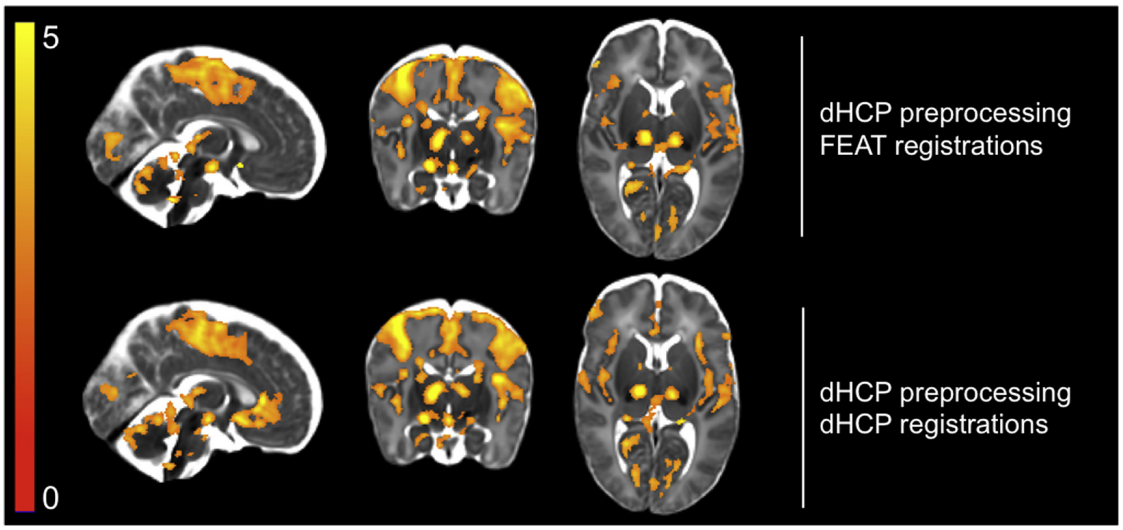

Figure 3.7: Comparison of spatial normalisation between FEAT and dHCP pipelines using the group-level GLM output for all 15 subjects. The significant activity in the thresholded t-statistic maps (TFCE, default parameters, 5% FWER corrected) more faithfully aligned with grey matter structures after applying the dHCP registration transformations. There was also better separation of activity in physically proximal brain regions that are separated by white matter, which should be devoid of activity e.g. the grey matter of the central sulcus and insular cortex.

and an overall measure of median accuracy of 98.6% (accuracy defined in FIX as $(3 \times TPR + TNR)/4$ (for details, see Salimi-Khorshidi et al., 2014). Using the dHCP pipeline, the effects of FIX denoising (using the semi-automatic classification approach outlined in the methods) can be seen at both the subject-level (Figure 3.8) and group-level (Figure 3.9), demonstrating a dramatic improvement in sensitivity to signal. Particularly in subjects with strong stimulus-correlated motion, BOLD and motion signal sources were clearly separated using ICA, and noise was successfully removed from the data using FIX (Figure 3.8). Additionally, white matter and CSF signal sources were readily visible as ICA components, negating the need to manually extract signal timeseries from these regions using ROIs to perform ROI-based nuisance regression clean-up. Eight examples of common ICA components identified in this dataset are provided as Supplementary Figures in Appendix A.

Including FIX denoising in the preprocessing pipeline resulted in an increase in the mean group-level t-statistic in all grey matter ROIs (Table 3.3). Comparing the number of significantly active voxels, all three thresholding approaches demonstrated dramatically increased number of active voxels using FIX denoising (Table 3.3). Using the thresholded group activity maps to qualitatively compare pipelines, an increase in significant activity was seen globally (Figure 3.9). Together, these results suggest FIX denoising dramatically increases sensitivity to signal by reducing noise variance. Comparing the differences in subject-level t-statistics, these differences were statistically significant (voxel-based thresholding, 5% FWER corrected), localised to all grey matter ROIs, except the ACC.

3.3.4 Spatial filtering assessment

Changing the spatial filter extent shifted the balance between sensitivity to signal and spatial specificity. Comparing the ROI analysis results, a consistent decrease in maximum t-statistic was observed across all grey matter ROIs as filter extent increased (Table 3.4, values in parentheses). Comparing the region-constrained analysis results, a statistically significant (voxel-based thresholding, 5% FWER correction) increase in t-statistics was observed across all grey matter ROIs as

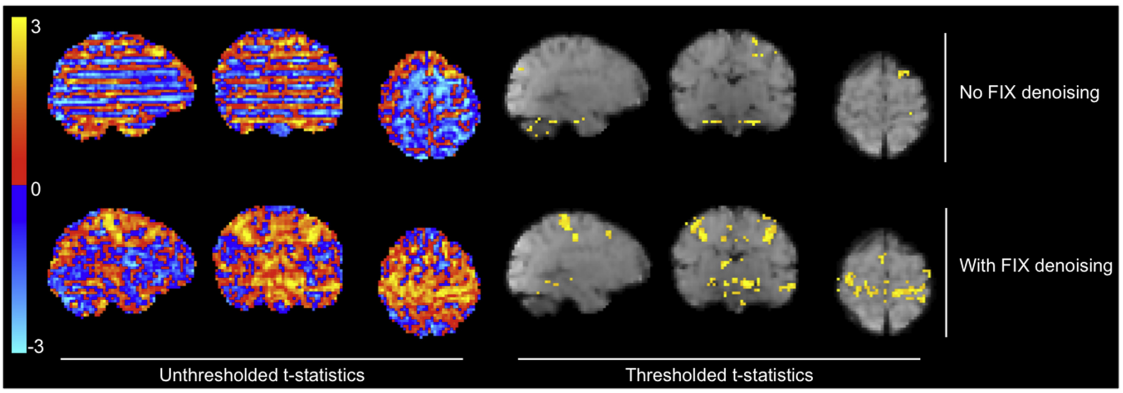

Figure 3.8: Visualization of the effect of FIX denoising using the subject-level GLM output for a representative subject with strong stimulus-correlated motion. Thresholding and correction for multiple comparisons of the t-statistic map was achieved using Gaussian random field theory cluster-based thresholding with a cluster-defining threshold of 2.3 and a 5% FWER correction. Top row: not using FIX denoising resulted in strong motion and striped multiband artefacts dominating the unthresholded t-statistic image, resulting in very poor sensitivity to signal in the thresholded image. Bottom row: using FIX denoising resulted in greatly reduced noise contamination of the unthresholded t-statistic image, and a significant improvement in sensitivity to signal in the thresholded image. Notable for this subject is the presence of stimulus-correlated motion; see Figure 3.3A to see stimulus and head motion timings for this subject. FIX denoising allowed separation of sources of BOLD signal from motion artefacts successfully removing noise while retaining the signal of interest.

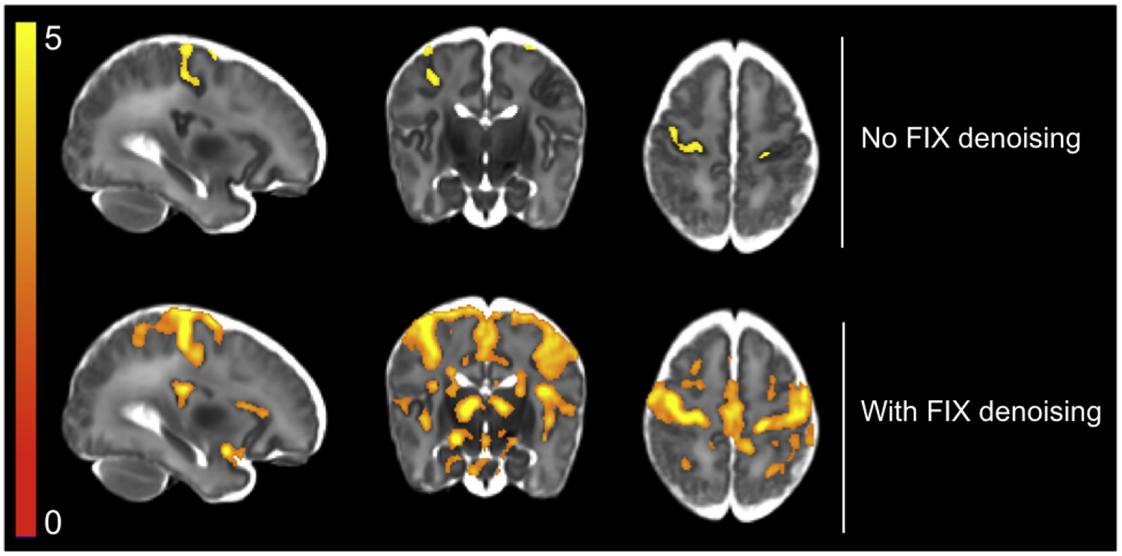

Figure 3.9: Comparison of the effects of FIX denoising on the group-level GLM output for all 15 subjects. Comparing thresholded group t-statistic maps (TFCE, default parameters, 5% FWER corrected), there is a dramatic increase in sensitivity to signal, in both cortical and subcortical regions, when using FIX denoising compared to no FIX denoising. Note, this FIX denoising comparison was assessed using dHCP pipeline outputs only.

Table 3.3: Comparison of the effects of FIX denoising on t-statistics from the GLM output and significantly active voxel counts for all 15 subjects. T-statistics: using FIX denoising resulted in an increase in the mean t-statistic in all grey matter ROIs (maximum t-statistic in parentheses). * = ROIs in which FIX denoising resulted in statistically significant increases in t-statistics (voxel-based thresholding, 5% FWER corrected). # active voxels: using FIX denoising also resulted in an increase in the number of significantly active voxels using all three thresholding approaches. Voxel is FDR corrected; Cluster and TFCE are FWER corrected. ACC = anterior cingulate cortex; PoCG = postcentral gyrus; TFCE = threshold free cluster enhancement.

		No FIX	**With FIX**
T-statistics	ACC	2.065 (4.822)	2.104 (4.030)
	Insula	2.199 (4.839)	2.977* (5.224)
	PoCG	2.249 (5.329)	2.959* (6.987)
	Thalamus	1.527 (7.387)	2.565* (6.797)
# active voxels	Voxel	0	66,984
	Cluster	30,837	82,837
	TFCE	1,311	66,756

filter extent increased (Table 3.4; * and † = statistically significant difference; * = 3 > 0 mm, † = 5 > 3 mm). The group-level ROI mean t-statistics results were slightly more variable: there was a consistent increase in mean t-statistic comparing 3 mm filter to no filter, but a 50/50 split in ROIs in which mean t-statistics increased when comparing 3 mm and 5 mm filters. These ROI and region-constrained results suggest that increasing the spatial filter extent tended to increase t-statistics overall by "smearing" activity.

Comparing the number of significantly active voxels and the proportion of significant activity incorrectly localised to white matter, all three thresholding approaches demonstrated increasing number of active voxels and proportion of mislocalised activity with increasing filter extent (Table 3.4). Using the thresholded activity maps to qualitatively compare filter extents, a clear shift in the balance between spatial specificity and sensitivity to signal was visible, consistent with the quantitative t-statistic and voxel count comparisons (Figure 3.10). Without filtering, activations in many regions were tightly localised to grey matter, and as filter extent increased, activations became increasingly blurred across functionally distinct brain areas. However, not filtering resulted in a thresholded activity map

lacking several functionally important regions, such as ipsilateral thalamus, ACC, hippocampus, brainstem, and cerebellum. The minimal filter extent of 3 mm (1.5 times voxel size) appeared to be a reasonable compromise between gaining sensitivity to signal at the cost of spatial specificity.

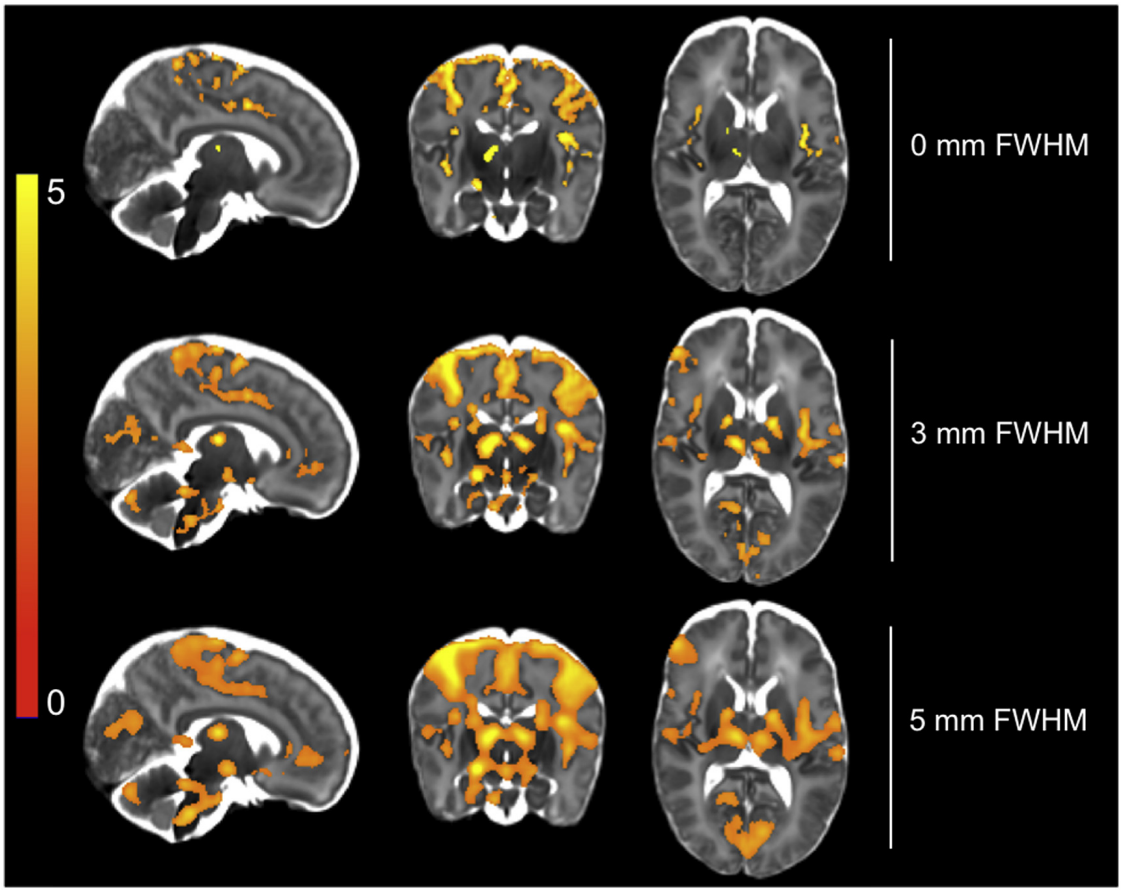

Figure 3.10: Comparison of the effects of spatial filter extent on the group-level GLM output for all 15 subjects. Comparing thresholded (TFCE, default parameters, 5% FWER corrected) group t-statistic maps, as spatial filter extent increased, the sensitivity to signal increased in all ROIs, but the spatial specificity decreased. Not using spatial filtering resulted in a lack of sensitivity to signal in both cortical and subcortical structures. Using the 5 mm FWHM kernel, significant activity was incorrectly localised to non-grey matter regions, and distinct clusters of activity fused into massive clusters spanning several brain regions. Note, this spatial filtering comparison was assessed using dHCP pipeline outputs only.

3.3.5 HRF modelling assessment

The effects of three HRF models on sensitivity to signal were assessed using the dHCP pipeline results. To explore the causes of variability in HRF model fits,

Table 3.4: Comparison of the effects of spatial filtering on t-statistics from the GLM output and significantly activated voxel counts for all 15 subjects. T-statistics: as spatial filter extent increased, peak t-statistics (in parentheses) decreased and statistically significant (voxel-based thresholding, 5% FWER corrected) increases in t-statistics (* and † symbols) were observed across all grey matter ROIs. The mean t-statistic in all ROIs increased after 3 mm filtering compared to no filtering, and increased in PoCG and thalamus after 5 mm filtering compared to 3 mm filtering. # active voxels and % white matter: as filter extent increased, the number of statistically significant voxels increased, measured using three thresholding approaches. Voxel is FDR corrected; Cluster and TFCE are FWER corrected. ACC = anterior cingulate cortex; PoCG = postcentral gyrus; TFCE = threshold free cluster enhancement; * and † = statistically significant difference.

		0 mm	**3 mm**	**5 mm**
T-statistics	ACC	1.956 (5.392)	2.104* (4.030)	2.066† (3.102)
	Insula	2.674 (6.222)	2.977* (5.224)	2.968† (4.317)
	PoCG	2.423 (8.778)	2.959* (6.987)	3.265† (5.874)
	Thalamus	1.928 (7.467)	2.565* (6.797)	2.927† (5.389)
# active voxels	Voxel	14,484	66,984	140,389
	Cluster	48,332	82,837	109,137
	TFCE	25,226	66,756	93,852
% white matter	Voxel	13.353	17.598	21.309
	Cluster	4.616	5.400	7.192
	TFCE	2.319	4.687	6.148

peristimulus timeseries plots were compared at both the subject and group levels in the PoCG and thalamus ROIs (Figure 3.11). In the raw data, a considerable undershoot was visible (Figure 3.11 Row 4, grey and black plots). Due to this undershoot not being modelled by the SG HRF, the SG was excluded from further peristimulus timeseries comparisons. The FLOBS model appeared to over-fit the data compared to the DG model. At the subject level, the fitted FLOBS HRFs exhibited a wide array of BOLD response shapes, including some biologically unlikely profiles (Figure 3.11 Row 2). The reduced flexibility of the DG HRF appeared to make this model more robust to noise (Figure 3.11 Row 3). To quantify these differences, the baseline-to-peak latency and upshoot-to-undershoot ratio of the group mean DG and FLOBS HRFs were compared to a double gamma function fit to the raw data group average (Figure 3.11 Row 4). In both ROIs, the DG HRF had an upshoot-to-undershoot ratio more closely resembling the raw data. In the PoCG,

the DG HRF had a latency to peak more closely resembling the raw data than the FLOBS HRF, but in the thalamus, this relationship was reversed. Comparing the between-subject variability in HRFs (measured as standard deviation), the FLOBS HRF had noticeably larger variability than the DG HRF, especially during the undershoot component of the BOLD response (Figure 3.11 Row 5). Together these results suggest the SG HRF is an overly simple model that under-fits the BOLD response. The FLOBS model may be too flexible for the level of noise in infant fMRI data, resulting in over-fitting and large between-subject variability. The DG appeared to be a reasonable compromise between under-fitting and over-fitting of the modelled response to the data.

Comparing the group-level ROI analysis results, the DG HRF produced the largest mean t-statistics in all grey matter ROIs compared to both the SG and FLOBS HRFs (Table 3.5). The region-constrained group comparisons of HRF models demonstrated statistically significantly (voxel-based thresholding, 5% FWER corrected) larger t-statistics in the DG HRF compared to the SG, localised to the insula. There were no statistically significant differences in t-statistics between the DG and FLOBS models. Similarly, the DG HRF resulted in the largest number of active voxels compared to both SG and FLOBS models using all three thresholding approaches (Table 3.5). Using the thresholded group activity maps to qualitatively compare HRF models, the DG HRF produced noticeably increased activity in both cortical and subcortical structures (Figure 3.12). Overall, these GLM-based results were in line with our peristimulus timeseries results, and suggested the DG HRF model provides greatest sensitivity to signal compared to both the SG and FLOBS models.

3.3.6 Overall comparison of preprocessing pipelines

Finally, the overall effect of the FEAT and dHCP pipelines on GLM results was compared. To generate the group-level results, both pipelines included FIX denoising, 3 mm FWHM spatial filtering, and the DG HRF. Comparing the mean t-statistics of the ROI analysis, the dHCP pipeline resulted in increased mean t-statistics in all grey

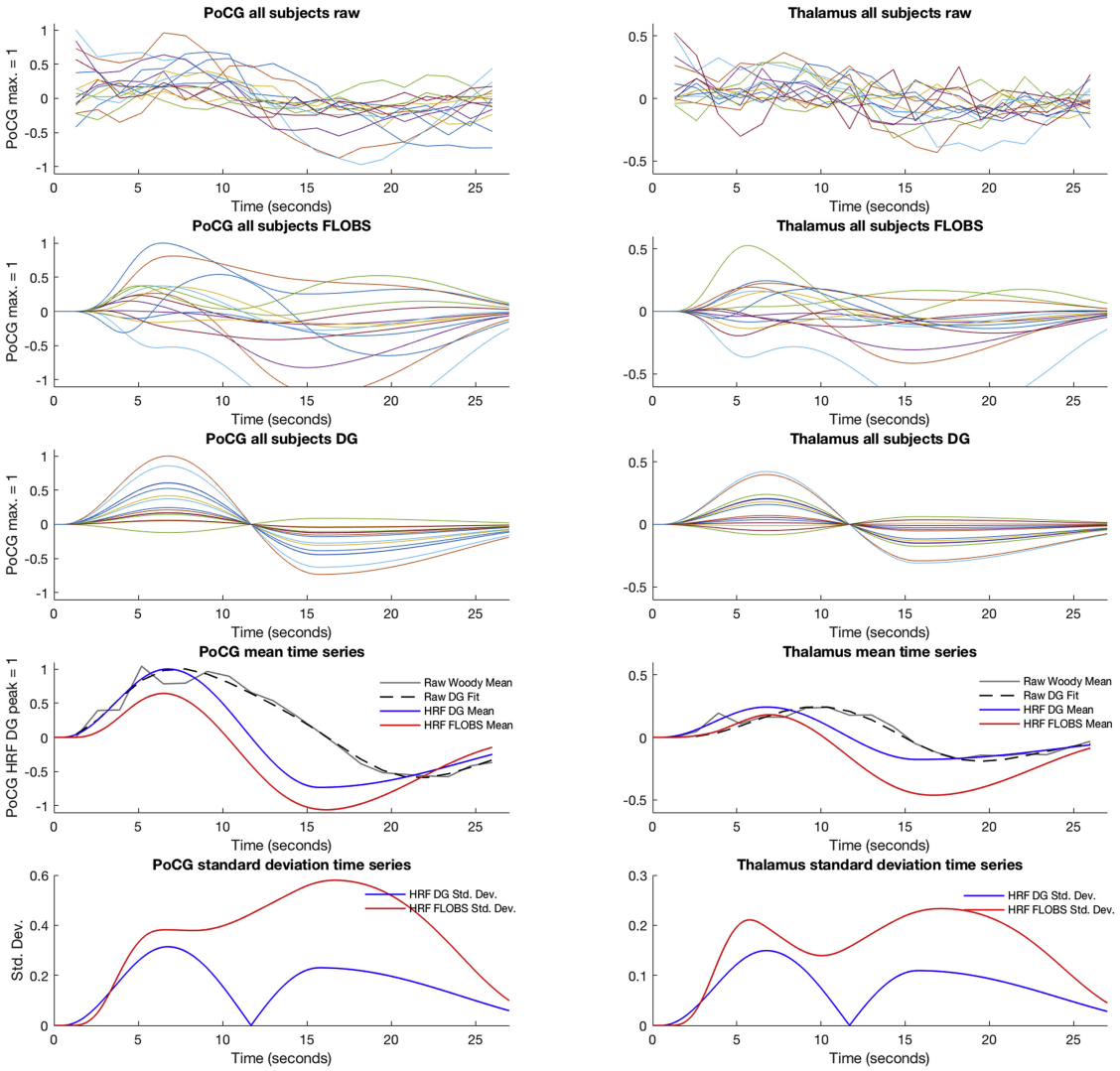

Figure 3.11: Peristimulus timeseries plot analysis for the PoCG ROI (left column) and the Thalamus ROI (right column). For all plots, the x-axis is time in seconds and displays a time window from t = 0 s (time of stimulus delivery) to t = 26 s. The y-axis for rows 1 – 4 is arbitrarily scaled so that the maximum value equals 1. Row 5 y-axis is in arbitrary units. Rows 1 – 3: Peristimulus timeseries plots for all individual subjects, averaged over all voxels in the region and all trials, using the raw data (Row 1), the three basis function (FLOBS) HRF (Row 2), and the double gamma (DG) HRF (Row 3). Row 4: Group mean timeseries plots. The raw data plots are arbitrarily scaled so that the double gamma function fit to the PoCG raw data has a maximum of 1. In both the PoCG and Thalamus, the DG HRF has a larger upshoot amplitude and a smaller undershoot amplitude compared to the FLOBS HRF. There are also differences in latencies to peak, as described in the main text. Row 5: Group standard deviation timeseries plots. In both the PoCG and Thalamus, the FLOBS HRF has larger cross-subject variability, most noticeably during the undershoot component of the BOLD response.

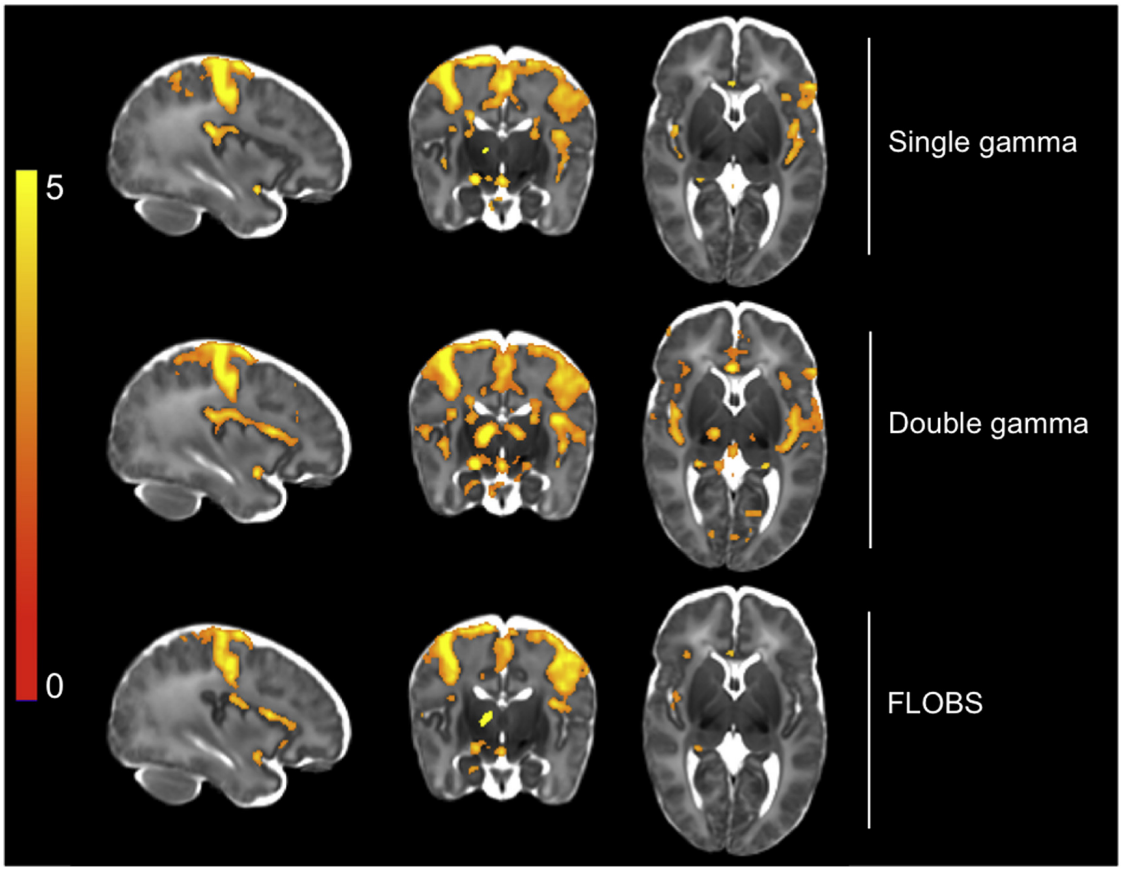

Figure 3.12: Comparison of the effects of HRF modelling on the group-level GLM output for all 15 subjects. Comparing thresholded (TFCE, default parameters, 5% FWER corrected) group t-statistic maps, the double gamma HRF had greatest sensitivity to signal in both cortical and subcortical structures, most noticeable in ipsilateral thalamus.

matter ROIs (Table 3.6). The region-constrained group comparisons demonstrated statistically significantly (voxel-based thresholding, 5% FWER corrected) larger t-statistics in the dHCP pipeline compared to the FEAT pipeline, localised to the PoCG. This is in contrast to our assessments of MCDC and spatial normalisation above (Sections 3.3.1 and 3.3.2), where statistically significant differences in t-statistics were not observed when contrasting these pipeline differences in isolation.

Similar to the mean t-statistic results, the dHCP pipeline resulted in a larger number of significantly active voxels compared to the FEAT pipeline using all three thresholding approaches (Table 3.6). The different thresholding approaches were less clear-cut in quantifying the proportion of activity mislocalised to white matter. Voxel-based thresholding was not appropriate here, because no activation survived

Table 3.5: Comparison of the effects of HRF modelling on t-statistics from the GLM output and significantly activated voxel counts for all 15 subjects. T-statistics: the double gamma HRF had greatest mean t-statistic in all grey matter ROIs (maximum t-statistic in parentheses). # active voxels: the double gamma HRF had the largest number of significantly activated voxels using all three thresholding approaches. Voxel is FDR corrected; Cluster and TFCE are FWER corrected. ACC = anterior cingulate cortex; DG = double gamma HRF; FLOBS = three basis function HRF; PoCG = postcentral gyrus; SG = single gamma HRF; TFCE = threshold free cluster enhancement; * = statistically significant difference.

		SG	**DG**	**FLOBS**
T-statistics	ACC	1.870 (4.889)	2.104 (4.030)	1.905 (3.953)
	Insula	2.773 (5.269)	2.977* (5.224)	2.141 (4.400)
	PoCG	2.877 (7.182)	2.959 (6.987)	2.729 (5.824)
	Thalamus	1.985 (5.709)	2.565 (6.797)	2.366 (5.920)
# active voxels	Voxel	33,421	66,984	57,569
	Cluster	72,987	82,837	46,648
	TFCE	38,896	66,756	29,665

thresholding with the FEAT pipeline. Cluster-based thresholding suggested a clear improvement in spatial specificity using the dHCP pipeline. TFCE thresholding revealed almost identical levels of mislocalised activity in both pipelines. However, given our assessment of spatial normalisation (Section 3.3.2) clearly demonstrated improved alignments using the dHCP pipeline, the similar proportions of activity in white matter measured here appear to be a TFCE "artefact". That is, given the increase in sensitivity to signal seen with the dHCP pipeline, we would expect this substantially larger grey matter activity to unavoidably 'enhance' neighbouring white matter t-statistics. Finally, using the thresholded group activity maps to qualitatively compare pipelines, the dHCP pipeline demonstrated noticeably increased activity in both cortical and subcortical structures (Figure 3.13). Taken together, these quantitative and qualitative GLM-based pipeline comparisons revealed dramatically improved spatial specificity and sensitivity to signal using the dHCP pipeline.

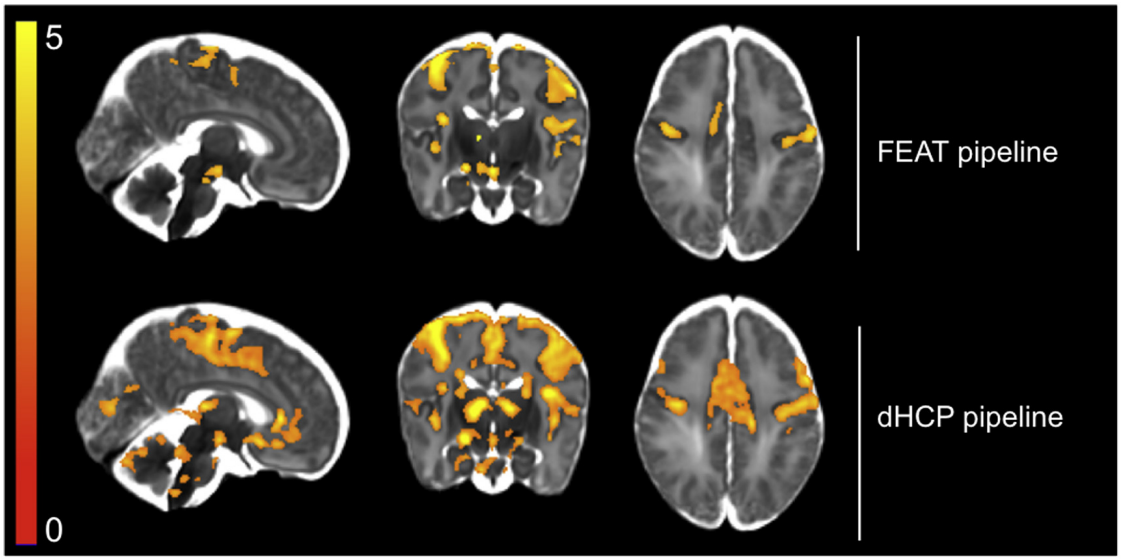

Figure 3.13: Overall comparison of FEAT and dHCP pipelines using the group-level GLM output for all 15 subjects. Comparing the thresholded group t-statistic maps (TFCE, default parameters, 5% FWER corrected), the dHCP pipeline resulted in increased sensitivity to signal in both cortical and subcortical regions. Compared to the unilateral activity detected in the thalamus and cingulate cortex using the FEAT pipeline, the bilateral activity detected in these regions using the dHCP pipeline could lead to a substantially different interpretation of the data.

3.4 Discussion

The central aim of this chapter was to identify optimal methods for preprocessing neonatal fMRI data and for subsequent modelling of the neonatal BOLD response to noxious stimulation. It was observed that the novel dHCP preprocessing pipeline provides significant improvements in motion correction, distortion correction, and spatial normalisation over the more standard adult FEAT preprocessing pipeline, modified for infants. Additionally, the value of incorporating FIX denoising was found to dramatically improve data quality, even for data heavily confounded by both spontaneous and stimulus-correlated motion. These advancements allowed the removal of motion-related artefacts due to slice misalignment (Andersson et al., 2017), susceptibility-by-movement distortions (Andersson et al., 2018), and spin history effects (Salimi-Khorshidi et al., 2014). This is noteworthy considering neonatal fMRI data are heavily motion-contaminated, especially noxious-response fMRI data, and the alternative analysis pipeline (the optimised FEAT pipeline,

Table 3.6: Overall comparison of FEAT and dHCP preprocessing pipelines using t-statistics from the GLM output and significantly activated voxel counts for all 15 subjects. T-statistics: using the dHCP pipeline, all grey matter ROIs showed an increase in mean t-statistic (maximum t-statistic in parentheses). This increase in t-statistics was revealed to be statistically significant (voxel-based thresholding, 5% FWER corrected) using the region-constrained analysis, localised to the PoCG. # active voxels: the dHCP pipeline also resulted in an increase in the number of significantly active voxels using all three thresholding approaches. % white matter: voxel-based thresholding was not valid due to zero voxels being activated. Cluster-based thresholding showed a clear reduction in mislocalised activity using the dHCP pipeline. TFCE-based thresholding showed a very modest increase in mislocalised activity using the dHCP pipeline, possibly due to unavoidable 'enhancement' of white matter voxels by the neighbouring grey matter. Voxel is FDR corrected; Cluster and TFCE are FWER corrected. ACC = anterior cingulate cortex; PoCG = postcentral gyrus; TFCE = threshold free cluster enhancement; * = statistically significant difference.

		FEAT	**dHCP**
T-statistics	ACC	1.529 (3.518)	2.104 (4.030)
	Insula	2.651 (5.244)	2.977 (5.224)
	PoCG	2.822 (6.568)	2.959* (6.987)
	Thalamus	2.168 (6.197)	2.565 (6.797)
# active voxels	Voxel	0	66,984
	Cluster	60,175	82,837
	TFCE	16,823	66,756
% white matter	Voxel	—	17.598
	Cluster	10.051	5.400
	TFCE	4.412	4.687

typically excluding FIX denoising) does not remove these particular artefacts. The improvements in spatial normalisation further reduce artefacts due to misalignment of data in standard space, which is common in neonatal fMRI data due to the large inter-subject variability in brain size, gyrification, and tissue contrast (Rivkin et al., 2004; Dubois and Dehaene-Lambertz, 2015). In addition to aiding in the development and validation of the novel dHCP fMRI preprocessing pipeline, and generalising its applicability to non-dHCP datasets, the identification of optimal preprocessing methods is of central importance to allowing meaningful analysis of individual differences in the BOLD response to noxious stimulation, a major aim of this thesis. As well as the core preprocessing pipeline comparisons, the

identification of a beneficial spatial filter extent (balancing SNR gain and spatial resolution loss) and HRF model (balancing over- and under-fitting) are processing stages of central importance to this goal.

The improvements in MCDC were primarily due to the novel application of FSL's EDDY, originally developed for diffusion MRI, to fMRI data. EDDY allows for the correction of additional motion-related artefacts (slice misalignments and susceptibility-by-movement distortions), but another important contrast with the FEAT pipeline is that EDDY performs motion correction and distortion correction simultaneously. The MCFLIRT-FUGUE-based MCDC corrections of the FEAT pipeline perform these corrections sequentially, which is suboptimal because motion artefacts and distortion artefacts are not independent, but closely interact. However, at the time of this study, these proposed advantages of EDDY for fMRI data were primarily theoretical, as the novel application was in development and without prior validation. This study was the first publication using these methods (Baxter et al., 2019). The adoption of novel methods comes with the risk of unknown detrimental effects, and the current implementation of EDDY to fMRI data makes the assumption that each fMRI volume is a spin-echo B0 image (when in fact, BOLD fMRI is typically a gradient-echo acquisition). While this spin-echo assumption does not account for signal dropout artefacts, this inaccurate assumption is weak, and after close manual inspection of all of our fMRI data, no unusual or unexpected effects were detectable. Additionally, the results of this study directly comparing the novel application of EDDY to the well-established use of MCFLIRT-FUGUE, demonstrated data quality improvements both pre- and post-HRF modelling of BOLD activity, strongly suggesting that EDDY successfully reduced noise variance without detrimental effect to or loss of the BOLD signal-of-interest. Since the publication of these results (Baxter et al., 2019), the pipeline has been released onto bioRxiv (Fitzgibbon et al., 2019) and the dHCP data made publicly available (`https://www.developingconnectome.org`). The high quality of the dHCP data, and the identification of RSNs consistent with existing literature (Doria et al., 2010; Mongerson et al., 2017) and sensible functional connectivity developmental

trends provide further validation for the novel methodology. The FIX results of this current study reaffirm the value of ICA-based clean-up methods (Griffanti et al., 2014; Salimi-Khorshidi et al., 2014). Thus, these results should hopefully encourage a more widespread adoption of these methods in the infant fMRI field, and exemplar signal and noise components have been provided (Supplementary Figures in Appendix A) to facilitate the adoption of these methods by other groups.

A major improvement in the registration pipeline was the refinement of intra-subject image alignment using BBR (Greve and Fischl, 2009). While the distinct effects of the functional-to-structural and structural-to-standard registrations were not individually assessed in this study, due to manual inspection of all alignment effects during the pipeline assessment, BBR was observed to improve alignments in every case. Since publication of these results (Baxter et al., 2019), an improved neonatal standard space template has been released (Schuh et al., 2018), and assessments internal to dHCP have found this template further improves the structural-to-template registration. Of the functional-to-structural and structural-to-standard registrations, the structural-to-standard step appears to be the weaker of the two. Of the 538 subjects analysed for the 2019 dHCP data release, 26 failed QC, 18 of which failed as a result of this registration step (Fitzgibbon et al., 2019). This is thus currently an active area of research to optimise the parameters of these registration tools. Since the publication of findings in this chapter, the noxious-response fMRI data used here and in later chapters have since been updated to include these improved registrations to the new standard space.

The use of spatial filtering is not without controversy, as the desired improvement in SNR is coupled with blurring of the image and thus a loss in spatial localisation (Turner and Geyer, 2014; Coalson, Van Essen, and Glasser, 2018). The use of spatial filtering should be context dependent, and there are certainly several instances in which it should be omitted. For example, studies into short-range functional connectivity (e.g. between neighbouring functional areas) must avoid introducing artefactually inflated functional connectivity measures due to inappropriate averaging of neighbouring functional regions via inappropriate spatial

filter use. Additionally, the spatial filter analysis performed in this study is specific to volumetric space, and functional data projected onto the cortical surface should not necessarily adopt our heuristic of 1.5 times the voxel size as a spatial filter extent. Finally, when performing group whole-brain voxelwise analyses, it is typical to use spatial statistics such as cluster-based and TFCE-based thresholding over voxel-based thresholding to improve sensitivity to signal. Inappropriate use of spatial filtering can induce a degree of blurring that results in the fusion of statistically significant 'clusters' of activity. This fusion of clusters when using spatial statistics causes severe issues for statistical inference, as the inference should technically be limited to the level of the cluster, not each individual functional region that might be spanned by a large significant cluster (Woo, Krishnan, and Wager, 2014). While spatial filtering should be omitted if not required, this step has been included in the following thesis chapter as it is indeed a valid step for the relevant research questions being addressed. The use of a modest sample size and the investigation of individual differences (as opposed to group comparison analysis) is an analysis setting in which the SNR gain of minimal spatial filtering very likely outweighs the data degradation due to image blurring.

There is a lack of research into the human neonate HRF, with only a single comprehensive study being performed to date (Arichi et al., 2012). The HRF parameterisation described in this study has not been incorporated into major analysis software packages, so there is currently no established 'canonical' infant HRF implementation for a defined age e.g. term neonate HRF (it is not possible to have a single canonical HRF for all infant age ranges due to the rapidly maturing haemodynamic response). While the group who originally developed the HRF predominantly use the three basis function (FLOBS) parameterisation (Allievi et al., 2016; Arichi et al., 2017), our group has used the single gamma (SG) variation in a previous publication (Goksan et al., 2015). Three parameterisations were studied here: the SG, double gamma (DG), and FLOBS HRF models. These results suggest that the DG parameterisation, intermediate in complexity to the SG and FLOBS parameterisations, was the preferred model for the infant event-related

noxious-response paradigm. While the FLOBS variant will undoubtedly fit the data more closely, the noise level of this infant fMRI dataset appeared to benefit from the more inflexible and robust DG HRF. However, it is worth noting that the experimental design used here and in the study in which the HRF was developed are very closely matched. They both involved somatosensory stimulation (one noxious, one innocuous), and both focused on primary somatosensory cortical region analyses. It is thus a possibility that our dataset's BOLD responses are very well fit by the first basis function alone of the FLOBS HRF, which is essentially the DG HRF. And these data might then not benefit from the additional flexibility afforded by the other basis functions. It is thus likely that other fMRI datasets using very different paradigms (e.g. visual or auditory stimulation) may indeed find the FLOBS parameterisation to be better suited than the DG HRF. In brief, more research into the infant HRF is certainly needed. Considering the data used throughout this thesis was collected using the noxious-response experimental paradigm described in this chapter, these HRF comparison results justify the adoption of the DG HRF throughout this thesis.

The optimisation of neonatal fMRI data analysis methods that explicitly address the challenges posed by this population is central to improving infant fMRI data interpretation, minimising the heterogeneity of fMRI analyses, and facilitating cross-study comparisons. As the fMRI field moves from group-level analysis to subject-level analysis (Vogt, 2015; Finn et al., 2015; Tavor et al., 2016; Parker Jones et al., 2017), the optimisation of data preprocessing for the infant population is imperative. While this work was a key first step in achieving the research aims of this thesis, it is clear these results will help form the foundations on which further infant fMRI research can be conducted, and provides a platform to address fundamental neuroscientific questions, such as investigating how environmental factors shape central nervous system function during early human development.

3.5 Acknowledgements

We are grateful for the provision of simultaneous multi-slice (multi-band) pulse sequence and reconstruction algorithms from the Centre for Magnetic Resonance

Research, University of Minnesota.

We are grateful for Tomoki Arichi's provision of the infant FLOBS HRF model and advice on deriving appropriate effect size statistics from the model.

4

Predicting the infant haemodynamic response to nociceptive input from resting-state activity

4.1 Introduction

Newborn infants undergo a multitude of necessary painful procedures shortly after birth as part of standard clinical care (Carbajal et al., 2008). It takes several years' training to acquire the clinical acumen to competently decide upon appropriate pain management strategies for infants undergoing clinically required painful procedures, such as type and titration of analgesia and the required intensity of post-procedural monitoring. However, infants' lack of verbal communication and limited behavioural repertoire, and the brief extra-uterine time window for establishing a medical history, ensure a considerable degree of uncertainty exists in clinical decision making for effective personalised analgesic care of the newborn infant. The inherent practical and ethical constraints in researching pain in newborns has resulted in a reliance on ambiguous behavioural and physiological responses to noxious events as indicators of an infant's pain experience and sensitivity (Lee and Stevens, 2013). This uncertainty and ambiguity results in pain often going under-detected and under-treated in the infant population (Roofthooft et al., 2014). It would thus be of immense value to identify novel informative indicators of a newborn's unique sensitivity to nociceptive input that do not rely on prior observations of pain-related responses to clinical procedures, and that could reduce the uncertainty and ambiguity during clinical

decision-making processes regarding infant pain management. In this study, a cohort of healthy newborn infants was studied in an experimental setting to identify novel predictors of the infants' neural responses to mild nociceptive input. The central question was whether it was possible to predict a newborn infant's cerebral haemodynamic response amplitude from non-nociceptive data.

The human brain plays a central role in nociceptive processing and pain experience. Previous studies using electroencephalography (EEG) and near infrared spectroscopy (NIRS) have underscored the value of using brain-based approaches to identify neural correlates of infant pain (Slater et al., 2006; Slater et al., 2010a; Verriotis et al., 2016), highlighting important dissociations between brain activity and observable pain-related behaviours (Slater et al., 2008; Slater et al., 2010b; Bembich et al., 2015). From the age of approximately 35 weeks postmenstrual age (PMA), it is possible to distinguish infants' neural responses to noxious and non-noxious stimulation using EEG (Fabrizi et al., 2011). In the current study, a cohort of healthy peri-term-aged infants (35.9 – 41.4 weeks PMA) were exposed to a mild non-skin-breaking experimental noxious stimulus applied to the infant's foot to mimic, without causing distress to the infant, the skin-breaking heel lance blood test procedure, one of the commonest clinical procedures to which newborns are exposed (Courtois et al., 2016). The experimental sharp-touch stimulus activates Aδ-fibres evoking a noxious-specific event related potential (ERP), that is also seen in response to the heel lance (Hartley et al., 2015; Hartley et al., 2017). Functional magnetic resonance imaging (fMRI) was used to measure the amplitude of the cerebral haemodynamic response to the nociceptive input. To date, only three fMRI studies on infant pain processing have been published (Williams et al., 2015; Goksan et al., 2015; Goksan et al., 2018). The challenges of studying infant pain in the MRI environment and analysing infant fMRI data have been major obstacles to the widespread use of the fMRI modality. However, the recent establishment of a feasible experimental paradigm to study infant nociception (Goksan et al., 2015), as well as the outline of an optimised analysis pipeline for infant noxious-response fMRI datasets (Baxter et al., 2019), provide favourable conditions for the pursuit

of fMRI use. Unlike EEG and NIRS, fMRI provides whole-brain coverage and unparalleled spatial resolution that will undoubtedly provide valuable insight into the neurophysiological basis for infant pain.

An infant's pattern of spontaneous brain activity is unique to an infant and reflects the brain's underlying structural and functional neural architectures. To identify novel predictors of infants' noxious-response amplitudes, spontaneous brain activity was recorded using resting-state fMRI (rs-fMRI), during which the infants lie passively in the scanner. A set of core resting-state networks were identified through replication in an independent dataset, a subset of healthy newborn infants from the Developing Human Connectome Project (dHCP) dataset (Fitzgibbon et al., 2019), and the pattern of network activity amplitudes was assessed for the capacity to predict the infants' noxious-response amplitudes. In adults, the functional connectome of the brain has been used to uniquely identify subjects (Finn et al., 2015) and is capable of predicting unique brain activity patterns to a wide range of cognitive and behavioural tasks (Bzdok et al., 2016; Tabor et al., 2017; Parker Jones et al., 2017). In infants, resting-state networks closely resembling those observed in adults during rest and task conditions (Smith et al., 2009) are established and functional prior to term age (Fransson et al., 2007; Fransson et al., 2009; Doria et al., 2010). However, the relationship between resting-state and pain and nociception has yet to be established in newborns. Resting-state activity has great potential to provide novel insight into the multiple dimensions of infant pain experience and expression, as it is unique to each individual infant, originates in the organ from which pain perception emerges, and acquisition does not require pain or nociception exposure for the infant.

In this study, focus was placed on network amplitudes as potential predictors over measures of functional connectivity for three reasons. There is a strong dependency of functional connectivity measures on signal amplitude (Cole et al., 2016; Duff et al., 2018; Friston, 2011), yet amplitude effects are currently under-explored in infant fMRI studies. There is now increased recognition of the functional importance of fMRI activity amplitude (Bijsterbosch et al., 2017; Duff et al., 2008; Mennes

et al., 2013; Xu et al., 2014; Zou et al., 2013), with resting-state activity amplitudes shown to associate with task activity in adults (Mennes et al., 2010; Mennes et al., 2011). The challenges of studying infant pain with MRI typically constrains study sample sizes to modest numbers ($n < 20$) (Williams et al., 2015; Goksan et al., 2015; Goksan et al., 2018), so studies attempting to identify novel indicators of infant pain and nociception should avoid issues due to combinatorial explosion (N networks produce ${}^{N}C_{2}$ functional connectivity measures, but only N amplitude measures). Another key methodological decision was to focus on assessing prediction over association (e.g. correlation or regression). While performing a correlation or regression analysis typically involves fitting a model that can then be used to generate predictions for novel data, establishing a statistically significant association does not confirm generalisable predictability for new data (Poldrack, Huckins, and Varoquaux, 2019). Establishing statistically significant predictive ability of associations between resting-state activity and noxious-response activity would more convincingly suggests that an infant's spontaneous brain activity could be a novel source of valuable aids to clinical decision making.

Infant diffusion MRI (dMRI) data were also analysed to assess potential structure-function relationships between an infant's haemodynamic response amplitude and the underlying white matter microstructure. While relationships between the structural and functional architectures of the brain provide neuroscientifically valuable insight in their own right, association with structure was also used to assess the temporal stability of the observed fMRI noxious-response amplitudes. An infant's neural response to stimuli is dependent on transient state effects such as wakefulness and emotional state (Jones et al., 2017; Slater et al., 2006), as well as relatively stable trait effects. In the context of identifying novel indicators of infant nociceptive sensitivity, temporally stable trait effects would be of greater clinical utility than temporally transient state effects.

Infants' white matter microstructural complexity was quantified using the mean kurtosis (MK) parameter derived using a diffusion kurtosis imaging (DKI) model (Jensen et al., 2005; Jensen and Helpern, 2010; Steven, Zhuo, and Melhem, 2014),

and assessed in a whole-brain voxelwise manner using tract-based spatial statistics (TBSS) (Ball et al., 2010; Smith et al., 2006). While the MK parameter has not been studied extensively in neonatal populations, its usefulness in studying brain development in both humans and rodents has already been previously demonstrated (Cheung et al., 2009; Paydar et al., 2014). MK was used in this study for its sensitivity to both isotropic and anisotropic diffusion barrier effects, as well as its ability to accurately quantify diffusional heterogeneity in white matter regions with complex fibre orientations, such as crossing-fibre arrangements (Jensen et al., 2005; Steven, Zhuo, and Melhem, 2014). Previous studies in infant structure-function relations highlighted the importance of primary project fibre bundles, the optic radiations in the context of these visual function studies (Bassi et al., 2008; Berman et al., 2009; Groppo et al., 2014). For sharp-touch noxious stimuli, many of the primary somatosensory projection bundles traverse regions of the centrum semiovale with multiple fibre orientations and complex crossing-fibre arrangements. The most commonly studied DTI diffusion parameters (mean diffusivity, fractional anisotropy, axial diffusivity, and radial diffusivity) are notoriously challenging to interpret in these complex white matter regions (Alexander et al., 2001; Tuch et al., 2003; Vos et al., 2012; Wheeler-Kingshott and Cercignani, 2009).

This work provides valuable insight into the neurophysiological basis for normative variability in the haemodynamic response to nociceptive input in a group of healthy newborn infants. A novel nociception-related neural structure-function relationship is revealed, and tight coupling between an infant's resting-state and noxious-response neural activities provides proof-of-concept that an infant's noxious-free brain activity can predict noxious-response brain activity.

4.2 Methods

4.2.1 Infant recruitment and demographics

Infant recruitment was as described in Section 3.2.1 (Chapter 3). A clinical investigator of the Paediatric Neuroimaging Group recruited healthy neonates from the postnatal ward at the John Radcliffe Hospital (Oxford University Hospital

NHS Trust) for an MRI scan. Infants were considered healthy if they were inpatients on the postnatal ward that never required admission to the neonatal unit, had no history of congenital conditions or neurological problems, and were clinically stable at the time of study. Written informed consent was obtained from parents prior to the study. Ethical approval was obtained from an NHS Research Ethics Committee (National Research Ethics Service, REC reference: 12/SC/0447), and research was conducted in accordance with standards set by Good Clinical Practice guidelines and the Declaration of Helsinki.

Infant demographics differ to that described in Section 3.2.1, due to the increased sample size and requirement of both noxious-response fMRI and rs-fMRI data per infant for the current study. Demographic details of this 18-infant sample are displayed in Table 4.1. The mean gestational age (GA) at birth was 38.3 weeks (range: 35.3 – 41.4 weeks), and the mean birth weight was 3,343 g (range: 1,910 – 4,570 g). At the time of scan, the mean postmenstrual age (PMA) was 38.7 weeks (range: 35.9 – 41.7 weeks), mean postnatal age (PNA) was 2.8 days (range 1 – 10 days), and mean total brain volume (TBV) was 289,823 mm^3 (range: 212,410 – 416,904 mm^3). The sample included eight females and ten males.

4.2.2 Experimental setup and design

Experimental setup and design were as described in Section 3.2.2 (Chapter 3). Neonates were transported to the Wellcome Centre for Integrative Neuroimaging (Oxford, UK), then screened for metal, fed, and swaddled prior to scanning. Infants were fitted with silicone ear plugs (Hush Plugz, Robinson Healthcare, Worksop, UK; 23 dB attenuation), ear-muffs (Minimuffs, Natus Medical Inc., Galway, Ireland; 7 dB attenuation), and ear-defenders (Em's 4 Bubs Baby Earmuffs, Em's 4 Kids, Brisbane, Australia; 22 dB attenuation), and placed on a vacuum-positioning mattress with additional soft padding around the head to restrict motion. Heart rate and blood oxygen saturation were monitored throughout scanning (Fibre Optic Pulse Oximeter, Nonin Medical, Plymouth, Minnesota).

Table 4.1: Demographic details of the 18-infant local dataset. PMA = postmenstrual age; GA = gestational age; PNA = postnatal age; BW = birth weight; TBV = total brain volume. * = excluded from dMRI analyses due to missing data.

Infant	**PMA (weeks)**	**GA (weeks)**	**PNA (days)**	**Sex**	**BW (grams)**	**TBV (mm^3)**
1	40.6	40.3	2	M	3,880	283,685
2	37.4	37.1	2	M	4,570	276,222
3	35.9	35.3	4	F	1,910	212,410
4	35.9	35.6	2	M	3,180	284,623
5	38.3	38.0	2	M	3,400	309,632
6	38.0	37.3	4	F	3,410	306,530
7	39.6	39.3	2	F	3,250	301,086
8	36.4	36.0	3	F	3,510	260,268
9	38.0	37.9	1	M	2,490	221,792
10	40.7	40.6	1	M	4,300	346,861
11	40.4	40.1	2	M	4,040	356,454
12	39.3	39.0	2	F	3,775	300,094
13	38.9	38.6	2	F	2,950	286,714
14*	41.7	41.4	2	M	3,400	297,004
15	40.4	39.0	10	F	3,750	416,904
16	38.3	38.0	2	M	2,780	247,199
17	37.4	36.4	7	F	2,235	257,069
18	39.0	38.9	1	M	3,350	252,271

For the nociception paradigm, an event-related experimental design was used, and the mild non-skin-breaking nociceptive stimulus was a 128 mN sharp-touch pinprick (PinPrick Stimulator, MRC Systems). Ten trials of the stimulus were delivered to the dorsum of the left foot, each trial was 1 s, and the minimum inter-stimulus interval was 25 s. This long inter-stimulus interval was used to minimise the influence of motion at the time of stimulus delivery. The stimuli were applied when the infants were naturally still, and were time-locked to the fMRI data using Neurobehavioural Systems software (Presentation, `https://www.neurobs.com`). The same two-researcher setup as described in Section 3.2.2 was used in this study. For all other scan types, infants lay passively in the scanner. No sedatives were used at any stage of this study.

4.2.3 MRI data acquisition

The T2 structural, GRE dual-echo fieldmap, and noxious-response fMRI data were acquired as described in Section 3.2.3 (Chapter 3). For completion, they are reproduced here. All data were collected on a 3T Siemens Prisma with an adult 32 channel receive coil.

The T2 structural data acquisition was: T2-weighted, TSE (factor 11), 150° flip angle, TE = 89 ms, TR = 14,740 ms, parallel imaging GRAPPA 3, 192×192 in-plane matrix size, 126 slices, 1 mm isotropic voxels, and 2 mins 13 s acquisition time.

The GRE dual-echo fieldmap data acquisition was: gradient echo, 2DFT readout, dual echo TE1/TE2 = 4.92/7.38 ms, TR = 550 ms, 46° flip angle, 90 × 90 in-plane matrix size, 56 slices, 2 mm isotropic voxels, and 1 min 40 s acquisition time.

The noxious-response fMRI data acquisition was: T2* BOLD-weighted, gradient echo, EPI readout, 70° flip angle, TE = 50 ms (Goksan et al., 2017), TR = 1,300 ms, multiband 4 (Moeller et al., 2010; Xu et al., 2013), 90 × 90 in-plane matrix size, 56 slices, 2 mm isotropic voxels, AP phase encode direction, a single-band reference (SBref) image was acquired at the start, and the mean acquisition time was approximately 6 min (approximately 277 volumes).

The rs-fMRI data acquisition was: T2* BOLD-weighted, gradient echo, EPI readout, 70° flip angle, TE = 50 ms, TR = 1,300 ms, multiband 4, 90 × 90 in-plane matrix size, 56 slices, 2 mm isotropic voxels, AP phase encode direction, a single-band reference (SBref) image was acquired at the start, and the acquisition time was 10 mins 50 s (500 volumes).

The dMRI data acquisition was: T2 diffusion-weighted, spin echo, EPI readout, 90° flip angle, TE = 73 ms, TR = 2,900 ms, multiband 3, 102 × 102 in-plane matrix size, 60 slices, 1.75 mm isotropic voxels, AP phase encode direction, multishell (b500, b1000, b2000), 13 b0 volumes, a total of 140 directions uniformly distributed over the whole sphere (b500×20, b1000×50, b2000×70 directions), and approximately 8 mins acquisition time. For distortion correction of the diffusion data, a blip-up/blip-down spin-echo-derived fieldmap was constructed from phase-reversed b0 images. Thirteen

AP/blip-down b0 images were interspersed in the diffusion data acquisition, and eight PA/blip-up b0 images were acquired in a separate acquisition immediately after.

4.2.4 MRI data preprocessing

The T2 structural and GRE dual-echo fieldmap data were preprocessed as described in Section 3.2.4. In brief, the structural data were processed (brain extraction, bias field correction, and tissue segmentation) using the MIRTK Draw-EM neonatal pipeline (Makropoulos et al., 2014), which is the tool forming the basis of the dHCP structural preprocessing pipeline (Makropoulos et al., 2018). The fieldmap data were processed using a modified version of fsl_prepare_fieldmap.

Both the noxious-response fMRI and rs-fMRI data were preprocessed using an extended version of the dHCP fMRI preprocessing pipeline (Baxter et al., 2019; Fitzgibbon et al., 2019) (also Chapter 3). The functional data were simultaneously corrected for motion and distortion using FSL's EDDY (Andersson et al., 2016; Andersson and Sotiropoulos, 2016). Volume-to-volume and slice-to-volume alignments corrected for inter- and intra-volume motion, respectively (Andersson et al., 2017). Estimated dynamic distortion correction, using a GRE dual-echo fieldmap, corrected for both susceptibility-induced off-resonance and susceptibility-by-movement distortions (Andersson et al., 2018). Noxious-response fMRI data were high-pass temporally filtered at 0.01 Hz, and rs-fMRI data at 0.005 Hz. Data were then denoised using FSL's FIX (Griffanti et al., 2014; Salimi-Khorshidi et al., 2014) to remove both noise independent components and 24 head motion parameter timeseries. Data were low-pass spatially filtered with a 3 mm FWHM filter using FSL's SUSAN (Smith and Brady, 1997), then grand mean scaling to a global spatiotemporal median of 10,000. For spatial normalisation to the neonatal standard space (Schuh et al., 2018), the data were first registered from functional space to the infant's T2 structural space, via the SBref, using 6 degrees of freedom (DOF) rigid-body alignment, refined using BBR (Greve and Fischl, 2009) with FSL's FLIRT (Jenkinson and Smith, 2001; Jenkinson et al., 2002). The registration from structural space to an age-matched standard template, and from this age-matched

standard template to the global standard 40-week template (Schuh et al., 2018) was performed using ANTs's SyN (Avants et al., 2008).

The diffusion data were analysed using the dHCP dMRI preprocessing pipeline (Bastiani et al., 2019). A non-insignificant aspect of the current study involved the implementation of this novel preprocessing pipeline. The blip-up and blip-down b0 images were used to generate the fieldmap using FSL's TOPUP (Andersson, Skare, and Ashburner, 2003; Smith et al., 2004). The diffusion data were simultaneously corrected for motion, distortion, and eddy currents using FSL's EDDY. As with the functional data, the motion correction involves both volume-to-volume and slice-to-volume alignment, and the distortion correction involves both susceptibility-induced off-resonance and susceptibility-by-movement corrections. Similar to the functional data analysis, spatial normalisation involved first registering from diffusion space to the infant's structural space using 6 DOF rigid-body alignment, refined using BBR with FSL's FLIRT. The registration from structural space to an age-matched standard template, and from this age-matched standard template to the global standard 40-week template is performed using ANTs's SyN.

4.2.5 Noxious-response maps and amplitudes

For each infant, a noxious-response map was generated using standard subject-level voxelwise GLM analysis in FSL's FEAT (Jenkinson et al., 2012). This GLM analysis included FEAT's default prewhitening (Woolrich et al., 2001), 0.01 Hz high-pass temporal filtering, and fitting of the term-neonate double-gamma HRF (Arichi et al., 2012; Baxter et al., 2019), as described in Chapter 3. Using the 18 infants' noxious-response regression parameter maps, a group average t-statistic map was generated. Individual infants' regression parameter maps were used as the subject-level noxious-response maps; the group-average t-statistic map was used as the group-level noxious-response map. To summarise an individual infant's noxious-response map (regression parameter map) to a single scalar measure of noxious-response amplitude, the group-level noxious-response map was regressed

onto the infant's noxious-response map. This spatial regression coefficient was used as the infant's noxious-response amplitude.

To assess each infant's HRF goodness-of-fit, peristimulus timeseries plots were constructed using the default method in FSL's FEAT. In brief, the ROI used was the noxious-response map voxel with largest z-statistic (both positive and negative). All raw time points from individual trials are presented along with their mean timeseries and the mean HRF timeseries. The mean raw timeseries and mean HRF timeseries were visually inspected to assess HRF fit quality.

4.2.6 Resting-state network maps and amplitudes

To identify a robust set of core resting-state networks in the infants' rs-fMRI data, identified resting-state networks were compared to those identified in a subset of the dHCP dataset, which had previously been produced as part of the Develop Human Connectome Project (Fitzgibbon et al., 2018) (`http://www.developingconnectome.org`). A robust set of core resting-state networks was defined as those consistently identified across both datasets. The dHCP rs-fMRI data were acquired as described in Price et al., 2015 and Hughes et al., 2017, the data were preprocessed and resting-state network analysis performed as described in Fitzgibbon et al., 2019. The dHCP data-subset analysed to produce the resting-state network maps used in this study included 242 healthy term-aged infants: mean GA at birth = 38.6 weeks; mean PMA at scan = 40.4 weeks; 112 females and 124 males. The resting-state network analysis performed on the local dataset was closely matched to that described for the dHCP dataset (Fitzgibbon et al., 2019). In brief, probabilistic functional mode (PFM) analysis using FSL's PROFUMO (Harrison et al., 2015; Harrison et al., 2019) was run on both datasets with a pre-specified dimensionality of 25, and using the infant double-gamma HRF (Arichi et al., 2012; Baxter et al., 2019) as the temporal prior.

Similar to ICA, PFM is a matrix factorisation method that decomposes the fMRI data into modes of spatial and temporal patterns. However, PFM does not impose spatial independence, and using a Bayesian approach, allows the incorporation

of the infant double-gamma HRF as a temporal prior to identify modes with BOLD-like timeseries. The group-level PFM modes are also directly informed by subject-level modes, as both levels are modelled simultaneously, so the PFM group-level networks are typically more representative of the individual subject data than standard temporal concatenation group-ICA analysis. In order to achieve functional correspondence between the PFM group-level and subject-level maps, a PFM mode only appears at the group-level if it is identifiable in the majority of the individual infant's data. The PROFUMO Bayesian model complexity penalties can thus eliminate modes, returning a number of group-level modes that can be less than the pre-specified dimensionality. In summary, the data-determined dimensionality can be less than the pre-specified dimensionality.

As will be described in Section 4.3.2, the dHCP resting-state network maps had greater SNR due to the significantly larger sample size. Thus, the dHCP resting-state network maps forming the robust set of core resting-state networks were used as the template maps to extract resting-state network amplitudes from our local dataset rs-fMRI data. These network maps were spatially regressed onto each infant's resting-state functional data using multiple regression, resulting in network timeseries. To quantify the resting-state network amplitudes from these timeseries, the typical metric used is the standard deviation, which is the default in FSL's FSLNets (Smith et al., 2011). However, the standard deviation is sensitive to outliers, which in this context, appear as timeseries spikes typically reflecting subject head motion. The presence of timeseries outliers was directly tested using the "quartiles method": outliers defined as values more than 1.5 interquartile ranges above the upper quartile or below the lower quartile. Despite the advanced motion and distortion correction and bespoke FIX denoising during preprocessing, all 18 infants had one or more network timeseries that included outliers – see Figure 4.4, outliers displayed in red. For this reason, resting-state network amplitudes were quantified using the median absolute deviation (MAD), due to the MAD's increased robustness to outliers.

4.2.7 Non-functional variables and fMRI imaging confounds

In addition to testing the association between noxious-response amplitudes and resting-state network amplitudes, associations with noxious-response amplitudes were also assessed for a set of resting-state imaging confounds and a set of non-functional variables. Testing the resting-state imaging confounds directly assessed whether functional coupling between resting-state and noxious-response activities could be explained by unwanted confound variables. In contrast, testing the non-functional variables assessed whether functional coupling between resting-state and noxious-response activities could be explained by biologically interesting mediator variables. When testing the association between noxious-response amplitudes and resting-state network amplitudes, the resting-state network amplitudes were adjusted for the resting-state imaging confounds. Similarly, the noxious-response amplitudes were adjusted for a set of noxious-response imaging confounds. The nature of and method for quantifying the non-functional variables, resting-state imaging confounds, and noxious-response imaging confounds is outlined below. Infant age is also discussed for its role as both an interesting non-functional mediator variable and an unwanted non-functional confound variable.

The set of non-functional variables included six metrics: postmenstrual age (PMA), gestational age (GA), postnatal age (PNA), birth weight (BW), total brain volume (TBV), and sex. These metrics are displayed in Table 4.1. The TBV was calculated from the infants' structural MRI data using the tissue segmentation outputs of the structural preprocessing pipeline.

The set of resting-state imaging confounds included three metrics extracted from each infant's rs-fMRI data: mean head motion, CSF signal amplitude, and white matter signal amplitude. Mean head motion was defined as the mean framewise displacement across the entire rs-fMRI scan. The CSF amplitudes were included to attempt capturing cross-subject variability in residual cardiac pulsatility signal that may remain post-FIX denoising (Eklund, Knutsson, and Nichols, 2019). The white matter amplitudes were included to attempt capturing cross-subject variability in global signal amplitudes, considering global signals contaminate rs-fMRI data and

are not removed by any of the preprocessing steps performed in this study (Glasser et al., 2018). To produce CSF and white matter signal timeseries with minimal partial volume contamination, conservative region-of-interest (ROI) masks were defined using the neonatal template atlas (Schuh et al., 2018) (Supplementary Figure B.1). Mean timeseries were extracted, and amplitudes quantified as the timeseries MAD values. These three resting-state imaging confounds were directly tested for association with noxious-response amplitudes, and used for confound-adjusting the resting-state network amplitudes.

The set of noxious-response imaging confounds included three metrics extracted from each infant's noxious-response fMRI data: mean head motion, stimulus-correlated head motion, and CSF signal amplitude. Mean head motion was defined as the mean framewise displacement across the entire noxious-response fMRI scan. Due to the experimenter control of the experimental pinprick application timing, an additional stimulus-correlated head motion metric could be estimated. This was highly relevant given the noxious nature of the stimulus, and infants' inability to suppress motion in response to noxious stimulation, which can result in instances of strong stimulus-correlated head motion (see Figure 3.3A and 3.8 for a clear example). Stimulus-correlated head motion was estimated as a multiple correlation coefficient z-statistic (Fisher r-to-z transformation) between the predicted BOLD response (stimulus application timeseries convolved with the HRF) and the 24 head motion parameter timeseries (estimated by EDDY during motion correction). Finally, CSF amplitude was estimated from each infant's noxious-response map (regression coefficient map) as the mean regression coefficient within the CSF ROI. These three noxious-response imaging confounds were used for confound-adjusting the noxious-response maps and amplitudes.

While postmenstrual age (PMA) is a biologically interesting non-functional variable, there are also several MRI imaging confounds that correlate with PMA. As PMA increases, both brain volume and cortical gyrification increase, and brain tissue composition and contrast also change. In the context of MRI, these effects can result in cross-infant variability in partial volume effects and spatial normalisation

quality that correlate with PMA. As part of the assessment of the non-functional variables, the biological value of age as a potential mediator variable for functional coupling between resting-state and noxious-response activities was assessed. As will be described in the results, the 18 infants' noxious-response amplitudes were directly correlated with age (PMA, GA, and PNA). No statistically significant correlations or trends toward significance were observed between noxious-response amplitudes and any age metric (see top row of Supplementary Figure B.3). Age was thus deemed unlikely to be an important mediator of potential functional coupling between resting-state and noxious-response activities in this sample of neonates. It is also worth noting that the current study was cross-sectional in design, and the current sample had a relatively narrow age-range: comparing infant PMA at time of scan, the youngest and oldest infant differ by only 40 days; comparing infant PNA at time of scan, the youngest and oldest infant differ by only 10 days. While the relationship between age and noxious-response amplitudes was tested at several points in this study, in several of the other analyses performed, age is adjusted for in order to weaken potential age-related confound effects. When and how age is adjusted for will be described explicitly.

4.2.8 Predicting noxious-response amplitudes

In this prediction analysis, the responses to be predicted were the infants' noxious-response amplitudes (defined in Section 4.2.5). Three sets of predictors were tested for predictive capacity: the set of nine resting-state network amplitudes (defined in Section 4.2.6), the set of six non-functional variables (defined in Section 4.2.7), and the set of three resting-state imaging confounds (defined in Section 4.2.7). For all three sets of predictors, a support vector regression (SVR) model with a linear kernel was used. Linear SVR was selected over a linear regression via ordinary least squares, due to SVR's greater robustness to outliers: the OLS linear regression uses an L2 loss function taking the square of the error, while SVR uses an L1 loss function taking the absolute value of the error. Out-of-sample predictions were generated using leave-one-out cross validation (LOO-CV). The noxious-response amplitudes were

confound-adjusted for the three noxious-response imaging confounds (mean head motion, stimulus-correlated head motion, and CSF amplitude). When generating predictions using the resting-state network amplitudes, this set of predictors was confound-adjusted for PMA and the three resting-state imaging confounds (mean head motion, CSF amplitude, and white matter amplitude).

The linear SVR model was fit in Python using scikit-learn packages (Pedregosa et al., 2011), with all steps performed in a LOO-CV manner. Confound adjustment of the resting-state network amplitudes and noxious-response amplitudes were performed using cross-validated confound regression, implemented using the publicly available code by Lukas Snoek (`https://github.com/lukassnoek/MVCA`), as described in the author's article (Snoek, Miletić, and Scholte, 2019). Responses were z-scaled using training set means and standard deviations. PCA was used on the resting-state network amplitude predictors to improve SNR, retaining 98% variance (generally the first seven components). The scikit-learn SVR parameters were: kernel = linear, loss function = epsilon insensitive, epsilon = 0.1, regularization = ridge, regularization strength = $\{0.001, 0.01, 0.1, 1\}$. Optimisation of the regularization strength parameter was performed using a grid search over this set of values. Regularisation tuning and SVR model training were optimised to minimise mean squared error.

The prediction accuracy was assessed using three summary metrics: root mean squared error (RMSE), sums-of-squares formulation of the coefficient of determination (R^2), and Spearman's rank correlation coefficient (R_{Sp}). The RMSE was selected as the primary metric of prediction accuracy, as it directly quantifies the error (the difference between predicted and actual observed values), and it is in original units. The R^2 was also reported, as its value is easily related to the success of the predictions. Unlike a coefficient of determination derived from squaring the Pearson correlation coefficient, this sums-of-squares formulation has a range of $-\infty < R^2 \leq 1$. R^2 is interpreted as the proportion of the total variation of the response (about its mean) that is accounted for by the fitted model, and R^2 can be negative when a model misspecification is used or a model fitting technique other

than linear least squares regression, such as a robust regression method, is used for data with outliers (Kvalseth, 1985). The R_{Sp} between predicted and observed noxious-response amplitudes was also reported. While R_{Sp} is not informative about additive or proportional differences between observations and predictions, it may be valuable to know a model's ability to correctly rank infants' noxious-response amplitudes, on a relative scale, from lowest to highest. To test the statistical significance of the RMSE, R^2, and R_{Sp} measures using null hypothesis testing, one-tailed significance tests were performed using permutation analysis, running 1,000 permutations through the full prediction pipeline.

4.2.9 Validating resting-state-based predictions

The SVR prediction analysis is relatively computationally complex, due to the need to maintain independence between training and test sets throughout e.g. using confound regression, using PCA on predictors, and standardising responses. If the LOO-CV procedure does not encompass all operations applied to the data, leakage between training and test sets can occur inflating measures of prediction accuracy (Poldrack, Huckins, and Varoquaux, 2019). In order to test for associations between noxious-response amplitudes and the three sets of predictors in a computationally simpler manner without LOO-CV, a traditional multiple linear regression (MLR) mapping analysis was performed. An MLR model was fit in a whole-dataset manner assessing association between predictors and responses. Unlike the SVR analysis, the MLR analysis was not used to generate predictions but was provided to complement the SVR prediction analysis. This MLR mapping analysis was performed using the 18 infants' noxious-response maps. The global null hypothesis between predictors and responses was tested in a whole-brain voxelwise manner using non-parametric combination (NPC) analysis (Winkler et al., 2016) with FSL's PALM (Winkler et al., 2014). Similar to the prediction analyses, all mapping analyses were adjusted for the three noxious-response fMRI imaging confounds, and only the resting-state network amplitude analysis was adjusted for PMA and the three resting-state imaging confounds. All NPC permutation analyses involved 10,000 permutations,

and results were thresholded using TFCE (Smith and Nichols, 2009) with default parameters and FWER-corrected for multiple testing. Both unthresholded and thresholded maps are provided. For thresholded results, final significance level thresholds used are mentioned as appropriate.

Owing to the observation of robust functional coupling between resting-state and noxious-response activities in both the SVR prediction and MLR mapping analyses, the nature of the relationship between the infants' noxious-response amplitudes (the scalar values) and resting-state network amplitudes was further assessed. The potential for an artefactual global signal present throughout the brain in both the resting-state and noxious-response runs could cause artefactual similarities between an infant's noxious-response and resting-state network amplitudes. Univariate Pearson correlation analyses were performed between the noxious-response amplitudes and each individual resting-state network amplitude to test for global or network-specific associations. The statistical significance of these correlations was assessed using two-tailed permutation tests with 10,000 permutations and results FWER-corrected for multiple testing using PALM (Winkler et al., 2014; Winkler et al., 2016).

4.2.10 Relating noxious-response amplitude to white matter microstructure

Examining structure-function relationships is a central neuroscientific endeavour to reveal the relationship between a subject's brain activity and behaviour and their brain's structural architecture. In this study, the infant's noxious-response amplitudes were assessed for structure-function associations through analysing white matter microstructure. This analysis was also valuable for assessing the temporal stability of the infants' noxious-response amplitudes. Considering the relative temporal stability of white matter microstructure and its insensitivity to infant wakefulness and emotional state, an observed structure-function relationship would suggest the noxious-response amplitudes were also temporally stable trait effects, rather than temporally transient state effects. The infants' dMRI data were used to assess structure, due to the T2 structural images not being of high enough quality; additionally, temporal stability was assessed using structure-function association

rather than looking at stability across multiple test occasions, as infants could only be tested on a single occasion.

White matter microstructural complexity was quantified as mean kurtosis (MK) using FSL's DTIFIT (Jenkinson et al., 2012) as part of the dHCP dMRI preprocessing pipeline (Bastiani et al., 2019). MK quantifies the heterogeneity of the diffusion environment or diffusional complexity, and was selected for its sensitivity to both isotropic and anisotropic diffusion barrier effects, as well as its ability to accurately quantify diffusional complexity in white matter regions with complex fibre orientations, such as crossing-fibre arrangements (Jensen et al., 2005; Jensen and Helpern, 2010; Steven, Zhuo, and Melhem, 2014). The relationship between noxious-response amplitudes and MK was assessed in a whole-brain voxelwise manner using tract-based spatial statistics (TBSS) (Ball et al., 2010; Smith et al., 2006). The TBSS white matter skeleton was generated from the infants' FA maps using a threshold of 0.15 (Ball et al., 2010). Permutation analysis was performed using FSL's PALM with 10,000 permutations, thresholded using TFCE with default parameters (for 2D data), and FWER-corrected for multiple comparisons. Similar to analyses associating noxious-response amplitude to resting-state activity, the TBSS analysis was confound-adjusted for PMA and the three noxious-response imaging confounds. Two additional diffusion imaging confounds were also adjusted for: mean head motion and total brain volume (TBV). The mean head motion values were calculated from the outputs of EDDY, and the TBVs were estimated from the infants' structural image tissue segmentations (see Table 4.1). MK-based quantification of microstructural complexity is influenced by white matter fibre packing density, so adjustment for TBV was included to remove density confounds due to global brain volume variance across infants.

Variability in MK can reflect differences in isotropic tissue barriers caused by variable extracellular matrix complexity and intracellular organelle proliferation, to which the most commonly used diffusion parameter, fractional anisotropy (FA), is insensitive (Suzuki et al., 2003; Jensen and Helpern, 2010). However, variability in MK can also reflect differences in fibre myelination, fibre density, and fibre

dispersion to which the FA parameter is indeed sensitive. To aid the biological interpretation of the MK-based TBSS results, the white matter skeleton region identified as exhibiting statistically significant associations between noxious-response amplitudes and MK was used to perform an ROI analysis for four diffusivity-based parameters: mean diffusivity (MD), fractional anisotropy (FA), axial diffusivity (AD), and radial diffusivity (RD). These diffusivity-based parameters were projected onto the TBSS skeleton for this ROI analysis. Due to the poor association between noxious-response amplitudes and these parameters, the correlation polarity (positive or negative) was tentatively used to facilitate interpretation. For example, a positive correlation between noxious-response amplitudes and both MK and FA would be interpreted as indication increasing response amplitude with increasing fibre density or myelination; and a positive correlation between noxious-response amplitudes and MK coupled with a negative correlation between noxious-response amplitudes and FA would be interpreted as indication increasing response amplitude with increasing fibre dispersion. However, these diffusivity-based parameters have limited sensitivity and interpretability in white matter regions with complex fibre orientation arrangements, such as crossing fibres in the centrum semiovale.

4.3 Results

4.3.1 Quantifying noxious-response amplitudes

The 18 infants' noxious-response maps (regression parameter maps) were combined to produce the group average t-statistic map i.e. the group-level noxious-response map (Figure 4.1A). This group-level noxious-response map (t-statistic map) was spatially regressed onto each infant's noxious-response map (regression parameter map) to produce a spatial regression coefficient representing the amplitude of each infant's response to the nociceptive input. Each infant's noxious-response map and associated amplitude are presented in Figure 4.1B. While 3 of the 18 infants showed negative BOLD responses, the majority of infants (15 out of 18) showed a wide-range of positive responses. Visual inspection revealed differences among infants in the functional topography of the activity pattern, but the dominant feature was differing

response amplitudes. Peristimulus timeseries plots for each infant were visually inspected for HRF goodness-of-fit (Figure 4.2). The fits were deemed appropriate, and there was no visible lag effect or other issue to suggest the positive or negative BOLD responses were inappropriate. The subject-level noxious-response amplitudes in 4.1B were the targets to be predicted from resting-state activity in this study.

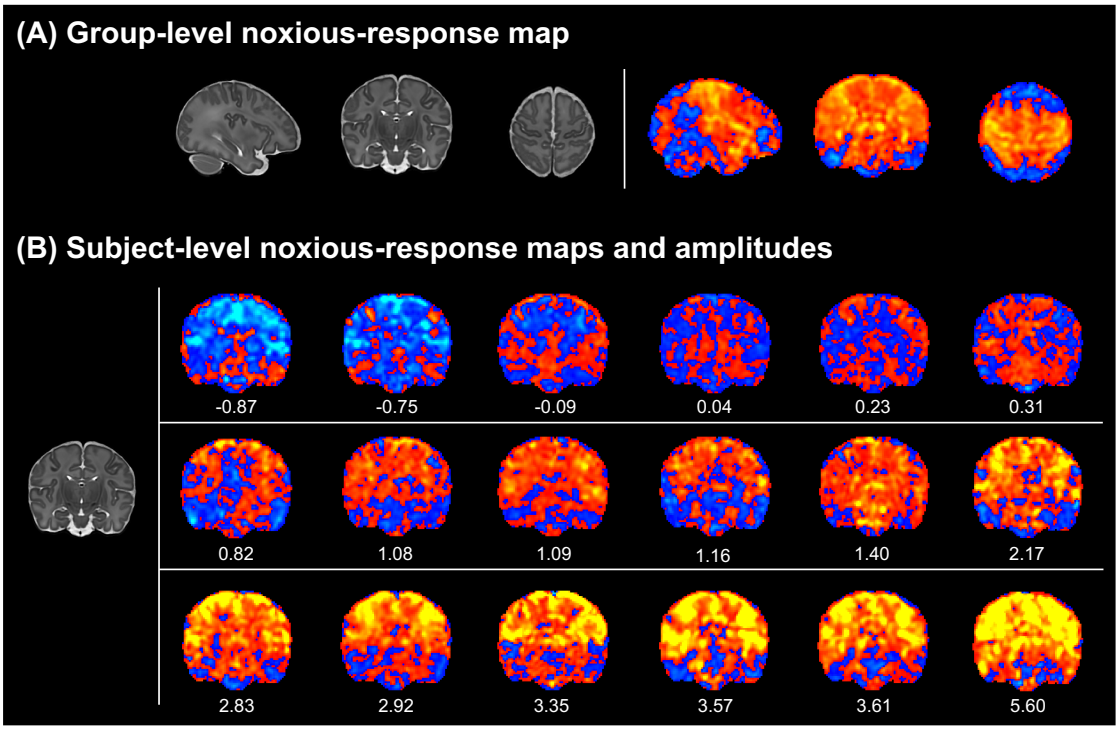

Figure 4.1: Noxious-response maps and amplitudes. For all activity maps, the stimulus was ten trials of a mild experimental pinprick to the left foot. The noxious-response activity was modelled using an infant double gamma HRF. Red-yellow indicates positive values, blue-cyan indicates negative values, and the accompanying grey scale images are the 40-week standard space infant template to provide an anatomical reference. (A) The group-level noxious-response map was the group average t-statistic map. Largest values, and thus strongest weights, were generally localised to primary somatosensory and motor cortices as well as somatosensory thalamus. (B) The subject-level noxious-response maps were the regression parameter maps. The scalar value below each map indicate the infant's noxious-response amplitude, which was quantified as the regression coefficient of the group-level noxious-response map regressed onto the infant's subject-level noxious-response map. Maps are ordered from left-to-right and top-to-bottom based on these noxious-response amplitudes.

4.3.2 Quantifying resting-state network amplitudes

In the PFM analysis of both the local dataset (n = 18) and dHCP dataset (n = 242), a dimensionality of 25 modes was pre-specified. In the dHCP dataset, a

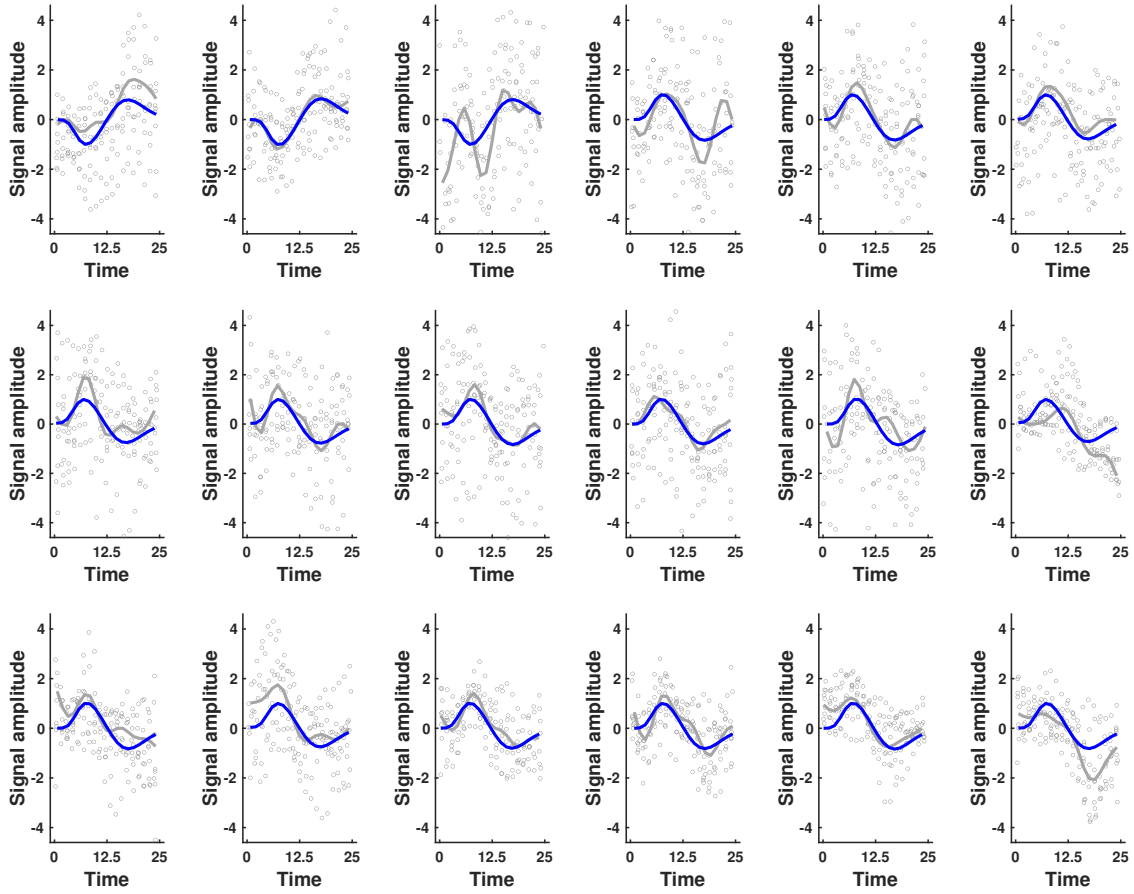

Figure 4.2: HRF peristimulus timeseries plots. The timeseries plots for all 18 infants are displayed in identical order to Figure 4.1B. Timeseries plots follow the FSL FEAT format: timeseries are from a single voxel with largest z-statistic. For all plots, the x-axis (time in seconds) runs from 0 to 25, where t = 0 s is the point of stimulus delivery, and the minimum inter-stimulus interval is 25 s. The y-axis is arbitrarily scaled to blue curve (HRF fit) peak equals one. Light grey circles are individual timepoints from all 10 trials, the solid grey line is the trial average timeseries, and the solid blue line is the HRF fit to the data. Using the noxious-response network amplitudes, infants 1 – 3 were identified as having negative BOLD responses, and infants 4 – 18 were identified as having positive BOLD responses. Infants with substantial BOLD stimulus responses have reasonable HRF fits to the data for both positive and negative BOLD responses. There appears to be no indication that either the negative or positive BOLD responses to the stimulus are an artefact of inaccurate HRF fit to the data.

full 25 modes were identified; in the local dataset, 11 modes were identified. Of these 11 modes, nine were highly consistent with those identified in the dHCP dataset, and these were taken as the set of replicable core neonatal resting-state networks (Figure 4.3A). Pearson correlation coefficients between matched networks were generated using unthresholded maps (mean = 0.78; range = 0.63 – 0.90). Of these nine resting-state networks, three types of sensory and motor networks were identified, which included two visual, two auditory, and two somatomotor networks. In addition, three cognitive networks were identified, which included the default mode, dorsal attention, and executive control networks.

Due to the higher SNR of the dHCP maps, these were used to quantify the nine resting-state network amplitudes for each infant. The nine networks' timeseries were plotted per infant and outlier values highlighted (Figure 4.4). Outliers were present in all infants, motivating the quantification of network activity amplitude using the median absolute deviation (MAD) rather than the standard deviation, due to MAD's greater robustness to outliers. The nine resting-state network amplitudes per infant, displayed in Figure 4.3B, were the resting-state features used to predict the infants' noxious-response amplitudes.

4.3.3 Predicting noxious-response from resting-state

The infants' resting-state network amplitudes were able to predict the noxious-response amplitudes with statistically significant accuracy: RMSE = 1.54, p = 0.001 (Figure 4.5 left). See Table 4.2 for prediction accuracy assessed using R^2 and R_{Sp}. Neither the resting-state imaging confounds nor the non-functional variables were able to predict the infants' noxious-response amplitudes (Figure 4.5 middle and right; Table 4.2).

Using the computationally simpler multiple linear regression mapping approach, results were consistent with the prediction analysis (Figure 4.6). The resting-state network amplitudes were significantly associated with widespread cortical and subcortical regions, while no significant association was observed for either the resting-state imaging confounds or the non-functional variables, even with

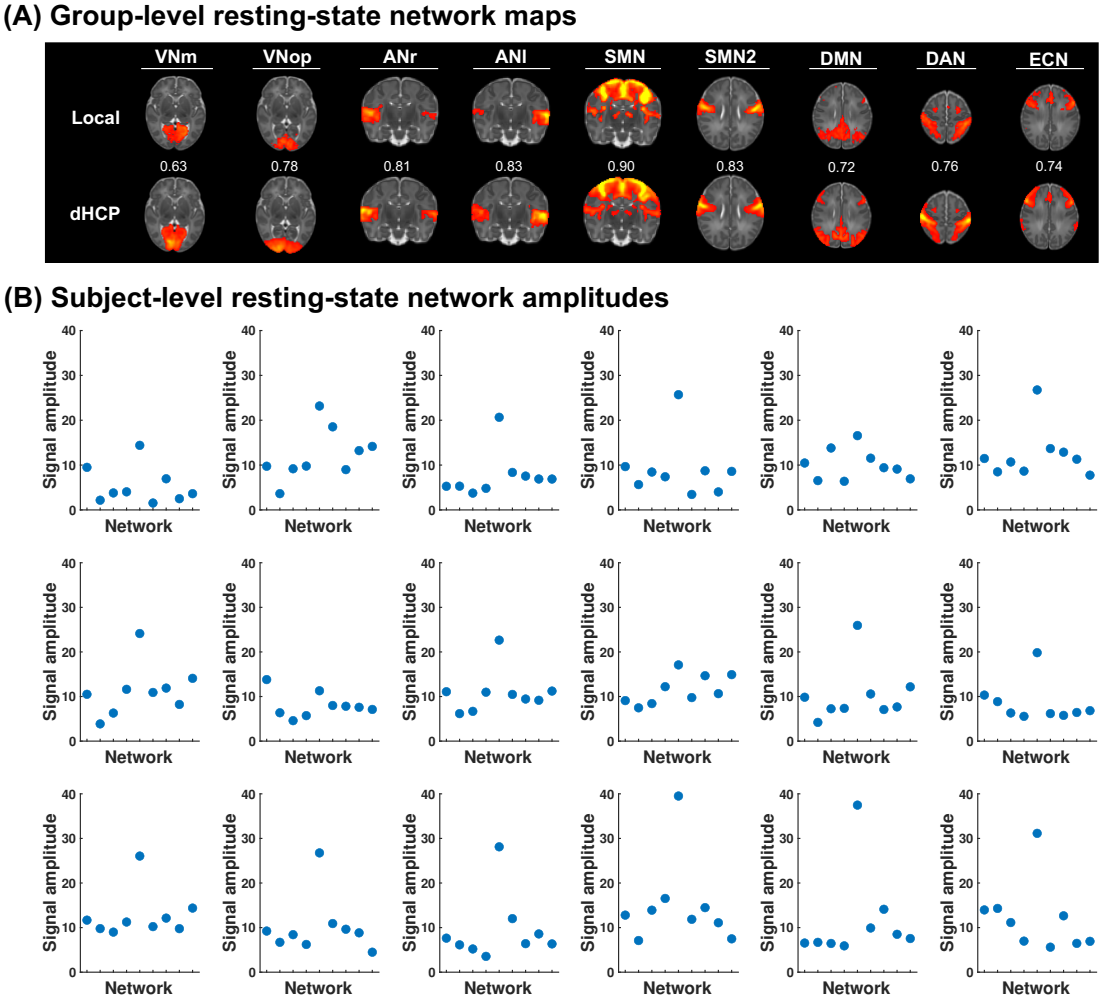

Figure 4.3: Resting-state network maps and amplitudes. (A) Nine resting-state networks were identified across two independent datasets using PFM: "Local" maps were generated using the noxious-response infant cohort (n = 18), "dHCP" maps were generated using a subset of healthy term infants from the dHCP dataset (n = 242). The scalar quantities are the spatial Pearson correlation coefficients between the matched unthresholded maps. Maps are arbitrarily thresholded for illustrative purposes. VNm = medial visual network; VNop = occipital pole visual network; ANr = right auditory network; ANl = left auditory network; SMN = somoatomotor network; DMN = default mode network; DAN = dorsal attention network; ECN = executive control network. (B) Each plot contains the nine resting-state network amplitudes (blue circles) for a single infant, where network timeseries were extracted using the dHCP resting-state network maps as network templates, and amplitudes were quantified as timeseries MADs. In all plots, the x-axis is the nine resting-state networks in identical order to that presented in (A), and the y-axis is the network MAD amplitudes. The 18 plots are arranged in identical order to that in Figure 4.1B.

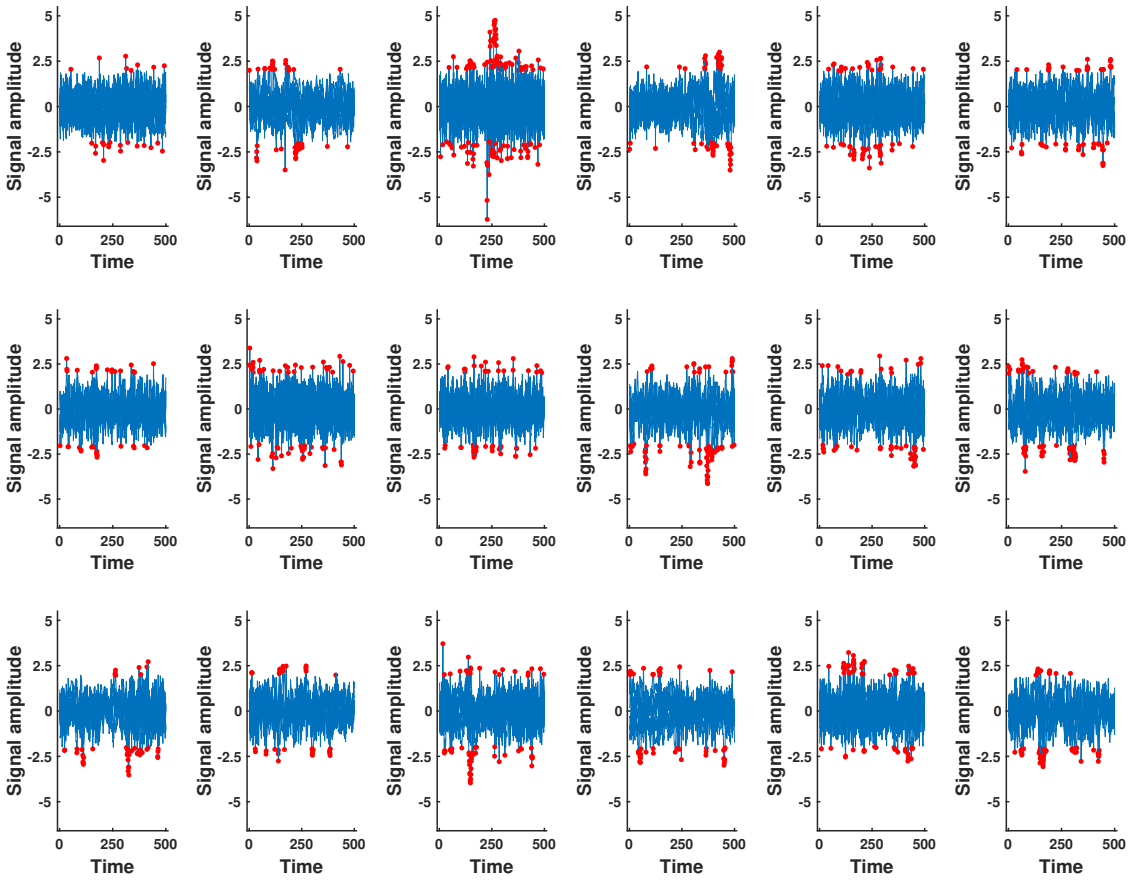

Figure 4.4: Resting-state network timeseries and outliers. The timeseries plots for all 18 infants are displayed in identical order to Figure 4.1B. For each infant's plot, the timeseries for all nine resting-state networks are plotted superimposed (blue), with the x-axis representing time in volumes (500 in total). Each individual network timeseries has been zero-centred around its median and scaled to have unit interquartial range (IQR), with the y-axis representing this centred and scaled signal amplitude. Outliers were identified for each network timeseries individually using the "quartiles method": outliers defined as values more than 1.5 interquartile ranges above the upper quartile or below the lower quartile. Outliers are highlighted in red. All 18 infants' network timeseries included outliers, motivating the use of a measure of data spread that is robust to outliers in order to quantify network activity amplitude. For this reason, resting-state network amplitudes were quantified using the median absolute deviation (MAD), instead of the more commonly used standard deviation.

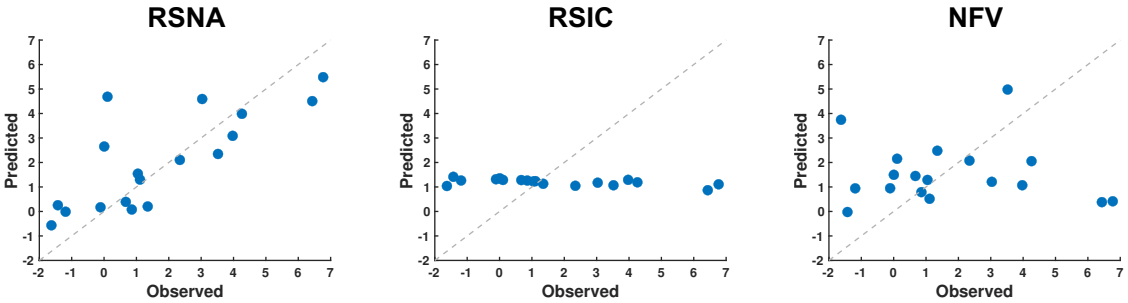

Figure 4.5: Noxious-response amplitude prediction plots. For all plots, each blue dot represents an out-of-sample prediction for a single infant (n = 18). The x-axis is the observed noxious-response amplitudes, the y-axis is the predicted amplitudes, and the grey dashed line is the $y = x$ line along which perfect predictions would lie. Predictions were made using three sets of predictors: resting-state network amplitudes (RSNA), resting-state imaging confounds (RSIC), and non-functional variables (NFV). RSNA included the nine network amplitudes in Figure 4.3. RSIC included mean head motion displacement, mean CSF timeseries amplitude, and mean white matter timeseries amplitude, all extracted from the resting-state scan. NFV included postmenstrual age, gestational age, postnatal age, birth weight, total brain volume, and sex (Table 4.1). Of the three sets of predictors, only the resting-state network amplitudes produced statistically significantly accurate predictions (see Table 4.2).

Table 4.2: Noxious-response amplitude prediction accuracies. Each row is the results for a specific set of predictors: RSNA = resting-state network amplitudes; RSIC = resting-state imaging confounds; NFV = non-functional variables. Each column is the prediction accuracy assessed using a specific metric: RMSE = root mean squared error; R^2 = coefficient of determination (sums-of-squares formulation); R_{Sp} = Spearman's rank correlation coefficient. P-values are presented in parentheses. * = statistically significant. Of the three sets of predictors, only the resting-state network amplitudes produced statistically significantly accurate predictions.

	RMSE	**R^2**	**R_{Sp}**
RSNA	1.54* (0.001)	0.64* (0.001)	0.77* (0.002)
RSIC	2.46 (0.60)	0.081 (0.60)	0.14 (0.40)
NFV	2.71 (0.47)	-0.12 (0.47)	0.0031 (0.51)

the significance level lowered to 0.1. The consistency of these results suggests that the ability to predict noxious-response amplitudes from resting-state network amplitudes was not an artefact of the prediction analysis pipeline, such as leakage between training and test sets.

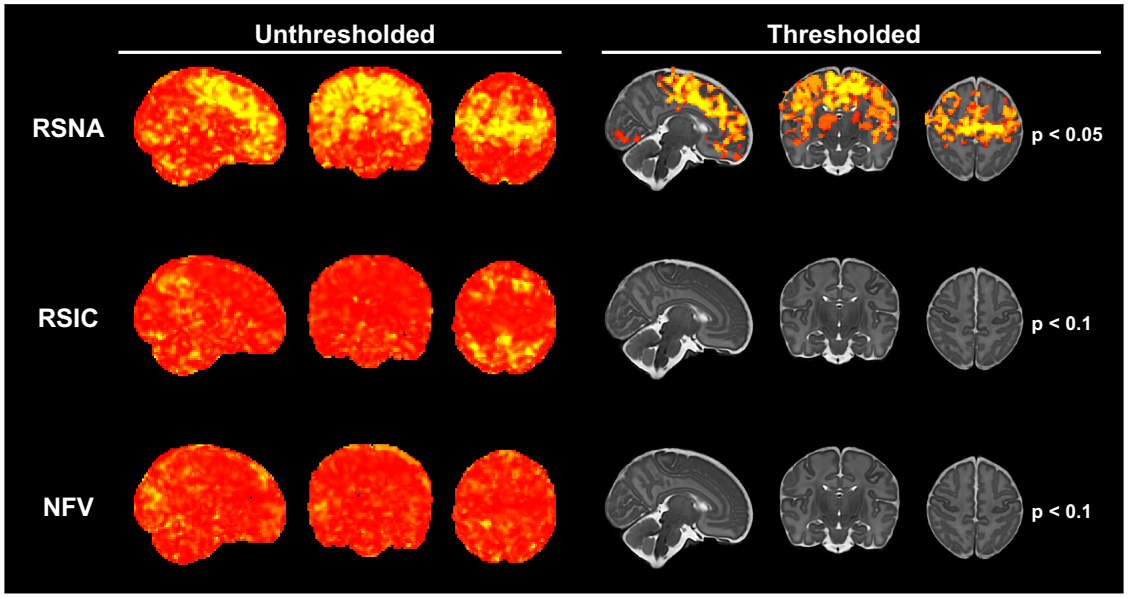

Figure 4.6: Noxious-response amplitude regression maps. The relationship between three sets of predictors and infants' noxious-response maps was assessed via multiple linear regression using NPC. For all maps, red-yellow represents $-log_{10}$(p-values). The three sets of predictors were RSNA = resting-state network amplitudes (row 1), RSIC = resting-state imaging confounds (row 2), NFV = non-functional variables (row 3). The left column displays the spatial patterns of $-log_{10}$(p-values) unthresholded and unadjusted for multiple testing. The right column displays the spatial patterns of $-log_{10}$(p-values) thresholded using TFCE with default parameters and adjusted for multiple testing using FWER. The RSNA p-values are thresholded at the typical significance level of $p < 0.05$, while the RSIC and NFV p-values are thresholded at the more liberal significance level of $p < 0.1$. Despite the liberal threshold for RSIC and NFV, only the RSNA set of predictors resulted in statistically significant associations, which were widespread in both cortical and subcortical regions. These NPC mapping results reflect the SVR prediction results.

Univariate correlation analysis between noxious-response amplitudes and individual resting-state network amplitudes revealed the coupling between noxious-response and resting-state activities was network-specific. Significant positive correlations existed with the somatomotor network (SMN) and the occipital pole visual network (VNop) (Figure 4.7), meaning infants with larger resting-state SMN and VNop amplitudes exhibited larger haemodynamic response amplitudes

to nociceptive input. This network-specificity suggests the prediction results were not a result of an artefactual global signal present throughout the brain in both the resting-state and noxious-response runs. The Pearson correlation results for all nine resting-state network amplitudes are displayed in Supplementary Figure B.2, corrected for multiple testing.

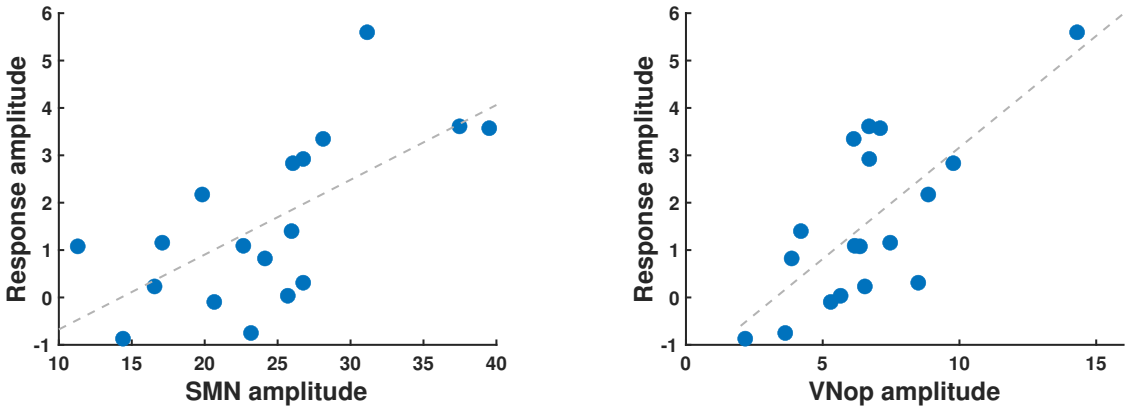

Figure 4.7: Network-specificity in functional coupling between resting-state and noxious-response activities. Of the nine resting-state network amplitudes, only the somatomotor network (SMN) and occipital pole visual network (VNop) were statistically significantly associated with infants' noxious-response amplitudes. All correlations were FWER-corrected for multiple testing. For both networks, strong positive correlations existed: $r = 0.66$ ($p = 0.025$) for SMN, and $r = 0.72$ ($p = 0.0052$) for VNop. Thus, infants with larger SMN and VNop activity amplitudes at rest exhibited larger haemodynamic responses to nociceptive input.

Taken together, the network-specificity of the resting-state network relationship and the lack of a relationship for the resting-state imaging confounds strongly suggest that the newborn infant brain exhibits robust functional coupling between resting-state and noxious-response activities mediated by commonalities in haemodynamic fluctuation amplitudes. The lack of association between noxious-response amplitudes and the non-functional variables means the observed functional coupling cannot be explained by an infant's age, birth weight, brain volume, or sex. The Pearson correlation results for all six non-functional variables are displayed in Supplementary Figure B.3. The p-values presented are uncorrected for multiple testing to clarify that the lack of association was not the result of a statistical correction.

For completeness, the univariate Pearson correlation results between noxious-response amplitudes and individual resting-state imaging confounds are displayed

in Supplementary Figure B.4. Similar to the non-functional variables above, no correlation was statistically significant. The p-values presented are uncorrected for multiple testing to clarify that the lack of association was not the result of a statistical correction.

4.3.4 Relating noxious-response amplitude to white matter microstructure

The infants' noxious-response amplitudes were significantly positively correlated with underlying white matter microstructural complexity (Figure 4.8A). This was established using a whole-brain voxelwise TBSS analysis with microstructural complexity quantified using mean kurtosis (MK). The association was restricted to the primary somatomotor projection tracts of the right centrum semiovale, meaning infants with greater white matter complexity in this region exhibited larger haemodynamic response amplitudes to nociceptive input. This result suggests two points. First, this observed structure-function relationship suggests that the size of an infant's cerebral haemodynamic response to nociceptive input is related to their brain's structural architecture. Second, this association with underlying temporally stable brain structure suggests the noxious-response amplitudes observed in this study were also temporally stable trait effects, rather than temporally transient state effects related to wakefulness and emotional state.

MK values can increase with increasing fibre density and myelination as well as with increasing fibre dispersion. To aid in disambiguating these possible interpretations, the noxious-response amplitudes were correlated with diffusivity-based parameters extracted from the region defined by the association with MK (the region outlined in Figure 4.8A). The four commonly studied diffusivity-based parameters that were assessed were mean diffusivity (MD), fractional anisotropy (FA), axial diffusivity (AD), and radial diffusivity (RD) (Figure 4.8B). While none of the Pearson correlations between noxious-response amplitudes and diffusivity parameters were statistically significant, the polarity of the correlations suggested that the underlying effect may be partially due to increasing fibre density and/or myelination. MD had a negligible correlation coefficient ($r < 0.001$), both FA and AD

had positive correlation coefficients (r = 0.22 and r = 0.18, respectively), and RD had a minor negative correlation coefficient (r = −0.07). Diffusivity-based parameters assessing directionality (FA, AD, and RD) have particularly limited sensitivity and interpretability in white matter regions with complex fibre orientation arrangements, such as crossing fibres. It was suggested that this would be particularly pertinent in the centrum semiovale. Given the region in which the association between noxious-response amplitude and MK exists was the centrum semiovale, the overlap between this region and fibre orientation complexity is explicitly demonstrated in Figure 4.9. The MK-outlined region overlaps sections of the centrum semiovale with a strong presence of multiple fibre orientations, limiting the sensitivity of directional diffusivity measures to possible structure-function relationships.

4.4 Discussion

Newborn infants lack the capacity of verbal communication, and their brief postnatal medical history further limits availability of information to guide effective pain management strategies during their stay in hospital. The high degree of uncertainty regarding an infant's pain experience and vulnerability often results in pain going under-detected and under-treated (Roofthooft et al., 2014). In this study, the central aim was to identify novel features that could be used to predict how an infant responds to a noxious event. Using a cohort of healthy newborn infants, the amplitude of the cerebral haemodynamic response to a mild experimental noxious stimulus was measured using fMRI. Infants' unique spontaneous brain activity, measured using rs-fMRI, and white matter microstructural architecture, measured using dMRI, were analysed to identify novel neural correlates of the noxious-response amplitudes. These data revealed that infants exhibiting larger haemodynamic response amplitudes to nociceptive input had larger resting-state network activity amplitudes in both the somatomotor network (SMN) and occipital pole visual network (VNop). Additionally, larger responses were associated with greater white matter microstructural complexity in primary somatomotor projection

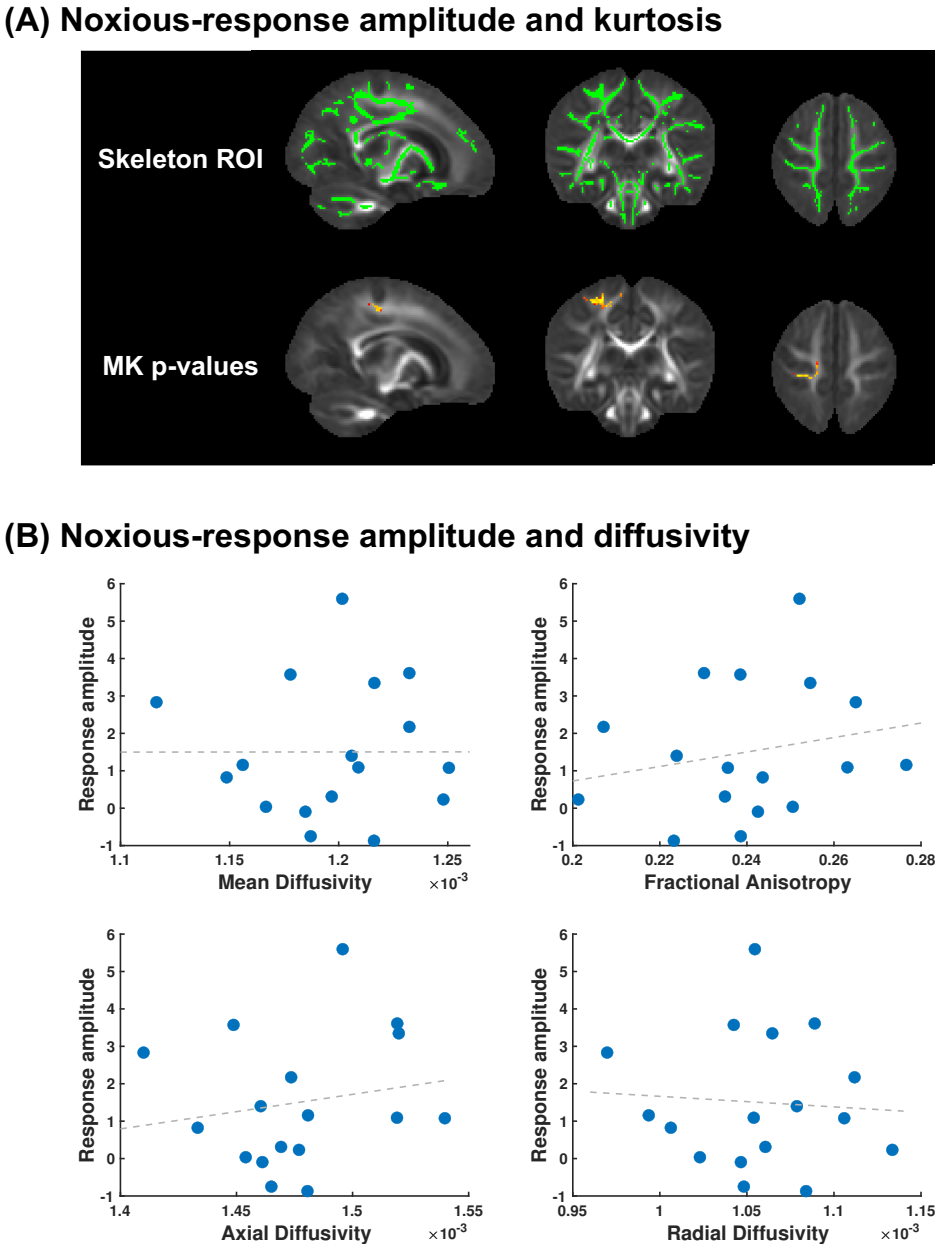

Figure 4.8: Relating noxious-response amplitude to white matter microstructure (n = 17). (A) Infants' noxious-response amplitudes were associated with underlying white matter complexity. Diffusional complexity was quantified using mean kurtosis (MK), and the association was assessed using a whole-brain voxelwise TBSS analysis. The white matter skeleton used in this analysis is presented in green, with the group mean fractional anisotropy map as the grey scale background. The statistically significant positive correlation between MK and noxious-response amplitude was localised to the primary somatomotor projection tracts of the right centrum semiovale. Red-yellow represents $-log_{10}$(p-values) thresholded using TFCE with default parameters and adjusted for multiple testing using FWER. (B) Using the MK-defined white matter region from (A), the association between noxious-response amplitude and four diffusivity-based parameters were assessed using simple Pearson correlation analysis. In all plots, the dashed grey line is the least-squares line. Mean diffusivity had negligible correlation with response amplitude, fractional anisotropy and axial diffusivity had mild positive correlations, and radial diffusivity had a minor negative correlation. See main text for further details.

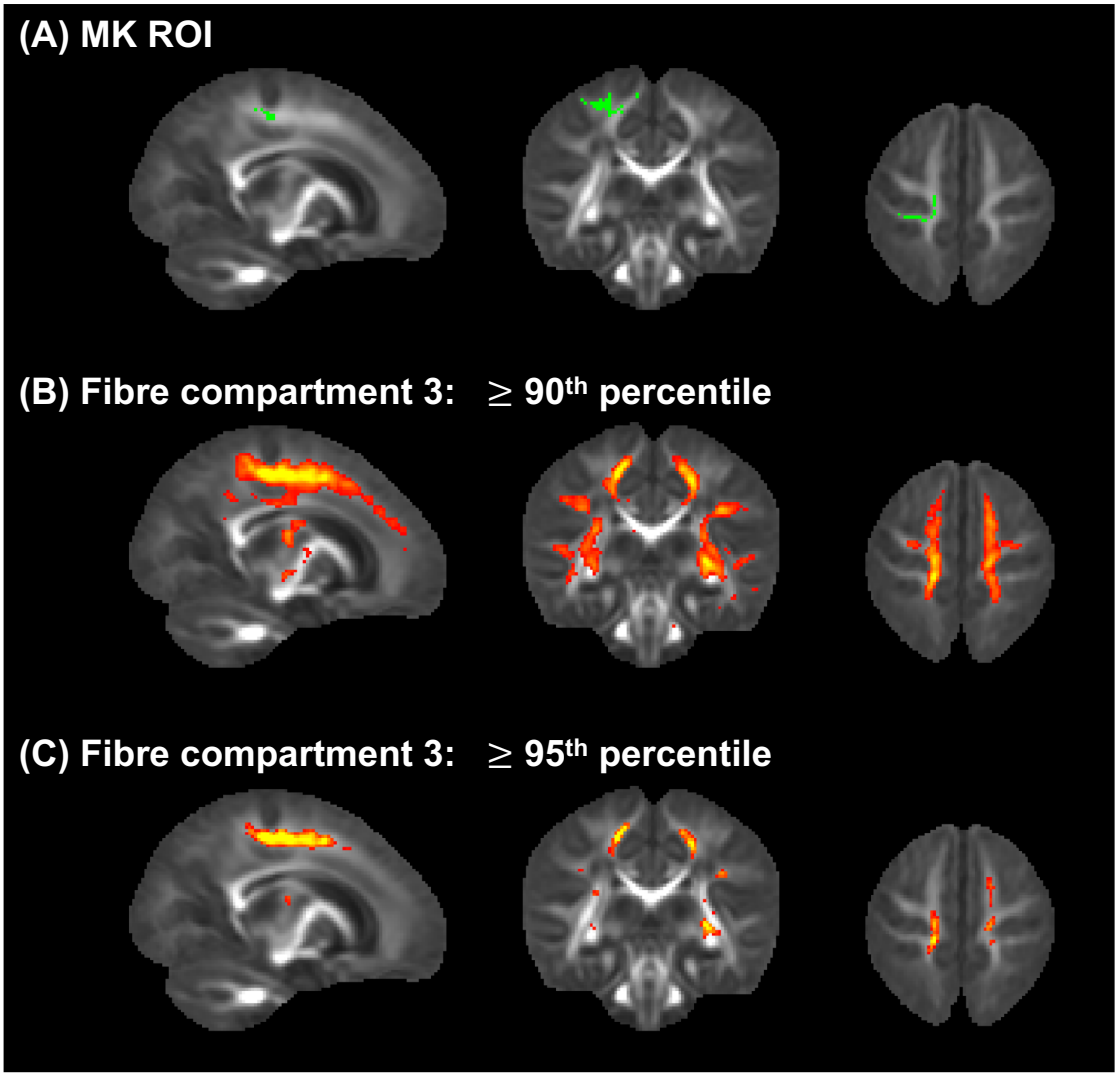

Figure 4.9: White matter fibre orientation complexity. In all images, the grey scale background is the group mean fractional anisotropy map. (A) The green region-of-interest (ROI) is the white matter region in which a statistically significant positive correlation was identified between noxious-response amplitudes and mean kurtosis (MK) (see Figure 4.8A). MK quantifies diffusional heterogeneity or microstructural complexity, and is sensitive to both isotropic and anisotropic diffusional barrier effects. (B and C) Using FSL's ball-and-zeppelin model (Sotiropoulos et al., 2016), three anisotropic fibre compartments were modelled per voxel with fibre compartment 1 modelling the principle diffusion direction accounting for the largest proportion of anisotropic variance. Fibre compartment 3 models the third anisotropic diffusion direction and accounts for the smallest proportion of anisotropic variance. Non-zero volume fractions for fibre compartment 3 are only required in white matter regions with complex fibre orientations with more than two dominant directions. The red-yellow in (B) and (C) are the group mean fibre compartment 3 volume fractions. These have been thresholded at the 90^{th} (B) and 95^{th} (C) percentile to highlight the regions of white matter with the greatest fibre orientation complexity. The MK ROI in (A) partially overlaps with regions of the centrum semiovale with the greatest fibre orientation complexity (C).

fibres of the right centrum semiovale. Most importantly, the infants' noxious-response amplitudes were predictable from their resting-state activity, where resting-state activity was recorded in a nociception-free manner. The ability to predict an infant's haemodynamic response to nociceptive input has the potential to advance the improvement of neonatal pain management, and this study highlights the importance of understanding resting-state brain activity in achieving this goal. An infant's unique stimulus-independent resting-state activity may have the potential to predict further valuable features (Parker Jones et al., 2017; Tavor et al., 2016), such as pain-related behavioural and physiological responses, that could greatly inform the clinical decision-making process regarding pain management strategies for newborn infants.

The correspondence between resting-state and task-evoked activities has been demonstrated in adults over a range of functional networks and tasks (Cole et al., 2014; Smith et al., 2009). While similar studies in neonates are lacking due to the inherent limitations in experimental design of infant task fMRI, the resting-state networks commonly observed in neonates, and those observed in this study, bear close resemblance to adult networks (Doria et al., 2010; Fitzgibbon et al., 2019; Fransson et al., 2007; Fransson et al., 2009), suggesting a similar correspondence exists in the infant brain. In the current experimental nociception paradigm, robust correlations were identified between the infants' noxious-response amplitudes and a specific subset of resting-state network amplitudes, the SMN and VNop. This network-specificity is consistent with a recent study in neonates (Goksan et al., 2018) that examined the relationship between noxious-response amplitude and functional connectivity of three networks in the immediate prestimulus period. The study observed a strong correlation between response amplitude and the prestimulus functional connectivity of the descending pain modulatory network, but no such relationship with the default mode network or with a 'negative control' network designed to test global brain effects. Similar studies performed in adults have also reported region-specific relationships between spontaneous brain activity in the immediate prestimulus period and trial-to-trial variability in pain perception (Boly et al., 2007; Ploner et al., 2010).

While the current study reveals the close functional coupling between resting-state and noxious-response brain activities in newborn infants that reflects this coupling observed in adults (Bzdok et al., 2016; Cole et al., 2014; Smith et al., 2009; Tavor et al., 2016), there are important differences between adult and infant cerebral haemodynamics which need to be considered when interpreting infant cerebral haemodynamic activity. The neural activity to which the haemodynamics are responding are dominated by electroencephalographic features unique to newborns in the preterm period, such as discontinuous neuronal bursting activity and the delta brush waveform (André et al., 2010; Arichi et al., 2017), and some of these features are typically still present, but to a much lesser degree, in the peri-term period (André et al., 2010). In the foetal brain, the dominant driver of spontaneous brain activity is peripheral in origin, rather than centrally generate as in adults (Colonnese and Khazipov, 2012). While centrally generated activity is present in the peri-term infant brain, similar to the presence of neonate-specific EEG features, the switch to adult-like resting-state neurodynamics is likely incomplete. Regarding the fMRI BOLD signal specifically, the developmental state of the newborn's neurovascular coupling mechanism and cerebral haemodynamic response results in immature functional hyperaemia and neonate-specific haemodynamic response waveforms (Arichi et al., 2012; Kozberg et al., 2013; Kozberg and Hillman, 2016).

Given these similarities and differences between the adult and infant cerebral haemodynamic response, interpreting the cause of the observed normative variability in noxious-response amplitude in this healthy sample of newborn infants is challenging. While a maximum of three of the infants displayed negative BOLD responses, the majority of the infants displayed positive BOLD responses. The infant double gamma HRF used in this study (Arichi et al., 2012; Baxter et al., 2019) appeared to fit the data appropriately, and the observation of positive BOLD responses being dominant in the peri-term age is consistent with observations in other groups using tactile but non-noxious experimental paradigms (Arichi et al., 2012). This is inconsistent with rodent studies using postnatal day 12 (P12) rats as rodent models of newborn term-aged infants (Kozberg et al., 2013),

which suggest that the newborn infant's BOLD response should be predominantly negative due to immature functional hyperaemia. However, it is worth noting that the haemodynamic response waveform reported in this rodent study for P15-18 rats closely resembles the biphasic neonatal BOLD HRF, suggesting the inconsistency between the human fMRI and rodent optical imaging results may be due to cross-species differences in the relative developmental trajectories of different physiological processes. Another potential confound to interpretation is that the BOLD response amplitude in neonates may reflect variability in systemic blood pressure effects due to functionally immature cerebral autoregulation (Greisen, 2009), where a stimulus-induced blood pressure increase could cause an immature negative BOLD response to resemble the adult-like positive BOLD response (Kozberg et al., 2013). However, the stimulus intensity required to induce this autoregulatory confound in a healthy peri-term-aged infant would be expected to cause a behavioural and physiological response "equivalent to a startle reflex that makes a newborn cry" (Hillman, 2014). While blood pressure recordings were not possible to collect in the current study, the low-intensity experimental noxious stimulus used does not cause behavioural or physiological (heart rate and blood oxygen saturation) distress, suggesting the observed haemodynamic responses are not contaminated by autoregulatory confounds.

Interpreting the cause of the observed noxious-response amplitude variability in isolation may not be possible. Fortunately, the flexibility of the MRI modality allowed the identification of several novel neural correlates of the noxious-response amplitudes, functional correlates using rs-fMRI and structural correlates using dMRI, which provide valuable insight. The functional correlates observed in this study included robust positive correlations between noxious-response amplitude and resting-state network activity amplitudes for both the SMN and VNop. While the correlation with resting-state SMN activity is unsurprising and may indicate a direct causal link, the correlation with resting-state VN activity was unexpected. This association between resting-state visual activity and noxious-evoked activity warrants further investigation. We do not believe the correlation is indicative of

a direct causal link in a manner akin to the SMN relationship, but may instead be a downstream consequence of a deeper common cause. (But why this potential underlying common cause would be limited to SMN and VNop is, again, an interesting question requiring further investigation.) In general, the developmental trend from infancy to adulthood observed using fMRI is increasing haemodynamic response amplitude (Arichi et al., 2012). Using NIRS to measure haemodynamic responses to pain in the perinatal period, the amplitude of these responses also progressively increases with age (Slater et al., 2006). And studies using fMRI to observe the developmental progression of resting-state activity have also found increased SMN and VN (as well as other networks) functional connectivity strength and activity amplitude with increasing age (Fitzgibbon et al., 2019; Smyser et al., 2010). The structural correlate observed in this study was a positive correlation between response amplitude and white matter microstructural complexity (mean kurtosis, MK) in primary somatomotor projection fibres of the right centrum semiovale. Similar to the neural functional measures, white matter MK increases with development, with MK observed to increase over the first two years (Paydar et al., 2014), and to increase with age in the white matter of primary projection bundles over the peri-term age-range (Bastiani et al., 2019). Taken together, the resting-state and white matter correlates suggest the noxious-response amplitude is a reflection of neural maturity, with larger response amplitudes indicating a more mature brain and negative and negligible response amplitudes indicating a more immature brain. This neural maturity hypothesis is appealing as it is consistent with the claim that the infants' noxious-response amplitudes are temporally stable trait effects due to their relationship with underlying microstructure. This also suggests a plausible explanation as to why the infant's resting-state activity is predictive of the noxious-response amplitudes: both the activity levels recorded at rest and in response to stimulation are a function of cerebral maturity and thus closely reflect each other due this common underlying cause.

Assuming increasing age is a reasonable proxy for increasing maturity, one can view an infant's age as an imperfect indicator of that infant's neural maturational

state. It is clear that two perfectly age-matched infants would not be expected to be perfectly matched for maturational state, due to individual variability in genetics and environmental influence. It is worth noting this important distinction between age and maturity when considering three key aspects of the present study: the sample used had a relatively narrow age-range (at time of scan, PMA range was 40 days, and PNA range was 10 days), there was no clear relationship between noxious-response amplitude and brain volume or age (PMA, GA, or PNA), and statistical associations were adjusted for PMA to correct for age-related MRI imaging artefacts. The results of the current study then suggest that an infant's unique MRI measures allow us to detect individual variability in the maturity of the structural and functional neural architecture of an infant's brain, which is not fully captured by proxy indicators of neural maturity such as age or brain volume.

The ability to use non-nociceptive data to predict the amplitude of an infant's cerebral haemodynamic response to nociceptive input offers an exciting proof-of-concept that an infant's unique spontaneous brain activity may be predictive of other behavioural and physiological pain-related responses. It will be interesting to see how the novel brain-based structural and functional measures identified in this study correspond to pain-related signals recorded using EEG and NIRS, which rely on thalamocortical neural circuits, as well as pain-related behaviours such as limb withdrawal reflexes and facial expressions, which rely predominantly on spinal and brainstem neural circuits (Fitzgerald, 2015). While it is likely that multimodal models incorporating EEG, behavioural, and autonomic measures may provide predictive information with greater clinical utility and feasibility than MRI (Vaart et al., 2019), the ability to identify measures of both structural and functional neural architecture in a whole-brain high spatial resolution manner means MRI will play a central role in furthering our basic understanding of the neurophysiological basis of infant pain.

4.5 Acknowledgements

We are grateful for the provision of simultaneous multi-slice (multi-band) pulse sequence and reconstruction algorithms from the Centre for Magnetic Resonance Research, University of Minnesota.

The dHCP resting-state network data were provided by the developing Human Connectome Project, KCL-Imperial-Oxford Consortium funded by the European Research Council under the European Union Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement no. [319456]. We are grateful to the families who generously supported this trial.

5
Thesis summary

5.1 Aims and outcomes

The purpose of the work in this thesis was to improve our understanding of the individual variability in newborn infants' response to nociceptive input. Infant pain scales are used in clinical practice to assess an individual's pain experience and sensitivity in order to inform selection of appropriate pain management strategies. Unfortunately, these scales rely predominantly on ambiguous behavioural and physiological measures. The lack of verbal communication and limited behavioural repertoire of newborn infants, the brief extra-uterine time window for establishing a medical history, and the ambiguity of infant pain scales introduce considerable uncertainty into the clinical decision making process. This inevitably leads to pain going under-detected and thus under-treated (Roofthooft et al., 2014). Understanding the neurophysiological basis of infants' response variability to nociceptive input will undoubtedly aid in improving these scales and improving personalised pain management for the newborn infant.

The work in this thesis involved studying a healthy cohort of newborn infants, using a mild experimental 128 mN pinprick stimulus, and recording infants' unique brain structure and function using MRI. The pinprick stimulus was used as it closely mimics the common heel lance blood test procedure, and evokes a noxious-specific ERP also seen in response to the heel lance (Hartley et al., 2015; Hartley et al., 2017). However, unlike the heel lance, the pinprick stimulus is non-skin breaking, does not cause any behavioural or physiological distress to the infant, and is thus not of clinical or ethical concern. Infants' brain structure and function were studied

using multiple MRI modalities to capitalise on the many unique features that MRI provides over other brain-based measures, such as EEG and NIRS, and non-brain-based measures, such as behavioural and physiological recordings. The key advantages of MRI as a method for studying infant pain include two major points. First, similar to EEG and NIRS, fMRI measures brain activity, overcoming many of the ambiguities associated with behavioural and physiological measures. Second, fMRI uses magnetic gradients to actively probe the infant brain, unlike the passive recordings of EEG and NIRS, allowing unparalleled acquisition of spatial information that results in unrivalled image reconstruction quality. MRI allows whole-brain coverage at relatively high spatial resolution, with signals accurately localised to the brain regions of origin. These properties allow the generation of activity maps with relatively high biological interpretability and facilitate the use of a wide range of sophisticated analysis methods. This spatial information is valuable due to the hierarchical organisation and functional segregation of the brain, with the resulting collection of networks subserving the multiple dimensions of pain being spatially distributed. Unfortunately, studying infant pain and nociception using MRI has several challenges and limitations. The lack of data analysis methods to deal with the infant-unique brain features, such as differences in tissue contrast and haemodynamic response profiles to adults, as well as the excessive subject head motion, typically results in data quality too poor to allow meaningful in-depth analysis.

As outlined in Chapter 1, there were three central aims to the work in this thesis. First, outlining and validating an optimised infant noxious-response fMRI data analysis pipeline to achieve high subject-level data quality. Second, assessing the potential to use an infant's unique resting-state brain activity to predict the size of the cerebral haemodynamic response to nociceptive input. Third, identifying novel functional and structural neural correlates of an infant's haemodynamic response to nociceptive input to improve our understanding of the neurophysiological basis of individual variability in response amplitude. These aims were achieved across two studies detailed in Chapters 3 and 4.

The first objective was addressed in Chapter 3, with the data analysis methods carried forward to Chapter 4. To optimise data analysis and improve subject-level data quality, five key stages of data analysis were studied: motion and distortion correction, spatial normalisation, FIX denoising, spatial filtering, and HRF modelling. Adopting cutting-edge motion and distortion correction preprocessing resulted in improved data quality through reduced motion and distortion contamination. Additionally, the adoption of FIX denoising resulted in the most dramatic improvement in data quality, even in instances of stimulus-correlated head motion. The spatial normalisation pipeline improved subject-subject alignment through the incorporation of BBR, and subject-template alignment through the use of age-appropriate templates and the incorporation of ANTs's SyN. The identification of an appropriate spatial filtering extent balanced gain in sensitivity to signal with loss of spatial specificity. Similarly, identification of a HRF model parameterisation that was optimal for the nociception experimental paradigm balanced over- and under-fitting of the infant cerebral haemodynamic response. While accomplishing this goal was a vital first step to the work in this thesis, two broader goals were achieved. Optimisation of the initial fMRI analysis steps for infant nociception studies has now resulted in a pipeline that can be readily adopted in future fMRI work. The successful adoption and extension of the dHCP preprocessing pipeline demonstrated that the dHCP pipeline was generalisable to non-dHCP datasets, was generalisable to non-resting-state datasets, and allowed for many additional validation tests that could be performed on a "task" fMRI dataset that could not be performed on the dHCP rs-fMRI dataset.

The second objective was addressed in Chapter 4. Each infant's resting-state brain activity was used as a source of infant-specific neural features, the acquisition of which does not involve concomitant or prior pain exposure, to predict an infant's haemodynamic response amplitude to nociceptive input. Prior to assessing each infant's resting-state activity, a highly robust set of nine networks were replicated in an independent dataset, an age-matched subset of the dHCP dataset. The relationship between an infant's resting-state network amplitudes and the size of the

infant's cerebral haemodynamic response to nociceptive input was demonstrated at both the voxel and whole-brain levels. This demonstrated the functional coupling between infants' resting-state and noxious-response brain activities. Importantly, the infants' resting-state network amplitudes were sufficiently information-rich to allow subject-level predictions of an infant's noxious-response amplitude. Subsequent structure-function analysis revealed a relationship between an infant's noxious-response amplitude and underlying white matter microstructure, suggesting that the individual infants' noxious-response amplitudes were temporally stable trait effects. Together, these results suggested that the fMRI-measured and trial-averaged amplitude of an infant's cerebral haemodynamic response was a trait effect that was predictable from resting-state activity. This provides a proof-of-concept that an infant's spontaneous brain activity could potentially be used to predict infant-specific features of clinical value, from brain-related activity to physiological and behavioural outcomes. Being able to predict how an infant may respond to a clinically-required painful procedure could help inform clinical decision making regarding appropriate and personalised pain management strategies.

The third objective was also addressed in Chapter 4. To improve our understanding of the neurophysiological basis of individual differences in the newborn infant's cerebral haemodynamic response amplitude to nociceptive input, the statistical significance and polarity of correlations with an infant's structural and functional neural architectural features were assessed. Assessing the relationship between noxious-response amplitude and resting-state network amplitudes, the functional coupling was observed to be network-specific. Positive correlations existed with the SMN and VNop amplitudes, such that infants with greater activity amplitudes in these networks exhibited a larger (more positive) response amplitude to the nociceptive input. Assessing the relationship between noxious-response amplitude and white matter MK, the structure-function relationship was well localised to a functionally relevant white matter bundle. A positive correlation existed within the primary somatomotor projection tract of the right centrum semiovale, such that infants with greater MK in this region exhibited a larger (more positive) response

amplitude to the nociceptive input. Mild non-significant positive correlations between noxious-response amplitude and fractional anisotropy and axial diffusivity suggested that the (significant) relationship with MK may be partially driven by myelination and fibre density effects. Based on the developmental trends in resting-state network activity amplitudes, white matter MK, and fibre density and myelination that have been reported in previous studies, it appeared that the individual variability in noxious-response amplitude was primarily due to individual variability in brain maturity. It was thus hypothesised that infants with more functionally and structurally mature brains exhibited larger cerebral haemodynamic responses to nociceptive input. The lack of a significant relationship between noxious-response amplitude and age (as well as brain volume, birth weight, and sex) thus highlight the value in using infant-specific structural and functional neural features to provide greater insight into the infant's processing of pain and nociception. While MRI lacks clinical feasibility and is one of the most challenging modalities to use to assess infant pain, its unique ability to provide whole-brain coverage and spatially rich insight into infant brain structure and function, as demonstrated in this thesis, means MRI will play a vital role in improving our understanding of the infant pain system.

5.2 Current limitations and future possibilities

The current thesis findings and associated limitations provide a platform from which future work can be based. The major limitations of the two studies in this thesis were considered in their respective discussion sections (Sections 3.4 and 4.4). These are summarised here to highlight common themes and to serve as an opportunity to discuss potential solutions and future work. The limitations and future work are addressed under five headings: preprocessing optimisation, biological interpretability, age and modality generalisability, heel lance and brain maturity assumptions, clinical utility assumption. The topics covered in the first two sections broadly cover desirable future improvements and extensions to the

present work, while the topics covered in the last three sections suggest novel future studies as natural continuations of the present work.

5.2.1 Preprocessing optimisation

In Chapter 3, an optimised pipeline for the earliest stages of fMRI data analysis was outlined and tested. While the optimisations proved to be highly valuable at the time of writing, it is important to note that "optimal" is a moving target. This pipeline and indeed others developed henceforth should not be considered static, but subject to continued improvement as the field progresses. As discussed in Section 3.4, limitations for each pipeline step exist. Preprocessing steps are discussed here, while the HRF modelling step is discussed in Section 5.2.3 below. In brief, the motion and distortion correction performed using FSL's EDDY is currently not optimised for fMRI data. The implementation of ANTs's SyN for spatial normalisation has not been optimised for the infant standard templates. The projection of fMRI data from volumetric space to surface space for reducing alignment errors, partial voluming effects, and spatial filtering issues has yet to be included in the pipeline. While different spatial filtering extents were tested in this study, the use of this preprocessing step depends on both the data and the research question, such that adoption of the particular filtering heuristic (of 1.5 $\times$ voxel size) should be limited to appropriate contexts. In studies using data projected to the cortical surface, spatial filtering should only be applied after the projection. The use of FSL's SUSAN would no longer be appropriate in these instances, and so filter extents smaller than 1.5 times voxel size using a simple Gaussian kernel would be a possibility.

Several fMRI preprocessing pipeline limitations are currently being worked on as part of the dHCP, including motion and distortion correction, FIX denoising, spatial normalisation, and surface projection (Fitzgibbon et al., 2019). It will be imperative to keep abreast of cutting-edge methods, and to incorporate improvements into the analysis pipelines as they arise. If data are to be re-used in future studies, improving data preprocessing will entail re-analysing the same data on multiple occasions. To reduce workload, ensure reproducibility across pipeline versions, and facilitate

the incorporation of novel methods, it would be ideal to adopt a standardised data format. Currently, the data structure is organised in a bespoke manner, but adopting the Brain Imaging Data Structure (BIDS) standard (Gorgolewski et al., 2016) would provide long-term benefits outside of adopting best preprocessing practices, such as the facilitation of data sharing.

5.2.2 Biological interpretability

In Chapter 4, three MRI modalities were used to study the newborn infants' structural and functional brain architectures: noxious-response fMRI, rs-fMRI, and dMRI. For all three modalities, neural features were extracted and given a biological interpretation. However, in all cases, the interpretation is ambiguous. While biologically ambiguous features can be invaluable in aiding appropriate clinical decision making, such as the currently available ambiguous infant pain scales used in infant pain assessment and management, it will ultimately be improvements in accurate understanding of the neurophysiological mechanisms of infant pain that will improve infant pain management. It is thus a central goal of infant pain research, and neuroscience in general, to improve our understanding of the infant brain, which involves identifying and resolving these ambiguities.

The noxious-response amplitudes extracted from the noxious-response fMRI data were derived by spatially regressing a group-level noxious-response map (group average t-statistic map) onto each infant's subject-level noxious-response map (regression parameter map). The pinprick stimulus used to evoke the haemodynamic response causes a multidimensional experience in adults that involves a wide range of brain regions. The stimulus is processed by the newborn infant brain using an equally wide range of brain regions (Goksan et al., 2015), suggesting a comparable range of stimulus response dimensions. These are likely to include sensory discrimination, saliency, attention, mild negative emotional valance, and nocifensive motor signals, among others. It is unclear which neural processes are represented by the noxious-response amplitudes used in this study, and disambiguating these multiple dimensions will be challenging in this non-verbal population. However, it is

possible to develop an alternative principled approach for noxious-response feature extraction that could decrease this ambiguity and improve biological interpretability. There are now several candidate fMRI 'neural signatures' for distinct components of adult pain and negative affect that could be applied to infant noxious-response data (Chang et al., 2015; Kragel et al., 2018; Krishnan et al., 2016; Wager et al., 2013; Woo, Krishnan, and Wager, 2014). While the limitations of this reverse inference approach will need to be carefully considered, applying these templates, which have been developed and validated in verbal populations, to assess brain activity in non-verbal infant populations has the potential to answer neuroscientific questions that would not be possible using infant data-derived templates alone.

The resting-state network amplitudes extracted from the rs-fMRI data were derived using a Dual Regression style approach, involving spatial regression of group-level maps onto subject-level timeseries data. The Dual Regression approach can confound the spatial, temporal, and amplitude effects of a subject's resting-state network. This results in subject-level Dual Regression measures of, for example, functional connectivity being heavily influenced by both amplitude (Duff et al., 2018; Friston, 2011) and spatial topography (Bijsterbosch et al., 2018; Bijsterbosch et al., 2019) effects. Thus, the resting-state network amplitudes used throughout this thesis will also include these ambiguities to some degree, limiting the biological interpretability of these resting-state features. Improvements in subject-level resting-state network modelling is needed, such as the PFM model implemented in FSL's PROFUMO (Harrison et al., 2015; Harrison et al., 2019). At the time of writing this thesis, PROFUMO had not been optimised for use with infant resting-state fMRI data. While PROFUMO produced highly reproducible group-level maps that were used throughout this thesis, the subject-level SNR penalties in the PFM model meant that the identification of certain networks failed in infants with poor SNR. This prevented the use of subject-level PFM-estimated network amplitudes in the resting-state analyses of this thesis. As described above, the group-level PFM maps were instead used in a Dual Regression style analysis, which guaranteed network amplitude estimates for all networks in all infants. Work is currently

underway to resolve these issues in PROFUMO. It would be valuable to revisit the resting-state analyses in this thesis in the future to take full advantage of the superior separation of spatial pattern and amplitude variation, and improved BOLD timeseries estimation, provided by PROFUMO. If the infants' PFM-estimated network amplitudes are also predictive of noxious-response amplitudes, this would strongly corroborate the current interpretation.

Finally, the white matter microstructural features extracted from the dMRI data were derived using the DKI model. The DKI model (like the DTI model) is a statistical model, which has the advantage of lacking any dependency on biophysical model assumptions, and as a result, the model parameters are physically well-defined. The primary diffusion parameter studied in this thesis was MK, which is simply "three times the square of the coefficient of variation for the distribution of compartmental diffusion coefficients. So qualitatively, the kurtosis is a measure of the heterogeneity of the diffusion environment." (Jensen and Helpern, 2010). MK is thus a measure of diffusional complexity that is sensitive to a multitude of microstructural properties, including both isotropic and anisotropic barrier effects. While MK's lack of dependency on biophysical model assumptions as well as its sensitivity to both isotropic and anisotropic barrier effects are attractive qualities and were motivations for selecting it as the microstructural parameter of primary interest, it is precisely these properties that limit biological interpretability. Complementing analysis of MK using diffusivity measures such as mean diffusivity and fractional anisotropy can aid interpretation, but this approach has its limitations, as was observed in Chapter 4. The adoption of biophysical models such as NODDI (Neurite Orientation Dispersion and Density Imaging) (Zhang et al., 2012) could improve the biological interpretability of the dMRI results. NODDI models the dMRI signal as a sum of restricted (intra-neurite), hindered (extra-neurite), and isotropic (CSF) pools of water, and also estimates an orientation dispersion index to account for fibre fanning. The explicit modelling of neurite density and dispersion could greatly improve the biological interpretability of the microstructural features underlying the relationship between noxious-response amplitude and diffusional complexity.

5.2.3 Age and modality generalisability

The identification of the double gamma HRF parameterisation as optimal in Chapter 3 and the observed functional coupling between resting-state and stimulus-response brain activities in Chapter 4 have limited generalisability due to the use of a single stimulus modality i.e. the 128 mN pinprick. In brief, the suitability of using the infant double gamma HRF was tested quantitatively at the group-level in Chapter 3 and also qualitatively at the subject-level in Chapter 4. While the double gamma HRF was found to be suitable for the noxious stimulus and experimental paradigm used throughout this thesis, the generalisability of this finding to alternative stimuli and paradigms is unclear. Additionally, the tight functional coupling between resting-state network amplitude and noxious-response amplitude was demonstrated at both the voxel and whole-brain levels in Chapter 4. While this functional coupling has been demonstrated for a wide range of tasks in adults, the findings in this thesis are currently limited to the haemodynamic noxious-response amplitude evoked by the sharp-touch nociceptive stimulus. A novel study in which both resting-state fMRI data and multimodal stimulus-response fMRI data are collected for each individual infant could resolve this limitation. While "task" fMRI experimental designs are severely limited in newborn infants, previous studies using non-nociceptive stimuli, such as non-noxious touch (Arichi et al., 2010), auditory (Anderson et al., 2001), and visual (Deen et al., 2017) stimuli have demonstrated the feasibility of a multimodal experimental design.

Additionally, the HRF parameterisation and functional coupling results have limited generalisability to premature infants, as the youngest infant included in this cohort was 35.9 weeks PMA at the time of study. A multimodal stimulus-response fMRI study would have to be conducted in a cohort of premature infants to test these findings in younger infants. However, this challenging experimental paradigm may lack feasibility in this fragile population. The use of a cot-side modality such as EEG may be more appropriate, from both the perspective of experimental feasibility and clinical utility, in order to examine the functional coupling between resting-state and stimulus-response activities.

5.2.4 Heel lance and brain maturity assumptions

Throughout this thesis, the experimental pinprick stimulus was used to probe an infant's response to nociceptive input due to the pinprick's similarity to the clinically required skin-breaking heel lance procedure. The pinprick delivers sharp tactile sensory input to the foot and evokes an EEG-recorded noxious-specific ERP at a lower amplitude to that seen in response to the heel lance (Hartley et al., 2015; Hartley et al., 2017). Due to the mild nature of the pinprick stimulus, the challenges in performing simultaneous EEG-fMRI recordings, and the challenges in performing clinical procedures in the MRI environment, it is not trivial to establish a relationship between the individual variability observed in the fMRI-recorded cerebral haemodynamic response amplitude to the pinprick and individual variability in pain-related behavioural, physiological, or electroencephalographic responses to the heel lance. However, establishing relationships such as these will be central to validating the assumption that MRI-identified neural correlates of the pinprick response do in fact apply to infants' responses to the heel lance procedure. In Chapter 4, three such novel MRI-identified neural correlates were outlined. The individual variability in cerebral haemodynamic response amplitude to the pinprick was significantly positively correlated with the resting-state network amplitudes of both the SMN and VNop, and with the white matter MK of the primary somatomotor projection tracts of the right centrum semiovale. These novel neural correlates were hypothesised to reflect individual variability in brain maturity. However, the hypothesis that infants with more mature brains exhibit larger cerebral haemodynamic responses to a heel lance was inferred and not directly tested.

In the near future, it should be possible to test the validity of both the pinprick stimulus fMRI experimental paradigm and the "neural maturity" hypothesis, as data collection required to perform these tests has begun being collected by another member of the Paediatric Neuroimaging Group. In this study, infants' responses to the pinprick are being recorded using fMRI. On a separate test occasion, resting-state data are being recorded using EEG followed by the infants' behavioural, physiological, and EEG responses to a clinically required heel lance. With this dataset, it will be

possible to validate the pinprick stimulus fMRI experimental paradigm by testing the relationship between the fMRI-recorded response to the pinprick with a multitude of responses to the heel lance. Additionally, established measures of resting-state EEG-derived markers of neural functional maturity, such as multiscale entropy (De Wel et al., 2017), can be used to directly test the "neural maturity" hypothesis.

As discussed in Section 4.4, the correlation between resting-state VNop activity and noxious-evoked activity was surprising, and we hypothesised that the association was not due to a direct causal link between the visual network and pinprick response network, but due to the deeper neural maturity cause. This distinction between correlation and causation is not limited to this particular observation, but to all of the identified functional and structural correlates of this infant MRI study. Given the haemodynamic nature of the fMRI signal, this distinction needs to be expanded further. While discussions up to this point primarily centred on brain activity (due to the reasonable assumption that the haemodynamic variability is relatively representative of the underlying neural activity variability), this may not be the case. Due to the haemodynamic nature of the fMRI signal, it is highly plausible that the cross-infant variability in noxious-response amplitudes is primarily a function of cerebral vasculature variability, as opposed to underlying neural activity. Using the neural maturity hypothesis, one would then claim that the observed variability in cerebral haemodynamic response to noxious input is due to individual variability in the maturity of the cerebral vascular system. This intriguing interpretation would require a more in-depth assessment of the newborn infant cerebral vasculature and perfusion. We are currently in the process of commencing a study to directly assess this perfusion interpretation using arterial spin labeling (ASL), an MRI modality that allows perfusion quantification. The results of this study should also shed greater insight into the nature of the observed negative and negligible BOLD responses observed in some newborn infants.

Finally, the generalisability of the maturity hypothesis to other infant populations needs further exploration. The work in this thesis focused on healthy newborn infants, with larger cerebral haemodynamic response amplitudes interpreted as

the more mature response. However, infants born prematurely and exposed to high levels of early-life pain and stress may have altered developmental trajectories that may limit the usefulness of a simplistic maturity hypothesis. There is a lack of clarity from research in this area due to the high correlation between prematurity, early-life pain burden, and other confounding comorbidities. There exists a single EEG study that demonstrated larger pain responses in ex-premature infants studied at term compared to healthy term infants (Slater et al., 2010c). However, this study was conducted with small sample sizes (n=8 term infants, and n=7 ex-premature infants), the ex-premature sample had major comorbid confounds (three infants with grade four intraventricular haemorrhage), and these EEG results cannot be simply extrapolated to the haemodynamic response. Unfortunately, there is currently no research comparing term and ex-premature infants' responses to pain using haemodynamic measures such as NIRS or fMRI. Comparing the haemodynamic response amplitudes of term infants to ex-premature infants is an area our group is beginning to look into directly, so hopefully an answer will be available in the near future.

Considering the negative impact that prematurity has on development, any group differences observed between these populations would have to be interpreted as a potentially negative or neutral difference caused by prematurity, where negative or neutral relates to functional outcomes in later life, such as cognitive capacity at school age. Hypothetically, if the ex-premature population had smaller pinprick response amplitudes, then this could be interpreted as prematurity potentially causing delayed cerebral vasculature maturation. Conversely, if the ex-premature population had larger pinprick response amplitudes, then this could be interpreted as prematurity potentially causing accelerated cerebral vasculature maturation. However, given that advanced maturity may have positive connotations, while prematurity is known to be detrimental, then the concept of maturation would have to increase in complexity to incorporate cases of dysmaturity for both delayed and accelerated maturity. Finally, even if a group difference is statistically significant, the effect size might not allow for meaningful separation been healthy term and

ex-premature infants to warrant labelling potential prematurity-induced delayed or accelerated maturity with the term "dysmature". While this likely scenario does not appear to undermine the maturity hypothesis (but simply necessitates an elaboration of the terminology), it would limit the usefulness of the noxious-evoked haemodynamic response amplitude as a useful marker of cerebral functional maturity. In addition to testing for statistically significant differences between healthy term and ex-premature infant populations, it would thus be valuable to assess the degree of separation between populations as a potential marker of prematurity-induced disruptions to cerebral maturation.

5.2.5 Clinical utility assumption

In Chapter 4, the infant resting-state brain activity was shown to be informative of the cerebral haemodynamic response amplitude to the pinprick stimulus. The motivation for testing the predictability of the neural response to nociceptive input from resting-state data was to establish a proof-of-concept of potential clinical utility. Specifically, an infant's unique brain activity can potentially inform the clinical decision making process regarding appropriate infant pain management strategies, thus helping to reduce the uncertainty in this challenging process. It is then of central importance to directly test this concept, a study I will be undertaking as part of my postdoctoral research position. The general approach will be to use EEG to study infant's resting-state brain activity and noxious-response brain activity with the same experimental pinprick stimulus used in this thesis. These recordings will be made on neurologically healthy infants who require inguinal hernia repair surgery. The resting-state and noxious-response EEG recordings will be made both pre- and post-surgically, with the post-surgical recordings occurring several hours after surgery to ensure full elimination of the caudal block and general anaesthetic. Monitoring of physiological variables and behavioural measures of post-surgical pain or discomfort will also take place.

This rich dataset will allow the testing of several exciting and important questions, of which three are listed here. First, the resting-state/noxious-response functional

coupling of the pre-surgical EEG recordings will be explored to test if this functional coupling observed in the fMRI data can be translated from haemodynamic signals to electrical signals. Second, changes in EEG features between the pre- and post-surgical periods will be assessed. Post-surgical recording of pain levels (assessed regularly using appropriate pain scales such as FLACC), analgesic requirement, and physiological instability will also be recorded, and their potential association with the affected EEG features will be quantified. These relationships are vitally important to determine, as an understanding of the neurophysiological basis for these post-surgical assessments is currently lacking. It would be highly valuable if existing post-surgical pain assessments could be grounded in neurophysiology, or alternatively, if novel neurophysiological measures of post-surgical outcomes could be identified. Third, the pre-surgical neurophysiological EEG recordings (both resting-state and noxious-response recordings) will be tested for their capacity to predict clinically valuable measures of post-surgical outcomes such as pain levels, analgesic requirement, and physiological instability.

5.3 Closing statement

Having summarised the core findings of this thesis and proposed that the inherent limitations can be a springboard for future work, it is clear that the results and concepts detailed here should continue to be explored and developed. It is very exciting and rewarding to see that researchers in the Paediatric Neuroimaging Group, including myself, see the potential in pursuing these findings, and are taking the opportunity to do so. While the challenges involved in studying infant pain and the limitations of the MRI modality have been discussed at length, the research described in this thesis more importantly highlights the abundant potential for further progress. At the time of writing, there are still fewer than five published fMRI studies with a focus on infant pain processing. The ability to directly tackle the challenges that have resulted in this dearth of research, and noting the abundance of exciting novel findings when one does so appropriately, should hopefully be a cause for optimism for those hoping to undertake MRI-based research into infant pain in the future.

Appendices

A

Chapter 3 Supplementary Figures

This appendix presents examples of spatial independent component analysis (sICA) components commonly seen in the neonatal noxious-response fMRI dataset. These components are independent signals in the fMRI data that originate from independent signal-sources, eight of which are shown below. For all figures, the spatial map of the component at several axial slices is presented on the left upper pane, and at a single sagittal and coronal slice on the right upper pane. The component's timeseries is presented in the left lower pane, and the power spectrum of the timeseries in the right lower pane. All components are displayed using FSL's FSLeyes. The nature and origin of these neonatal sICA components are no different from that of adults. The examples below are deliberately displayed and named to resemble the adult examples in Griffanti et al., 2017 to facilitate comparison.

It is important to note that at the spatial and temporal resolution of these data, it is not always possible to unequivocally identify the signal-source of a component. For this reason, the component in Figure A.4 is labelled 'Cerebrospinal fluid pulsation (and arteries?)'. Similarly, a summary head motion timeseries (red dotted timeseries) is provided in the left lower pane of Figure A.2 and A.7 to demonstrate that the subject's head motion is clearly the signal-source for the 'Motion artefact' and 'Multiband artefact' components.

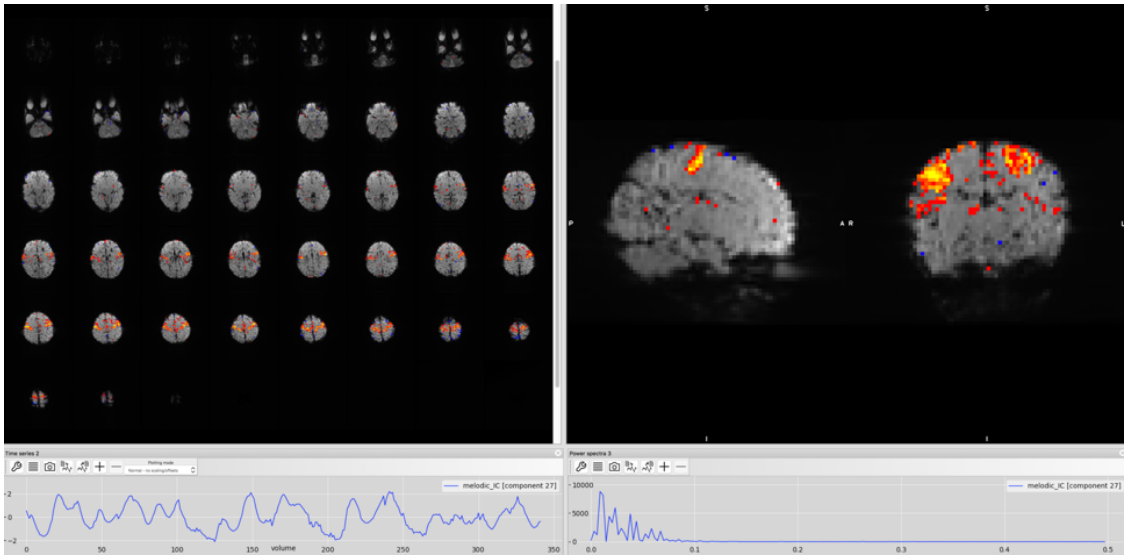

Figure A.1: Signal

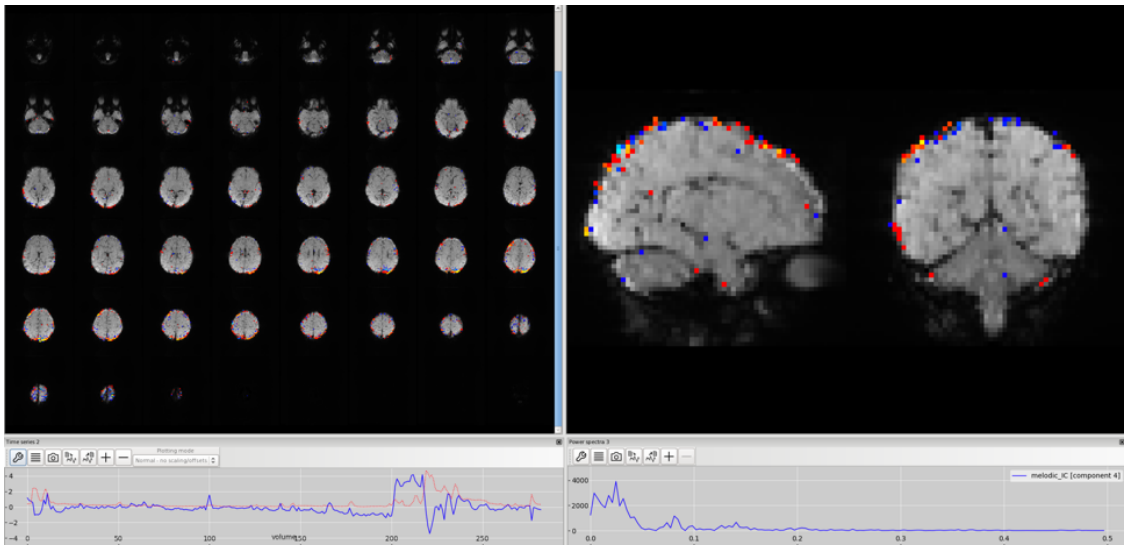

Figure A.2: Motion

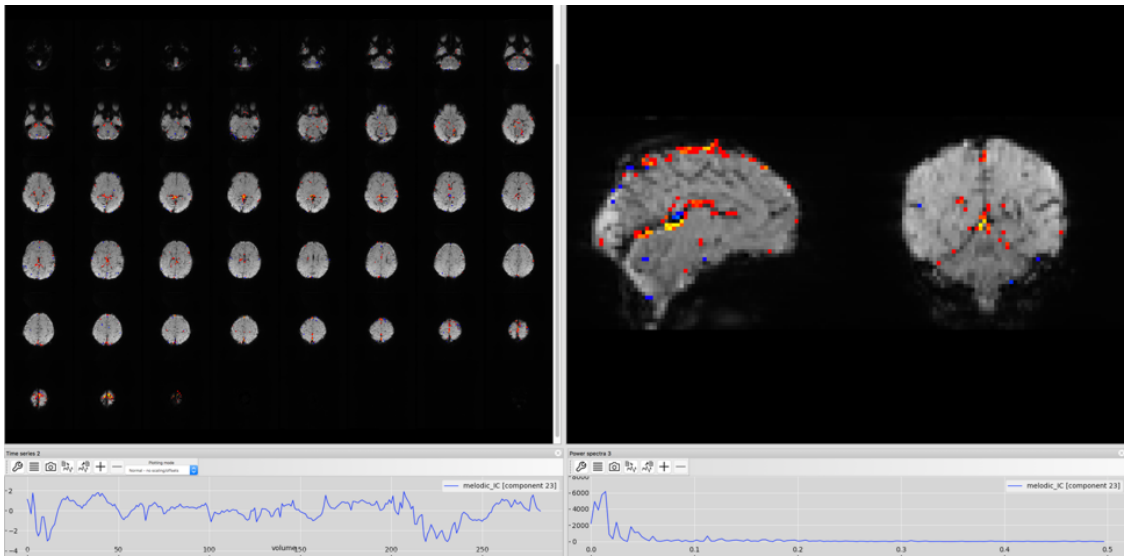

Figure A.3: Vein

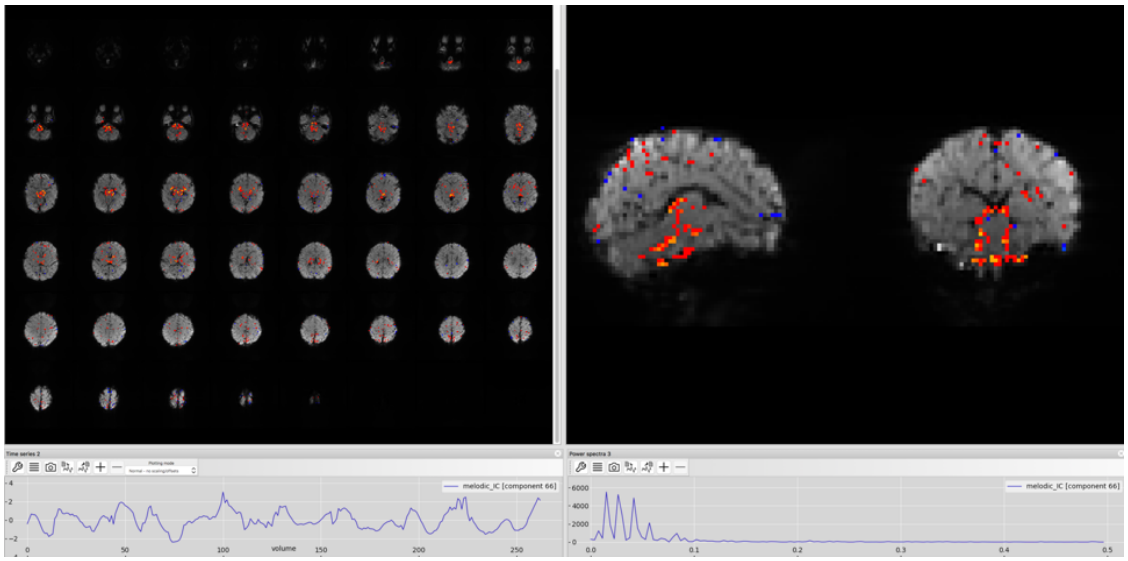

Figure A.4: Cerebrospinal fluid pulsation (and arteries?)

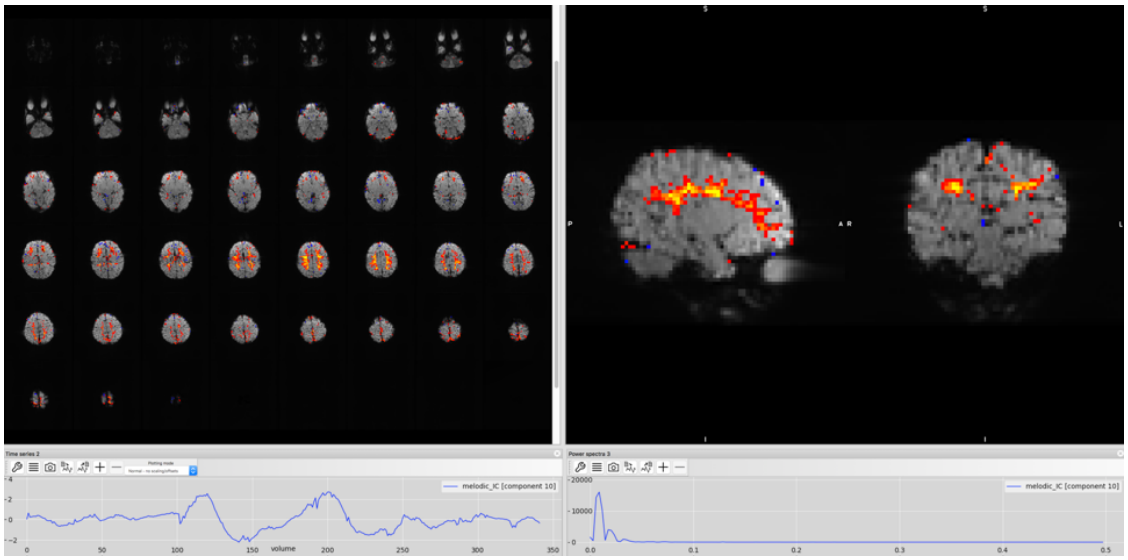

Figure A.5: Fluctuations in subependymal (and transmedullary) veins

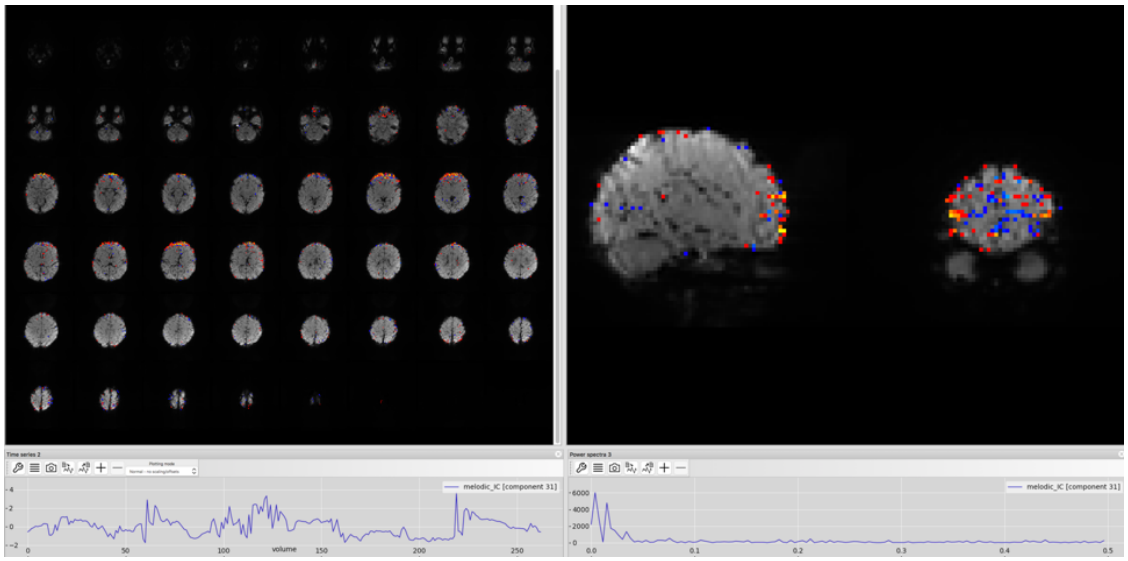

Figure A.6: Susceptibility artefacts

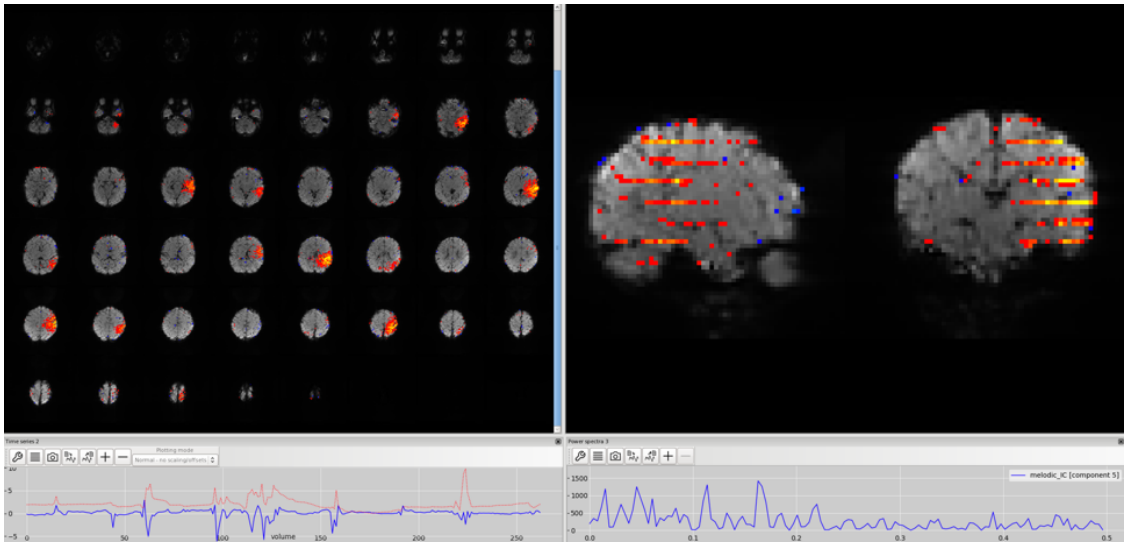

Figure A.7: Multiband artefact

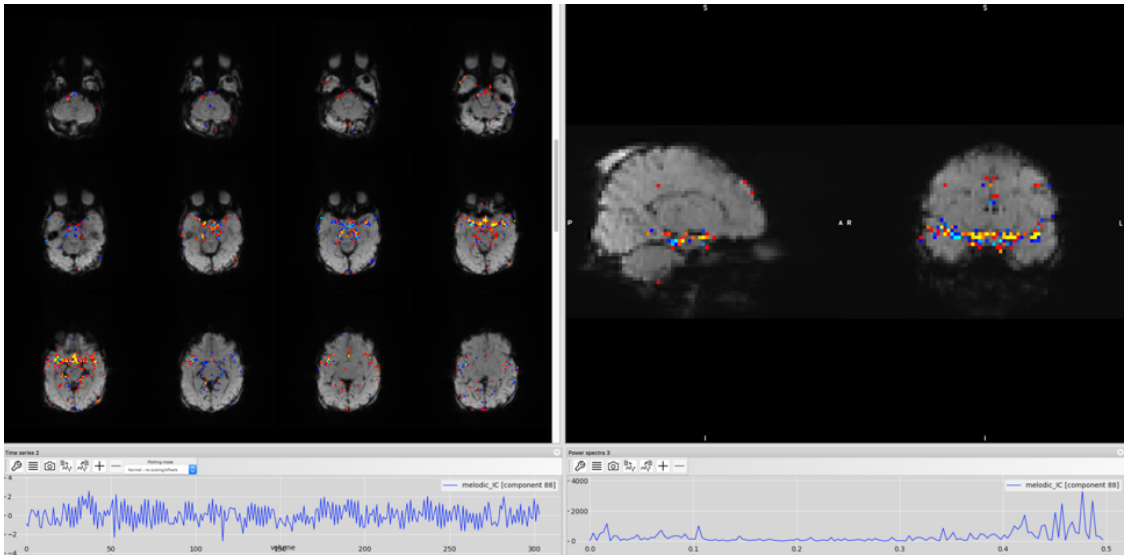

Figure A.8: MRI-related artefact

B

Chapter 4 Supplementary Figures

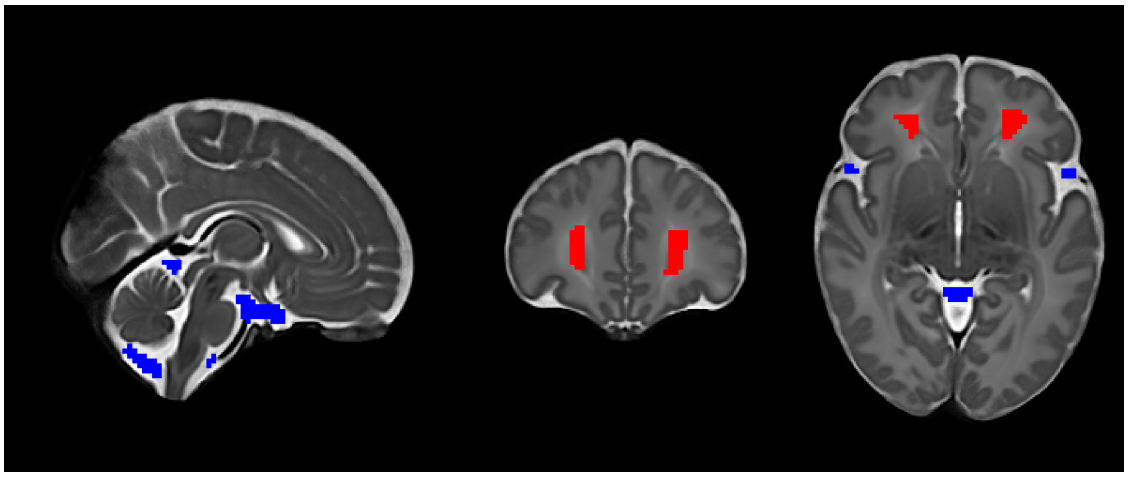

Figure B.1: Anatomically defined imaging confound regions-of-interest (ROIs). Both the cerebrospinal fluid (CSF) and white matter ROIs were defined using the standard infant template (Schuh et al., 2018). The CSF ROI (blue) is restricted predominantly to the fluid surrounding the brainstem and cerebellum and is a single region crossing the midline. Ventricular CSF was excluded from this ROI to minimize unwanted partial voluming effects due to the small size of infant ventricles. The white matter ROIs (red) are restricted to the largest white matter regions in the cerebrum. The left and right ROIs are tightly localized to the centre of white matter regions to minimize unwanted partial voluming effects with nearby grey matter tissue.

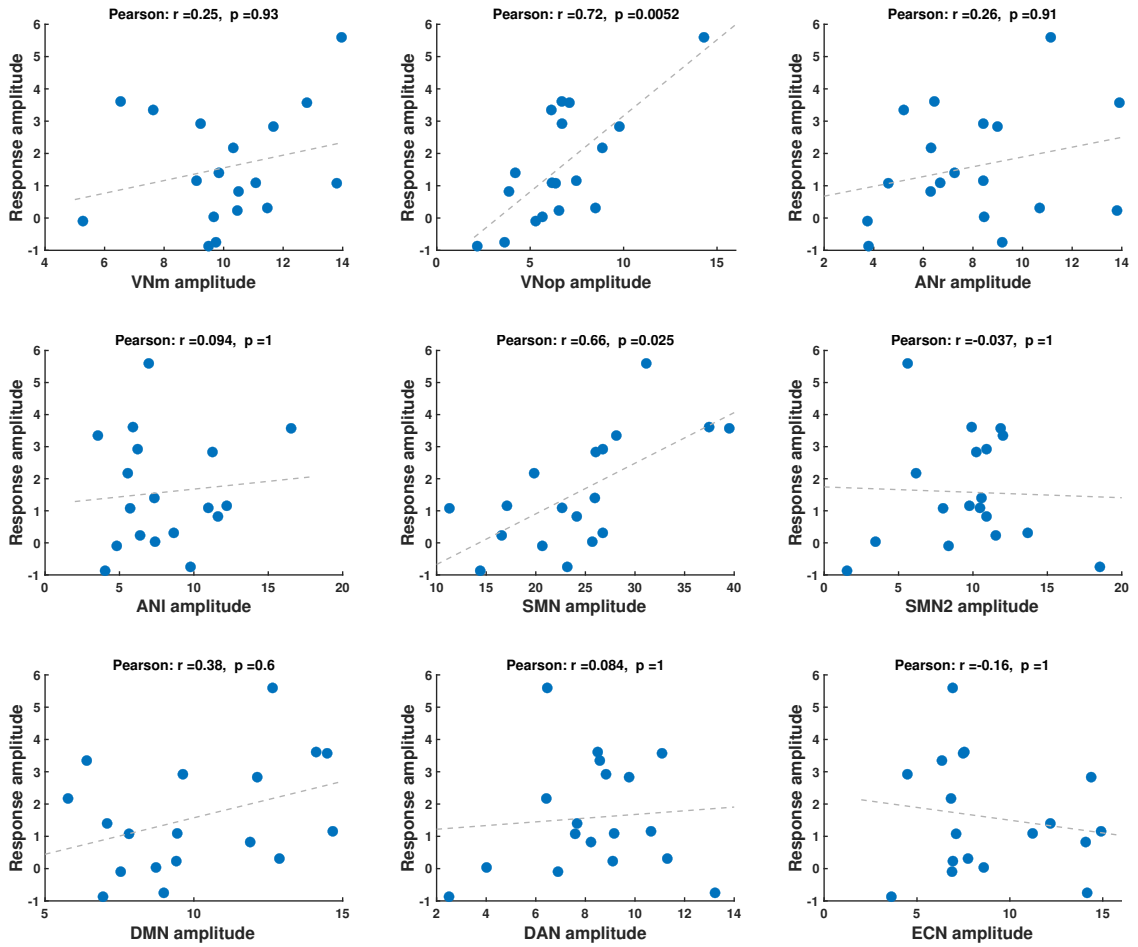

Figure B.2: Noxious-response amplitudes vs resting-state network amplitudes. In all plots, the x-axis is the resting-state network amplitude, the y-axis is the noxious-response amplitude, and the grey dashed line is the least-squares line. The networks are ordered from left-to-right then top-to-bottom in identical order to that presented from left-to-right in Figure 4.3A. Above each plot is presented the Pearson correlation coefficient (r) with its associated two-tailed p-value (p). All p-values are FWER-corrected for multiple testing. The SMN and VNop are the only two networks with statistically significant correlations (positive). VNm = medial visual network; VNop = occipital pole visual network; ANr = right auditory network; ANl = left auditory network; SMN = somoatomotor network; DMN = default mode network; DAN = dorsal attention network; ECN = executive control network.

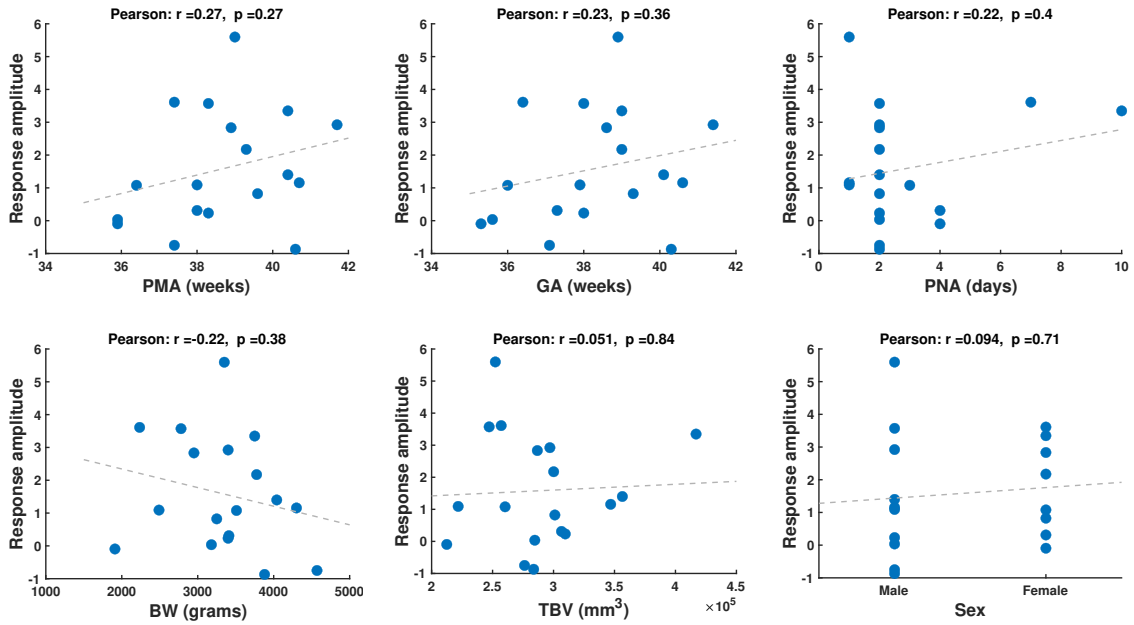

Figure B.3: Noxious-response amplitudes vs non-functional variables. In all plots, the x-axis is the non-functional variable, the y-axis is the noxious-response amplitude, and the grey dashed line is the least-squares line. Above each plot is presented the Pearson correlation coefficient (r) with its associated two-tailed p-value (p). Due to the lack of any statistically significant correlation, p-values are presented uncorrected for multiple testing. PMA = postmenstrual age; GA = gestational age; PNA = postnatal age; BW = birth weight; TBV = total brain volume.

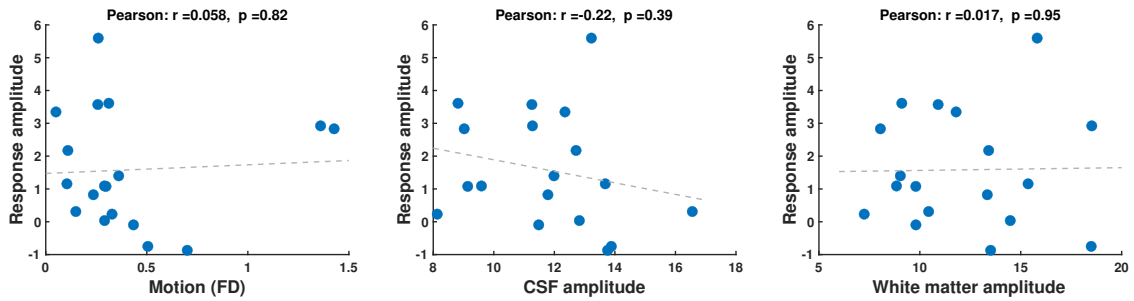

Figure B.4: Noxious-response amplitudes vs resting-state imaging confounds. In all plots, the x-axis is the resting-state imaging confound, the y-axis is the noxious-response amplitude, and the grey dashed line is the least-squares line. Above each plot is presented the Pearson correlation coefficient (r) with its associated two-tailed p-value (p). Due to the lack of any statistically significant correlation, p-values are presented uncorrected for multiple testing. FD = framewise displacement in units of millimetres; CSF = cerebrospinal fluid.

List of References

Alexander, A. L. et al. (May 2001). "Analysis of Partial Volume Effects in Diffusion-Tensor MRI". In: *Magnetic Resonance in Medicine* 45.5, pp. 770–780.

Allievi, A. G. et al. (Jan. 2016). "Maturation of Sensori-Motor Functional Responses in the Preterm Brain". In: *Cerebral Cortex (New York, N.Y.: 1991)* 26.1, pp. 402–413.

Anand, K. J. and K. D. Craig (Sept. 1996). "New Perspectives on the Definition of Pain". In: *Pain* 67.1, 3–6; discussion 209–211.

Anchisi, D. and M. Zanon (2015). "A Bayesian Perspective on Sensory and Cognitive Integration in Pain Perception and Placebo Analgesia". In: *PloS One* 10.2, e0117270.

Anderson, A. W. et al. (Jan. 2001). "Neonatal Auditory Activation Detected by Functional Magnetic Resonance Imaging". In: *Magnetic Resonance Imaging* 19.1, pp. 1–5.

Andersson, J, M. Jenkinson, and S. M. Smith (June 28, 2007). "Non-Linear Registration, Aka Spatial Normalisation". In: *FMRIB technical report TR07JA2*.

Andersson, J. L. R., S. Skare, and J. Ashburner (Oct. 2003). "How to Correct Susceptibility Distortions in Spin-Echo Echo-Planar Images: Application to Diffusion Tensor Imaging". In: *NeuroImage* 20.2, pp. 870–888.

Andersson, J. L. R. and S. N. Sotiropoulos (Nov. 15, 2015). "Non-Parametric Representation and Prediction of Single- and Multi-Shell Diffusion-Weighted MRI Data Using Gaussian Processes". In: *NeuroImage* 122, pp. 166–176.

— (Jan. 15, 2016). "An Integrated Approach to Correction for Off-Resonance Effects and Subject Movement in Diffusion MR Imaging". In: *NeuroImage* 125, pp. 1063–1078.

Andersson, J. L. R. et al. (Nov. 1, 2016). "Incorporating Outlier Detection and Replacement into a Non-Parametric Framework for Movement and Distortion Correction of Diffusion MR Images". In: *NeuroImage* 141, pp. 556–572.

Andersson, J. L. R. et al. (May 15, 2017). "Towards a Comprehensive Framework for Movement and Distortion Correction of Diffusion MR Images: Within Volume Movement". In: *NeuroImage* 152, pp. 450–466.

Andersson, J. L. R. et al. (Jan. 5, 2018). "Susceptibility-Induced Distortion That Varies Due to Motion: Correction in Diffusion MR without Acquiring Additional Data". In: *NeuroImage* 171, pp. 277–295.

André, M. et al. (May 1, 2010). "Electroencephalography in Premature and Full-Term Infants. Developmental Features and Glossary". In: *Neurophysiologie Clinique/Clinical Neurophysiology* 40.2, pp. 59–124.

Apkarian, A. V. and C. J. Hodge (Oct. 15, 1989). "Primate Spinothalamic Pathways: I. A Quantitative Study of the Cells of Origin of the Spinothalamic Pathway". In: *The Journal of Comparative Neurology* 288.3, pp. 447–473.

Apkarian, A. V. et al. (Aug. 2005). "Human Brain Mechanisms of Pain Perception and Regulation in Health and Disease". In: *European Journal of Pain (London, England)* 9.4, pp. 463–484.

Arichi, T. et al. (Feb. 1, 2010). "Somatosensory Cortical Activation Identified by Functional MRI in Preterm and Term Infants". In: *NeuroImage* 49.3, pp. 2063–2071.

Arichi, T. et al. (Nov. 1, 2012). "Development of BOLD Signal Hemodynamic Responses in the Human Brain". In: *NeuroImage* 63.2, pp. 663–673.

Arichi, T. et al. (Dec. 9, 2017). "Localization of Spontaneous Bursting Neuronal Activity in the Preterm Human Brain with Simultaneous EEG-fMRI". In: *eLife* 6.

Avants, B. B. et al. (Feb. 2008). "Symmetric Diffeomorphic Image Registration with Cross-Correlation: Evaluating Automated Labeling of Elderly and Neurodegenerative Brain". In: *Medical Image Analysis* 12.1, pp. 26–41.

Ball, G. et al. (Jan. 1, 2016). "Machine-Learning to Characterise Neonatal Functional Connectivity in the Preterm Brain". In: *NeuroImage* 124 (Pt A), pp. 267–275.

Ball, G. et al. (Oct. 15, 2010). "An Optimised Tract-Based Spatial Statistics Protocol for Neonates: Applications to Prematurity and Chronic Lung Disease". In: *NeuroImage* 53.1, pp. 94–102.

Bartocci, M. et al. (May 2006). "Pain Activates Cortical Areas in the Preterm Newborn Brain". In: *Pain* 122.1-2, pp. 109–117.

Basbaum, A. (2011). *Specificity Versus Patterning Theory: Continuing the Debate*. URL: `https://www.painresearchforum.org/forums/discussion/7347-specificity-versus-patterning-theory-continuing-debate`.

Bassi, L. et al. (Feb. 2008). "Probabilistic Diffusion Tractography of the Optic Radiations and Visual Function in Preterm Infants at Term Equivalent Age". In: *Brain: A Journal of Neurology* 131 (Pt 2), pp. 573–582.

Bastiani, M. et al. (Jan. 15, 2019). "Automated Processing Pipeline for Neonatal Diffusion MRI in the Developing Human Connectome Project". In: *NeuroImage* 185, pp. 750–763.

Baxter, L. et al. (Jan. 2, 2019). "Optimising Neonatal fMRI Data Analysis: Design and Validation of an Extended dHCP Preprocessing Pipeline to Characterise Noxious-Evoked Brain Activity in Infants". In: *NeuroImage* 186, pp. 286–300.

Beckmann, C. F. and S. M. Smith (Feb. 2004). "Probabilistic Independent Component Analysis for Functional Magnetic Resonance Imaging". In: *IEEE transactions on medical imaging* 23.2, pp. 137–152.

Beckmann, C. F. et al. (2009). "Group Comparison of Resting-State FMRI Data Using Multi-Subject ICA and Dual Regression". In: *OHBM*.

Behrens, T. E. J. et al. (Nov. 2003). "Characterization and Propagation of Uncertainty in Diffusion-Weighted MR Imaging". In: *Magnetic Resonance in Medicine* 50.5, pp. 1077–1088.

Behrens, T. E. J. et al. (Jan. 1, 2007). "Probabilistic Diffusion Tractography with Multiple Fibre Orientations: What Can We Gain?" In: *NeuroImage* 34.1, pp. 144–155.

Behrens, T. E. J., S. N. Sotiropoulos, and S. Jbabdi (Jan. 1, 2014). "Chapter 19 - MR Diffusion Tractography". In: *Diffusion MRI (Second Edition)*. Ed. by H. Johansen-Berg and T. E. J. Behrens. San Diego: Academic Press, pp. 429–451.

Bembich, S. et al. (Nov. 2015). "Pain Activates a Defined Area of the Somatosensory and Motor Cortex in Newborn Infants". In: *Acta Paediatrica (Oslo, Norway: 1992)* 104.11, e530–533.

Ben-Ari, Y. (Oct. 24, 2014). "The GABA Excitatory/Inhibitory Developmental Sequence: A Personal Journey". In: *Neuroscience* 279, pp. 187–219.

Berman, J. I. et al. (Jan. 2009). "Quantitative Fiber Tracking Analysis of the Optic Radiation Correlated with Visual Performance in Premature Newborns". In: *AJNR. American journal of neuroradiology* 30.1, pp. 120–124.

Bernard, J. F. and J. M. Besson (Mar. 1990). "The Spino(Trigemino)Pontoamygdaloid Pathway: Electrophysiological Evidence for an Involvement in Pain Processes". In: *Journal of Neurophysiology* 63.3, pp. 473–490.

Bester, H. et al. (Feb. 1995). "Spino (Trigemino) Parabrachiohypothalamic Pathway: Electrophysiological Evidence for an Involvement in Pain Processes". In: *Journal of Neurophysiology* 73.2, pp. 568–585.

Bijsterbosch, J. et al. (Jan. 10, 2017). "Investigations into Within- and between-Subject Resting-State Amplitude Variations". In: *NeuroImage* 159, pp. 57–69.

Bijsterbosch, J. D. et al. (Feb. 16, 2018). "The Relationship between Spatial Configuration and Functional Connectivity of Brain Regions". In: *eLife* 7.

Bijsterbosch, J. D. et al. (Aug. 5, 2019). "The Relationship between Spatial Configuration and Functional Connectivity of Brain Regions Revisited". In: *eLife* 8.

Biswal, B. et al. (Oct. 1995). "Functional Connectivity in the Motor Cortex of Resting Human Brain Using Echo-Planar MRI". In: *Magnetic Resonance in Medicine* 34.4, pp. 537–541.

Boly, M. et al. (July 17, 2007). "Baseline Brain Activity Fluctuations Predict Somatosensory Perception in Humans". In: *Proceedings of the National Academy of Sciences* 104.29, pp. 12187–12192.

Borsboom, D., G. J. Mellenbergh, and J. vanHeerden (Oct. 2004). "The Concept of Validity". In: *Psychological Review* 111.4, pp. 1061–1071.

Bremner, L., M. Fitzgerald, and M. Baccei (June 2006). "Functional GABA(A)-Receptor-Mediated Inhibition in the Neonatal Dorsal Horn". In: *Journal of Neurophysiology* 95.6, pp. 3893–3897.

Bullmore, E. T. et al. (Jan. 1999). "Global, Voxel, and Cluster Tests, by Theory and Permutation, for a Difference between Two Groups of Structural MR Images of the Brain". In: *IEEE transactions on medical imaging* 18.1, pp. 32–42.

Burgess, P. R. and E. R. Perl (June 1967). "Myelinated Afferent Fibres Responding Specifically to Noxious Stimulation of the Skin". In: *The Journal of Physiology* 190.3, pp. 541–562.

Burle, B. et al. (Sept. 1, 2015). "Spatial and Temporal Resolutions of EEG: Is It Really Black and White? A Scalp Current Density View". In: *International Journal of Psychophysiology*. On the Benefits of Using Surface Laplacian (Current Source Density) Methodology in Electrophysiology 97.3, pp. 210–220.

Burton, H. and E. G. Jones (July 15, 1976). "The Posterior Thalamic Region and Its Cortical Projection in New World and Old World Monkeys". In: *The Journal of Comparative Neurology* 168.2, pp. 249–301.

Bzdok, D. et al. (June 2016). "Formal Models of the Network Co-Occurrence Underlying Mental Operations". In: *PLoS computational biology* 12.6, e1004994.

Büchel, C. et al. (Mar. 19, 2014). "Placebo Analgesia: A Predictive Coding Perspective". In: *Neuron* 81.6, pp. 1223–1239.

Carbajal, R. et al. (July 2, 2008). "Epidemiology and Treatment of Painful Procedures in Neonates in Intensive Care Units". In: *JAMA* 300.1, pp. 60–70.

Carlton, S. M. et al. (Aug. 22, 1992). "GABA-Immunoreactive Terminals Synapse on Primate Spinothalamic Tract Cells". In: *The Journal of Comparative Neurology* 322.4, pp. 528–537.

Carp, J. (Oct. 15, 2012). "The Secret Lives of Experiments: Methods Reporting in the fMRI Literature". In: *NeuroImage* 63.1, pp. 289–300.

Cavanaugh, D. J. et al. (June 2, 2009). "Distinct Subsets of Unmyelinated Primary Sensory Fibers Mediate Behavioral Responses to Noxious Thermal and Mechanical Stimuli". In: *Proceedings of the National Academy of Sciences of the United States of America* 106.22, pp. 9075–9080.

Chang, L. J. et al. (June 2015). "A Sensitive and Specific Neural Signature for Picture-Induced Negative Affect". In: *PLoS biology* 13.6, e1002180.

Chappell, M., T. Okell, and M. Jenkinson (2020). *Short Introduction to MRI Physics for Neuroimaging*. Oxford Neuroimaging Primer Appendices. Independently published.

Cheung, M. M. et al. (Apr. 1, 2009). "Does Diffusion Kurtosis Imaging Lead to Better Neural Tissue Characterization? A Rodent Brain Maturation Study". In: *NeuroImage* 45.2, pp. 386–392.

Christensen, B. N. and E. R. Perl (Mar. 1970). "Spinal Neurons Specifically Excited by Noxious or Thermal Stimuli: Marginal Zone of the Dorsal Horn". In: *Journal of Neurophysiology* 33.2, pp. 293–307.

Clark, A. (June 2013). "Whatever next? Predictive Brains, Situated Agents, and the Future of Cognitive Science". In: *The Behavioral and Brain Sciences* 36.3, pp. 181–204.

Coalson, T. S., D. C. Van Essen, and M. F. Glasser (Mar. 7, 2018). "The Impact of Traditional Neuroimaging Methods on the Spatial Localization of Cortical Areas". In: *Proceedings of the National Academy of Sciences of the United States of America* 115.27, E6356–E6365.

Cole, M. W. et al. (July 2, 2014). "Intrinsic and Task-Evoked Network Architectures of the Human Brain". In: *Neuron* 83.1, pp. 238–251.

Cole, M. W. et al. (Feb. 1, 2016). "Functional Connectivity Change as Shared Signal Dynamics". In: *Journal of Neuroscience Methods* 259, pp. 22–39.

Colonnese, M. and R. Khazipov (Oct. 1, 2012). "Spontaneous Activity in Developing Sensory Circuits: Implications for Resting State fMRI". In: *NeuroImage*. Connectivity 62.4, pp. 2212–2221.

COSMIN (2019). *COnsensus-Based Standards for the Selection of Health Status Measurement INstruments*. URL: `https://www.cosmin.nl/`.

Courtois, E. et al. (July 1, 2016). "Epidemiology and Neonatal Pain Management of Heelsticks in Intensive Care Units: EPIPPAIN 2, a Prospective Observational Study". In: *International Journal of Nursing Studies* 59 (Supplement C), pp. 79–88.

Cowen, R. et al. (July 2015). "Assessing Pain Objectively: The Use of Physiological Markers". In: *Anaesthesia* 70.7, pp. 828–847.

Craig, K. D. (July 2015). "Social Communication Model of Pain". In: *Pain* 156.7, pp. 1198–1199.

Davidson, S., H. Truong, and G. J. Giesler (2010). "Quantitative Analysis of Spinothalamic Tract Neurons in Adult and Developing Mouse". In: *The Journal of Comparative Neurology* 518.16, pp. 3193–3204.

De Wel, O. et al. (Oct. 2017). "Complexity Analysis of Neonatal EEG Using Multiscale Entropy: Applications in Brain Maturation and Sleep Stage Classification". In: *Entropy* 19.10, p. 516.

Deen, B. et al. (Oct. 1, 2017). "Organization of High-Level Visual Cortex in Human Infants". In: *Nature Communications* 8, p. 13995.

deVet, H. C. W. et al. (2011). "Validity". In: *Measurement in Medicine: A Practical Guide*. Cambridge University Press, pp. 150–201.

Djouhri, L. (Feb. 2016). "Adelta-Fiber Low Threshold Mechanoreceptors Innervating Mammalian Hairy Skin: A Review of Their Receptive, Electrophysiological and Cytochemical Properties in Relation to Adelta-Fiber High Threshold Mechanoreceptors". In: *Neuroscience and Biobehavioral Reviews* 61, pp. 225–238.

Doria, V. et al. (Nov. 16, 2010). "Emergence of Resting State Networks in the Preterm Human Brain". In: *Proceedings of the National Academy of Sciences of the United States of America* 107.46, pp. 20015–20020.

Dubois, J. and G. Dehaene-Lambertz (2015). "Fetal and Postnatal Development of the Cortex: MRI and Genetics". In: *Brain Mapping: An Encyclopedic Reference*. Vol. 2. 3 vols. Elsevier.

Dubois, J. et al. (Sept. 12, 2014). "The Early Development of Brain White Matter: A Review of Imaging Studies in Fetuses, Newborns and Infants". In: *Neuroscience* 276, pp. 48–71.

Duff, E. P. et al. (July 2008). "The Power of Spectral Density Analysis for Mapping Endogenous BOLD Signal Fluctuations". In: *Human Brain Mapping* 29.7, pp. 778–790.

Duff, E. P. et al. (June 2018). "Disambiguating Brain Functional Connectivity". In: *NeuroImage* 173, pp. 540–550.

Dum, R. P., D. J. Levinthal, and P. L. Strick (Nov. 11, 2009). "The Spinothalamic System Targets Motor and Sensory Areas in the Cerebral Cortex of Monkeys". In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 29.45, pp. 14223–14235.

Eippert, F. et al. (Aug. 27, 2009). "Activation of the Opioidergic Descending Pain Control System Underlies Placebo Analgesia". In: *Neuron* 63.4, pp. 533–543.

Eklund, A., H. Knutsson, and T. E. Nichols (May 2019). "Cluster Failure Revisited: Impact of First Level Design and Physiological Noise on Cluster False Positive Rates". In: *Human Brain Mapping* 40.7, pp. 2017–2032.

Engle, W. A. and American Academy of Pediatrics Committee on Fetus and Newborn (Nov. 2004). "Age Terminology during the Perinatal Period". In: *Pediatrics* 114.5, pp. 1362–1364.

Fabrizi, L. et al. (Sept. 27, 2011). "A Shift in Sensory Processing That Enables the Developing Human Brain to Discriminate Touch from Pain". In: *Current biology: CB* 21.18, pp. 1552–1558.

Fazeli, S. and C. Büchel (July 18, 2018). "Pain-Related Expectation and Prediction Error Signals in the Anterior Insula Are Not Related to Aversiveness". In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 38.29, pp. 6461–6474.

FDA-NIH Biomarker Working Group (2016). *BEST (Biomarkers, EndpointS, and Other Tools) Resource*. Silver Spring (MD): Food and Drug Administration (US).

Finn, E. S. et al. (Nov. 2015). "Functional Connectome Fingerprinting: Identifying Individuals Using Patterns of Brain Connectivity". In: *Nature Neuroscience* 18.11, pp. 1664–1671.

Fitzgerald, M. and E. Jennings (July 6, 1999). "The Postnatal Development of Spinal Sensory Processing". In: *Proceedings of the National Academy of Sciences of the United States of America* 96.14, pp. 7719–7722.

Fitzgerald, M. and J. Swett (Dec. 30, 1983). "The Termination Pattern of Sciatic Nerve Afferents in the Substantia Gelatinosa of Neonatal Rats". In: *Neuroscience Letters* 43.2-3, pp. 149–154.

Fitzgerald, M. (Dec. 2015). "What Do We Really Know about Newborn Infant Pain?" In: *Experimental Physiology* 100.12, pp. 1451–1457.

Fitzgibbon, S et al. (2018). "The Developing Human Connectome Project Automated Functional Processing Framework for Neonates". In: *OHBM*.

Fitzgibbon, S. P. et al. (Sept. 12, 2019). "The Developing Human Connectome Project (dHCP) Automated Resting-State Functional Processing Framework for Newborn Infants". In: *bioRxiv*, p. 766030.

Fransson, P. et al. (Sept. 25, 2007). "Resting-State Networks in the Infant Brain". In: *Proceedings of the National Academy of Sciences of the United States of America* 104.39, pp. 15531–15536.

Fransson, P. et al. (Sept. 2009). "Spontaneous Brain Activity in the Newborn Brain during Natural Sleep–an fMRI Study in Infants Born at Full Term". In: *Pediatric Research* 66.3, pp. 301–305.

Fransson, P. et al. (Jan. 1, 2011). "The Functional Architecture of the Infant Brain as Revealed by Resting-State fMRI". In: *Cerebral Cortex* 21.1, pp. 145–154.

François, A. et al. (Feb. 22, 2017). "A Brainstem-Spinal Cord Inhibitory Circuit for Mechanical Pain Modulation by GABA and Enkephalins". In: *Neuron* 93.4, 822–839.e6.

Friston, K. J. et al. (Oct. 1997). "Psychophysiological and Modulatory Interactions in Neuroimaging". In: *NeuroImage* 6.3, pp. 218–229.

Friston, K. J. (2011). "Functional and Effective Connectivity: A Review". In: *Brain Connectivity* 1.1, pp. 13–36.

Geuter, S. et al. (May 19, 2017). "Functional Dissociation of Stimulus Intensity Encoding and Predictive Coding of Pain in the Insula". In: *eLife* 6.

Glasser, M. F. et al. (Jan. 11, 2018). "Using Temporal ICA to Selectively Remove Global Noise While Preserving Global Signal in Functional MRI Data". In: *NeuroImage* 181, pp. 692–717.

Glover, G. H. (Apr. 1999). "Deconvolution of Impulse Response in Event-Related BOLD fMRI". In: *NeuroImage* 9.4, pp. 416–429.

Goksan, S. et al. (Apr. 21, 2015). "fMRI Reveals Neural Activity Overlap between Adult and Infant Pain". In: *eLife* 4.

Goksan, S. et al. (Aug. 2017). "Optimal Echo Time for Functional MRI of the Infant Brain Identified in Response to Noxious Stimulation". In: *Magnetic Resonance in Medicine* 78.2, pp. 625–631.

Goksan, S. et al. (Nov. 9, 2018). "The Influence of the Descending Pain Modulatory System on Infant Pain-Related Brain Activity". In: *eLife* 7.

Gorgolewski, K. J. et al. (June 21, 2016). "The Brain Imaging Data Structure, a Format for Organizing and Describing Outputs of Neuroimaging Experiments". In: *Scientific Data* 3, p. 160044.

Gracely, R. (2013). "Studies of Pain in Human Subjects". In: *Wall and Melzack's Textbook of Pain*. sixth. Saunders, pp. 283–300.

Greisen, G. (Dec. 2009). "To Autoregulate or Not to Autoregulate–That Is No Longer the Question". In: *Seminars in Pediatric Neurology* 16.4, pp. 207–215.

Greve, D. N. and B. Fischl (Oct. 15, 2009). "Accurate and Robust Brain Image Alignment Using Boundary-Based Registration". In: *NeuroImage* 48.1, pp. 63–72.

Griffanti, L. et al. (July 15, 2014). "ICA-Based Artefact Removal and Accelerated fMRI Acquisition for Improved Resting State Network Imaging". In: *NeuroImage* 95, pp. 232–247.

Griffanti, L. et al. (Jan. 7, 2017). "Hand Classification of fMRI ICA Noise Components". In: *NeuroImage* 154, pp. 188–205.

Groppo, M. et al. (July 2014). "Development of the Optic Radiations and Visual Function after Premature Birth". In: *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior* 56, pp. 30–37.

Haber, L. H., B. D. Moore, and W. D. Willis (May 1, 1982). "Electrophysiological Response Properties of Spinoreticular Neurons in the Monkey". In: *The Journal of Comparative Neurology* 207.1, pp. 75–84.

Harris, J. J., C. Reynell, and D. Attwell (July 2011). "The Physiology of Developmental Changes in BOLD Functional Imaging Signals". In: *Developmental Cognitive Neuroscience* 1.3, pp. 199–216.

Harrison, S. J. et al. (Apr. 1, 2015). "Large-Scale Probabilistic Functional Modes from Resting State fMRI". In: *NeuroImage* 109, pp. 217–231.

Harrison, S. J. et al. (Feb. 8, 2019). "Modelling Subject Variability in the Spatial and Temporal Characteristics of Functional Modes". In: *bioRxiv*, p. 544817.

Hartley, C. et al. (July 31, 2015). "The Relationship between Nociceptive Brain Activity, Spinal Reflex Withdrawal and Behaviour in Newborn Infants". In: *Scientific Reports* 5, p. 12519.

Hartley, C. et al. (Aug. 8, 2016). "Changing Balance of Spinal Cord Excitability and Nociceptive Brain Activity in Early Human Development". In: *Current biology: CB* 26.15, pp. 1998–2002.

Hartley, C. et al. (Mar. 5, 2017). "Nociceptive Brain Activity as a Measure of Analgesic Efficacy in Infants". In: *Science Translational Medicine* 9.388.

Hathway, G. J. et al. (June 15, 2009). "The Changing Balance of Brainstem-Spinal Cord Modulation of Pain Processing over the First Weeks of Rat Postnatal Life". In: *The Journal of Physiology* 587 (Pt 12), pp. 2927–2935.

Hathway, G. et al. (Jan. 2006). "A Postnatal Switch in GABAergic Control of Spinal Cutaneous Reflexes". In: *The European Journal of Neuroscience* 23.1, pp. 112–118.

Heinricher, M. M. et al. (Apr. 2009). "Descending Control of Nociception: Specificity, Recruitment and Plasticity". In: *Brain Research Reviews* 60.1, pp. 214–225.

Hillman, E. M. C. (2014). "Out for Blood". In: *Scientific American Mind* 25.4, pp. 58–65.

Holmes, A. P. et al. (Jan. 1996). "Nonparametric Analysis of Statistic Images from Functional Mapping Experiments". In: *Journal of Cerebral Blood Flow and Metabolism: Official Journal of the International Society of Cerebral Blood Flow and Metabolism* 16.1, pp. 7–22.

Hoshi, Y. (July 2003). "Functional Near-Infrared Optical Imaging: Utility and Limitations in Human Brain Mapping". In: *Psychophysiology* 40.4, pp. 511–520.

Hosobuchi, Y., J. E. Adams, and R. Linchitz (July 8, 1977). "Pain Relief by Electrical Stimulation of the Central Gray Matter in Humans and Its Reversal by Naloxone". In: *Science (New York, N.Y.)* 197.4299, pp. 183–186.

Huettel, S. A., A. W. Song, and G. McCarthy (Aug. 18, 2014). *Functional Magnetic Resonance Imaging*. Third Edition. Oxford, New York: Oxford University Press. 573 pp.

Hughes, E. J. et al. (Aug. 2017). "A Dedicated Neonatal Brain Imaging System". In: *Magnetic Resonance in Medicine* 78.2, pp. 794–804.

Hylden, J. L. et al. (Apr. 9, 1986). "Serotonin Innervation of Physiologically Identified Lamina I Projection Neurons". In: *Brain Research* 370.2, pp. 401–404.

IASP (2019). *International Association for the Study of Pain*. URL: `https://www.iasp-pain.org`.

Jenkinson, M. and S. Smith (June 2001). "A Global Optimisation Method for Robust Affine Registration of Brain Images". In: *Medical Image Analysis* 5.2, pp. 143–156.

Jenkinson, M. et al. (Oct. 2002). "Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images". In: *NeuroImage* 17.2, pp. 825–841.

Jenkinson, M. et al. (Aug. 15, 2012). "FSL". In: *NeuroImage*. 20 YEARS OF fMRI 62.2, pp. 782–790.

Jennings, E. and M. Fitzgerald (Dec. 1996). "C-Fos Can Be Induced in the Neonatal Rat Spinal Cord by Both Noxious and Innocuous Peripheral Stimulation". In: *Pain* 68.2-3, pp. 301–306.

Jensen, J. H. and J. A. Helpern (Aug. 2010). "MRI Quantification of Non-Gaussian Water Diffusion by Kurtosis Analysis". In: *NMR in biomedicine* 23.7, pp. 698–710.

Jensen, J. H. et al. (June 2005). "Diffusional Kurtosis Imaging: The Quantification of Non-Gaussian Water Diffusion by Means of Magnetic Resonance Imaging". In: *Magnetic Resonance in Medicine* 53.6, pp. 1432–1440.

Jones, A. K. et al. (Apr. 22, 1991). "Cortical and Subcortical Localization of Response to Pain in Man Using Positron Emission Tomography". In: *Proceedings. Biological Sciences* 244.1309, pp. 39–44.

Jones, E. (2007). "Intralaminar Nuclei". In: *The Thalamus*. Second.

Jones, L. et al. (2017). "Nociceptive Cortical Activity Is Dissociated from Nociceptive Behavior in Newborn Human Infants under Stress". In: *Current biology: CB* 27.24, 3846–3851.e3.

Kenshalo, D. R. and O. Isensee (Dec. 1983). "Responses of Primate SI Cortical Neurons to Noxious Stimuli". In: *Journal of Neurophysiology* 50.6, pp. 1479–1496.

Kenshalo, D. R. et al. (June 1980). "Responses of Neurons in Primate Ventral Posterior Lateral Nucleus to Noxious Stimuli". In: *Journal of Neurophysiology* 43.6, pp. 1594–1614.

Khazipov, R. et al. (Dec. 9, 2004). "Early Motor Activity Drives Spindle Bursts in the Developing Somatosensory Cortex". In: *Nature* 432.7018, pp. 758–761.

Konietzny, F. et al. (1981). "Sensory Experiences in Man Evoked by Intraneural Electrical Stimulation of Intact Cutaneous Afferent Fibers". In: *Experimental Brain Research* 42.2, pp. 219–222.

Kostović, I. and N. Jovanov-Milosević (Dec. 2006). "The Development of Cerebral Connections during the First 20-45 Weeks' Gestation". In: *Seminars in Fetal & Neonatal Medicine* 11.6, pp. 415–422.

Kozberg, M. and E. Hillman (2016). "Neurovascular Coupling and Energy Metabolism in the Developing Brain". In: *Progress in Brain Research* 225, pp. 213–242.

Kozberg, M. G. et al. (Mar. 12, 2013). "Resolving the Transition from Negative to Positive Blood Oxygen Level-Dependent Responses in the Developing Brain". In: *Proceedings of the National Academy of Sciences of the United States of America* 110.11, pp. 4380–4385.

Kragel, P. A. et al. (July 25, 2018). "Representation, Pattern Information, and Brain Signatures: From Neurons to Neuroimaging". In: *Neuron* 99.2, pp. 257–273.

Krahé, C. et al. (2013). "The Social Modulation of Pain: Others as Predictive Signals of Salience - a Systematic Review". In: *Frontiers in Human Neuroscience* 7, p. 386.

Krishnan, A. et al. (June 14, 2016). "Somatic and Vicarious Pain Are Represented by Dissociable Multivariate Brain Patterns". In: *eLife* 5.

Kvalseth, T. O. (1985). "Cautionary Note about R2". In: *The American Statistician* 39.4, pp. 279–285.

Kwiat, G. C. and A. I. Basbaum (1992). "The Origin of Brainstem Noradrenergic and Serotonergic Projections to the Spinal Cord Dorsal Horn in the Rat". In: *Somatosensory & Motor Research* 9.2, pp. 157–173.

Lee, G. and B. Stevens (2013). "Neonatal and Infant Pain Assessment". In: *Oxford Textbook of Paediatric Pain*. Oxford University Press, pp. 353–369.

Light, A. R. and E. R. Perl (July 15, 1979a). "Reexamination of the Dorsal Root Projection to the Spinal Dorsal Horn Including Observations on the Differential Termination of Coarse and Fine Fibers". In: *The Journal of Comparative Neurology* 186.2, pp. 117–131.

— (July 15, 1979b). "Spinal Termination of Functionally Identified Primary Afferent Neurons with Slowly Conducting Myelinated Fibers". In: *The Journal of Comparative Neurology* 186.2, pp. 133–150.

Logothetis, N. K. (May 15, 2003). "The Underpinnings of the BOLD Functional Magnetic Resonance Imaging Signal". In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 23.10, pp. 3963–3971.

Lowe, M. J. and J. A. Sorenson (May 1997). "Spatially Filtering Functional Magnetic Resonance Imaging Data". In: *Magnetic Resonance in Medicine* 37.5, pp. 723–729.

Lumpkin, E. A. and M. J. Caterina (Feb. 22, 2007). "Mechanisms of Sensory Transduction in the Skin". In: *Nature* 445.7130, pp. 858–865.

Makropoulos, A. et al. (Sept. 2014). "Automatic Whole Brain MRI Segmentation of the Developing Neonatal Brain". In: *IEEE transactions on medical imaging* 33.9, pp. 1818–1831.

Makropoulos, A. et al. (Jan. 15, 2016). "Regional Growth and Atlasing of the Developing Human Brain". In: *NeuroImage* 125, pp. 456–478.

Makropoulos, A. et al. (Jan. 31, 2018). "The Developing Human Connectome Project: A Minimal Processing Pipeline for Neonatal Cortical Surface Reconstruction". In: *NeuroImage* 173, pp. 88–112.

Mantyh, P. W. (Oct. 10, 1982). "The Ascending Input to the Midbrain Periaqueductal Gray of the Primate". In: *The Journal of Comparative Neurology* 211.1, pp. 50–64.

McGlone, F., J. Wessberg, and H. Olausson (May 21, 2014). "Discriminative and Affective Touch: Sensing and Feeling". In: *Neuron* 82.4, pp. 737–755.

Mehler, W. R., M. E. Feferman, and W. J. Nauta (Dec. 1960). "Ascending Axon Degeneration Following Anterolateral Cordotomy. An Experimental Study in the Monkey". In: *Brain: A Journal of Neurology* 83, pp. 718–750.

Melzack, R. and P. D. Wall (Nov. 19, 1965). "Pain Mechanisms: A New Theory". In: *Science (New York, N.Y.)* 150.3699, pp. 971–979.

Mendell, L. M. (Nov. 1966). "Physiological Properties of Unmyelinated Fiber Projection to the Spinal Cord". In: *Experimental Neurology* 16.3, pp. 316–332.

Mendell, L. M. and P. D. Wall (Apr. 3, 1965). "Responses of Single Dorsal Cord Cells to Peripheral Cutaneous Unmyelinated Fibres". In: *Nature* 206, pp. 97–99.

Mendell, L. M. (Feb. 2014). "Constructing and Deconstructing the Gate Theory of Pain". In: *Pain* 155.2, pp. 210–216.

Mennes, M. et al. (May 1, 2010). "Inter-Individual Differences in Resting-State Functional Connectivity Predict Task-Induced BOLD Activity". In: *NeuroImage* 50.4, pp. 1690–1701.

Mennes, M. et al. (Feb. 14, 2011). "Linking Inter-Individual Differences in Neural Activation and Behavior to Intrinsic Brain Dynamics". In: *NeuroImage* 54.4, pp. 2950–2959.

Mennes, M. et al. (Jan. 2013). "The Extrinsic and Intrinsic Functional Architectures of the Human Brain Are Not Equivalent". In: *Cerebral Cortex (New York, N.Y.: 1991)* 23.1, pp. 223–229.

Merskey, H et al. (June 1979). "Pain Terms: A List with Definitions and Notes on Usage. Recommended by the IASP Subcommittee on Taxonomy". In: *Pain* 6.3, p. 249.

Milh, M. et al. (July 2007). "Rapid Cortical Oscillations and Early Motor Activity in Premature Human Neonate". In: *Cerebral Cortex (New York, N.Y.: 1991)* 17.7, pp. 1582–1594.

Miller, K. E. and A. T. Salvatierra (Aug. 1998). "Apposition of Enkephalin- and Neurotensin-Immunoreactive Neurons by Serotonin-Immunoreactive Varicosities in the Rat Spinal Cord". In: *Neuroscience* 85.3, pp. 837–846.

Mitra, A. et al. (2017). "Resting-State fMRI in Sleeping Infants More Closely Resembles Adult Sleep than Adult Wakefulness". In: *PloS One* 12.11, e0188122.

Moayedi, M. and K. D. Davis (Jan. 2013). "Theories of Pain: From Specificity to Gate Control". In: *Journal of Neurophysiology* 109.1, pp. 5–12.

Moeller, S. et al. (May 2010). "Multiband Multislice GE-EPI at 7 Tesla, with 16-Fold Acceleration Using Partial Parallel Imaging with Application to High Spatial and Temporal Whole-Brain fMRI". In: *Magnetic Resonance in Medicine* 63.5, pp. 1144–1153.

Mokkink, L. B. et al. (July 2010). "The COSMIN Study Reached International Consensus on Taxonomy, Terminology, and Definitions of Measurement Properties for Health-Related Patient-Reported Outcomes". In: *Journal of Clinical Epidemiology* 63.7, pp. 737–745.

Mongerson, C. R. L. et al. (Aug. 14, 2017). "Resting-State Functional Connectivity in the Infant Brain: Methods, Pitfalls, and Potentiality". In: *Frontiers in Pediatrics* 5.

Monti, M. M. (2011). "Statistical Analysis of fMRI Time-Series: A Critical Review of the GLM Approach". In: *Frontiers in Human Neuroscience* 5, p. 28.

Niazy, R. K. et al. (2015). "Resting-State Networks". In: *fMRI: From Nuclear Spins to Brain Functions*. Ed. by K. Uludag, K. Ugurbil, and L. Berliner. Biological Magnetic Resonance. Boston, MA: Springer US, pp. 387–425.

Nickerson, L. D. et al. (2017). "Using Dual Regression to Investigate Network Shape and Amplitude in Functional Connectivity Analyses". In: *Frontiers in Neuroscience* 11, p. 115.

Ochoa, J. and E. Torebjörk (Aug. 1989). "Sensations Evoked by Intraneural Microstimulation of C Nociceptor Fibres in Human Skin Nerves". In: *The Journal of Physiology* 415, pp. 583–599.

Ongaro, G. and T. J. Kaptchuk (Jan. 2019). "Symptom Perception, Placebo Effects, and the Bayesian Brain". In: *Pain* 160.1, pp. 1–4.

Parker Jones, O. et al. (2017). "Resting Connectivity Predicts Task Activation in Pre-Surgical Populations". In: *NeuroImage. Clinical* 13, pp. 378–385.

Parkes, L. et al. (Jan. 5, 2018). "An Evaluation of the Efficacy, Reliability, and Sensitivity of Motion Correction Strategies for Resting-State Functional MRI". In: *NeuroImage* 171, pp. 415–436.

Paus, T. et al. (Feb. 2001). "Maturation of White Matter in the Human Brain: A Review of Magnetic Resonance Studies". In: *Brain Research Bulletin* 54.3, pp. 255–266.

Paydar, A. et al. (Apr. 2014). "Diffusional Kurtosis Imaging of the Developing Brain". In: *AJNR. American journal of neuroradiology* 35.4, pp. 808–814.

Pedregosa, F. et al. (2011). "Scikit-Learn: Machine Learning in Python". In: *Journal of Machine Learning Research* 12, pp. 2825–2830.

Perl, E. R. (Aug. 1968). "Myelinated Afferent Fibres Innervating the Primate Skin and Their Response to Noxious Stimuli". In: *The Journal of Physiology* 197.3, pp. 593–615.

— (1996). "Cutaneous Polymodal Receptors: Characteristics and Plasticity". In: *Progress in Brain Research* 113, pp. 21–37.

Pillai Riddell, R. et al. (2013). "Psychological Theories and Biopsychosocial Models in Paediatric Pain". In: *Oxford Textbook of Paediatric Pain*. Oxford University Press, pp. 85–94.

Ploner, M. et al. (Jan. 5, 2010). "Prestimulus Functional Connectivity Determines Pain Perception in Humans". In: *Proceedings of the National Academy of Sciences of the United States of America* 107.1, pp. 355–360.

Poldrack, R. A., G. Huckins, and G. Varoquaux (Nov. 27, 2019). "Establishment of Best Practices for Evidence for Prediction: A Review". In: *JAMA psychiatry*.

Power, J. D. et al. (Feb. 1, 2012). "Spurious but Systematic Correlations in Functional Connectivity MRI Networks Arise from Subject Motion". In: *NeuroImage* 59.3, pp. 2142–2154.

Power, J. D. et al. (Jan. 1, 2014). "Methods to Detect, Characterize, and Remove Motion Artifact in Resting State fMRI". In: *NeuroImage* 84, pp. 320–341.

Price, A. N. et al. (2015). "Accelerated Neonatal fMRI Using Multiband EPI". In: *ISMRM*.

Price, D. D., J. D. Greenspan, and R. Dubner (Dec. 2003). "Neurons Involved in the Exteroceptive Function of Pain". In: *Pain* 106.3, pp. 215–219.

Reuter, M. et al. (Feb. 15, 2015). "Head Motion during MRI Acquisition Reduces Gray Matter Volume and Thickness Estimates". In: *NeuroImage* 107, pp. 107–115.

Rivkin, M. J. et al. (June 2004). "Prolonged T*2 Values in Newborn versus Adult Brain: Implications for fMRI Studies of Newborns". In: *Magnetic Resonance in Medicine* 51.6, pp. 1287–1291.

Rolls, E. T. et al. (Mar. 2003). "Representations of Pleasant and Painful Touch in the Human Orbitofrontal and Cingulate Cortices". In: *Cerebral Cortex (New York, N.Y.: 1991)* 13.3, pp. 308–317.

Roofthooft, D. W. E. et al. (2014). "Eight Years Later, Are We Still Hurting Newborn Infants?" In: *Neonatology* 105.3, pp. 218–226.

Ruda, M. A., J. Coffield, and R. Dubner (Aug. 1984). "Demonstration of Postsynaptic Opioid Modulation of Thalamic Projection Neurons by the Combined Techniques of Retrograde Horseradish Peroxidase and Enkephalin Immunocytochemistry". In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 4.8, pp. 2117–2132.

Salimi-Khorshidi, G. et al. (Apr. 15, 2014). "Automatic Denoising of Functional MRI Data: Combining Independent Component Analysis and Hierarchical Fusion of Classifiers". In: *NeuroImage* 90, pp. 449–468.

Sandrini, G. et al. (Dec. 2005). "The Lower Limb Flexion Reflex in Humans". In: *Progress in Neurobiology* 77.6, pp. 353–395.

Saper, C. B. (Mar. 20, 1995). "The Spinoparabrachial Pathway: Shedding New Light on an Old Path". In: *The Journal of Comparative Neurology* 353.4, pp. 477–479.

Satterthwaite, T. D. et al. (Mar. 2012). "Impact of In-Scanner Head Motion on Multiple Measures of Functional Connectivity: Relevance for Studies of Neurodevelopment in Youth". In: *NeuroImage* 60.1, pp. 623–632.

Satterthwaite, T. D. et al. (Jan. 1, 2013). "An Improved Framework for Confound Regression and Filtering for Control of Motion Artifact in the Preprocessing of Resting-State Functional Connectivity Data". In: *NeuroImage* 64, pp. 240–256.

Scheef, L. et al. (2017). "Functional Laterality of Task-Evoked Activation in Sensorimotor Cortex of Preterm Infants: An Optimized 3 T fMRI Study Employing a Customized Neonatal Head Coil". In: *PloS One* 12.1, e0169392.

Scherrer, G. et al. (June 12, 2009). "Dissociation of the Opioid Receptor Mechanisms That Control Mechanical and Heat Pain". In: *Cell* 137.6, pp. 1148–1159.

Schuh, A. et al. (Jan. 28, 2018). "Unbiased Construction of a Temporally Consistent Morphological Atlas of Neonatal Brain Development". In: *bioRxiv*, p. 251512.

Seeber, M. et al. (Feb. 14, 2019). "Subcortical Electrophysiological Activity Is Detectable with High-Density EEG Source Imaging". In: *Nature Communications* 10.1, p. 753.

Serag, A. et al. (Feb. 1, 2012). "Construction of a Consistent High-Definition Spatio-Temporal Atlas of the Developing Brain Using Adaptive Kernel Regression". In: *NeuroImage* 59.3, pp. 2255–2265.

Seymour, B. (Mar. 20, 2019). "Pain: A Precision Signal for Reinforcement Learning and Control". In: *Neuron* 101.6, pp. 1029–1041.

Shi, F. et al. (Apr. 14, 2011). "Infant Brain Atlases from Neonates to 1- and 2-Year-Olds". In: *PloS One* 6.4, e18746.

Shi, F. et al. (Jan. 4, 2018). "Functional Brain Parcellations of the Infant Brain and the Associated Developmental Trends". In: *Cerebral Cortex (New York, N.Y.: 1991)* 28.4, pp. 1358–1368.

Siegel, J. S. et al. (Sept. 1, 2017). "Data Quality Influences Observed Links Between Functional Connectivity and Behavior". In: *Cerebral Cortex (New York, N.Y.: 1991)* 27.9, pp. 4492–4502.

Slater, R. et al. (Apr. 5, 2006). "Cortical Pain Responses in Human Infants". In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 26.14, pp. 3662–3666.

Slater, R. et al. (June 24, 2008). "How Well Do Clinical Pain Assessment Tools Reflect Pain in Infants?" In: *PLoS medicine* 5.6, e129.

Slater, R. et al. (Mar. 2010a). "Evoked Potentials Generated by Noxious Stimulation in the Human Infant Brain". In: *European Journal of Pain (London, England)* 14.3, pp. 321–326.

Slater, R. et al. (Oct. 9, 2010b). "Oral Sucrose as an Analgesic Drug for Procedural Pain in Newborn Infants: A Randomised Controlled Trial". In: *Lancet (London, England)* 376.9748, pp. 1225–1232.

Slater, R. et al. (Aug. 15, 2010c). "Premature Infants Display Increased Noxious-Evoked Neuronal Activity in the Brain Compared to Healthy Age-Matched Term-Born Infants". In: *NeuroImage* 52.2, pp. 583–589.

Slugg, R. M., J. N. Campbell, and R. A. Meyer (May 12, 2004). "The Population Response of A- and C-Fiber Nociceptors in Monkey Encodes High-Intensity Mechanical Stimuli". In: *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 24.19, pp. 4649–4656.

Smith, S. M. (Nov. 2002). "Fast Robust Automated Brain Extraction". In: *Human Brain Mapping* 17.3, pp. 143–155.

Smith, S. M. and J. M. Brady (May 1, 1997). "SUSAN—A New Approach to Low Level Image Processing". In: *International Journal of Computer Vision* 23.1, pp. 45–78.

Smith, S. M., G. Kindlmann, and S. Jbabdi (Jan. 1, 2014). "Chapter 10 - Cross-Subject Comparison of Local Diffusion MRI Parameters". In: *Diffusion MRI (Second Edition)*. Ed. by H. Johansen-Berg and T. E. J. Behrens. San Diego: Academic Press, pp. 209–239.

Smith, S. M. and T. E. Nichols (Jan. 1, 2009). "Threshold-Free Cluster Enhancement: Addressing Problems of Smoothing, Threshold Dependence and Localisation in Cluster Inference". In: *NeuroImage* 44.1, pp. 83–98.

Smith, S. M. et al. (2004). "Advances in Functional and Structural MR Image Analysis and Implementation as FSL". In: *NeuroImage* 23 Suppl 1, S208–219.

Smith, S. M. et al. (Mar. 2005). "Variability in fMRI: A Re-Examination of Inter-Session Differences". In: *Human Brain Mapping* 24.3, pp. 248–257.

Smith, S. M. et al. (July 15, 2006). "Tract-Based Spatial Statistics: Voxelwise Analysis of Multi-Subject Diffusion Data". In: *NeuroImage* 31.4, pp. 1487–1505.

Smith, S. M. et al. (Aug. 4, 2009). "Correspondence of the Brain's Functional Architecture during Activation and Rest". In: *Proceedings of the National Academy of Sciences of the United States of America* 106.31, pp. 13040–13045.

Smith, S. M. et al. (Jan. 15, 2011). "Network Modelling Methods for FMRI". In: *NeuroImage* 54.2, pp. 875–891.

Smyser, C. D. et al. (Dec. 2010). "Longitudinal Analysis of Neural Network Development in Preterm Infants". In: *Cerebral Cortex (New York, N.Y.: 1991)* 20.12, pp. 2852–2862.

Snoek, L., S. Miletić, and H. S. Scholte (Jan. 1, 2019). "How to Control for Confounds in Decoding Analyses of Neuroimaging Data". In: *NeuroImage* 184, pp. 741–760.

Sotiropoulos, S. N. et al. (Jan. 7, 2016). "Fusion in Diffusion MRI for Improved Fibre Orientation Estimation: An Application to the 3T and 7T Data of the Human Connectome Project". In: *NeuroImage* 134, pp. 396–409.

Stepniewska, I. et al. (Jan. 13, 2003). "Somatosensory Input to the Ventrolateral Thalamic Region in the Macaque Monkey: Potential Substrate for Parkinsonian Tremor". In: *The Journal of Comparative Neurology* 455.3, pp. 378–395.

Steven, A. J., J. Zhuo, and E. R. Melhem (Jan. 2014). "Diffusion Kurtosis Imaging: An Emerging Technique for Evaluating the Microstructural Environment of the Brain". In: *AJR. American journal of roentgenology* 202.1, W26–33.

Stevens, B. et al. (1996). "Premature Infant Pain Profile: Development and Initial Validation". In: *The Clinical Journal of Pain* 12.1, pp. 13–22.

Stevens, B. et al. (2010). "The Premature Infant Pain Profile: Evaluation 13 Years after Development". In: *The Clinical Journal of Pain* 26.9, pp. 813–830.

Stevens, B. J. et al. (Mar. 2014). "The Premature Infant Pain Profile-Revised (PIPP-R): Initial Validation and Feasibility". In: *The Clinical Journal of Pain* 30.3, pp. 238–243.

Stevens, R. T., S. M. London, and A. V. Apkarian (1993). "Spinothalamocortical Projections to the Secondary Somatosensory Cortex (SII) in Squirrel Monkey". In: *Brain Research* 631.2, pp. 241–246.

Sugiura, Y., C. L. Lee, and E. R. Perl (Oct. 17, 1986). "Central Projections of Identified, Unmyelinated (C) Afferent Fibers Innervating Mammalian Skin". In: *Science (New York, N.Y.)* 234.4774, pp. 358–361.

Sugiura, Y., N. Terui, and Y. Hosoya (Oct. 1989). "Difference in Distribution of Central Terminals between Visceral and Somatic Unmyelinated (C) Primary Afferent Fibers". In: *Journal of Neurophysiology* 62.4, pp. 834–840.

Sullivan, M. J. L. (May 2008). "Toward a Biopsychomotor Conceptualization of Pain: Implications for Research and Intervention". In: *The Clinical Journal of Pain* 24.4, pp. 281–290.

Suzuki, Y. et al. (Aug. 2003). "Absolute Eigenvalue Diffusion Tensor Analysis for Human Brain Maturation". In: *NMR in biomedicine* 16.5, pp. 257–260.

Tabor, A. and C. Burr (Apr. 1, 2019). "Bayesian Learning Models of Pain: A Call to Action". In: *Current Opinion in Behavioral Sciences*. Pain and Aversive Motivation 26, pp. 54–61.

Tabor, A. et al. (Jan. 2017). "Pain: A Statistical Account". In: *PLoS computational biology* 13.1, e1005142.

Tavor, I. et al. (Apr. 8, 2016). "Task-Free MRI Predicts Individual Differences in Brain Activity during Task Performance". In: *Science* 352.6282, pp. 216–220.

Tracey, I., C. J. Woolf, and N. A. Andrews (June 3, 2019). "Composite Pain Biomarker Signatures for Objective Assessment and Effective Treatment". In: *Neuron* 101.5, pp. 783–800.

Treede, R. D. et al. (Feb. 1999). "The Cortical Representation of Pain". In: *Pain* 79.2-3, pp. 105–111.

Treede, R.-D. (Mar. 2018). “The International Association for the Study of Pain Definition of Pain: As Valid in 2018 as in 1979, but in Need of Regularly Updated Footnotes”. In: *Pain Reports* 3.2, e643.

Tuch, D. S. et al. (Dec. 4, 2003). “Diffusion MRI of Complex Neural Architecture”. In: *Neuron* 40.5, pp. 885–895.

Turner, R. and S. Geyer (Sept. 2014). “Comparing like with like: The Power of Knowing Where You Are”. In: *Brain Connectivity* 4.7, pp. 547–557.

Vaart, M. et al. (2019). “Multimodal Pain Assessment Improves Discrimination between Noxious and Non-noxious Stimuli in Infants”. In: *Paediatric and Neonatal Pain* 1.1, pp. 21–30.

vanPraag, H. and H. Frenk (Dec. 17, 1991). “The Development of Stimulation-Produced Analgesia (SPA) in the Rat”. In: *Brain Research. Developmental Brain Research* 64.1-2, pp. 71–76.

Verriotis, M. et al. (2016). “Mapping Cortical Responses to Somatosensory Stimuli in Human Infants with Simultaneous Near-Infrared Spectroscopy and Event-Related Potential Recording”. In: *eNeuro* 3.2.

Vogt, N. (Dec. 2015). “fMRI Goes Individual”. In: *Nature Methods* 12.12, pp. 1112–1113.

Vos, S. B. et al. (Feb. 1, 2012). “The Influence of Complex White Matter Architecture on the Mean Diffusivity in Diffusion Tensor MRI of the Human Brain”. In: *NeuroImage* 59.3, pp. 2208–2216.

Wager, T. D. et al. (Apr. 11, 2013). “An fMRI-Based Neurologic Signature of Physical Pain”. In: *The New England Journal of Medicine* 368.15, pp. 1388–1397.

Wall, P. D. (Mar. 1, 1996). “Comments after 30 Years of the Gate Control Theory”. In: *Pain Forum* 5.1, pp. 12–22.

Wanigasekera, V. et al. (Feb. 2018). “Disambiguating Pharmacological Mechanisms from Placebo in Neuropathic Pain Using Functional Neuroimaging”. In: *British Journal of Anaesthesia* 120.2, pp. 299–307.

Westlund, K. N. et al. (1991). “Noradrenergic Innervation of Somatosensory Thalamus and Spinal Cord”. In: *Progress in Brain Research* 88, pp. 77–88.

Wheeler-Kingshott, C. A. M. and M. Cercignani (May 2009). “About "Axial" and "Radial" Diffusivities”. In: *Magnetic Resonance in Medicine* 61.5, pp. 1255–1260.

Whitehead, K. et al. (May 1, 2019). “The Emergence of Hierarchical Somatosensory Processing in Late Prematurity”. In: *Cerebral Cortex (New York, N.Y.: 1991)* 29.5, pp. 2245–2260.

Williams, G. et al. (Feb. 1, 2015). “Functional Magnetic Resonance Imaging Can Be Used to Explore Tactile and Nociceptive Processing in the Infant Brain”. In: *Acta Paediatrica* 104.2, pp. 158–166.

Winkler, A. M. et al. (May 15, 2014). “Permutation Inference for the General Linear Model”. In: *NeuroImage* 92, pp. 381–397.

Winkler, A. M. et al. (Apr. 2016). “Non-Parametric Combination and Related Permutation Tests for Neuroimaging”. In: *Human Brain Mapping* 37.4, pp. 1486–1511.

Woo, C.-W., A. Krishnan, and T. D. Wager (May 1, 2014). "Cluster-Extent Based Thresholding in fMRI Analyses: Pitfalls and Recommendations". In: *NeuroImage* 91, pp. 412–419.

Woody, C. D. (Nov. 1, 1967). "Characterization of an Adaptive Filter for the Analysis of Variable Latency Neuroelectric Signals". In: *Medical and biological engineering* 5.6, pp. 539–554.

Woolf, C. J. (Nov. 2010). "What Is This Thing Called Pain?" In: *The Journal of Clinical Investigation* 120.11, pp. 3742–3744.

Woolrich, M. W. et al. (Dec. 2001). "Temporal Autocorrelation in Univariate Linear Modeling of FMRI Data". In: *NeuroImage* 14.6, pp. 1370–1386.

Woolrich, M. W., T. E. J. Behrens, and S. M. Smith (Apr. 2004). "Constrained Linear Basis Sets for HRF Modelling Using Variational Bayes". In: *NeuroImage* 21.4, pp. 1748–1761.

Xu, J. et al. (Mar. 28, 2014). "Spontaneous Neuronal Activity Predicts Intersubject Variations in Executive Control of Attention". In: *Neuroscience* 263, pp. 181–192.

Xu, J. et al. (Dec. 2013). "Evaluation of Slice Accelerations Using Multiband Echo Planar Imaging at 3 T". In: *NeuroImage* 83, pp. 991–1001.

Yendiki, A. et al. (Mar. 2014). "Spurious Group Differences Due to Head Motion in a Diffusion MRI Study". In: *NeuroImage* 88, pp. 79–90.

Yezierski, R. P. (Jan. 1, 1988). "Spinomesencephalic Tract: Projections from the Lumbosacral Spinal Cord of the Rat, Cat, and Monkey". In: *The Journal of Comparative Neurology* 267.1, pp. 131–146.

Yoshimura, M. and H. Furue (June 2006). "Mechanisms for the Anti-Nociceptive Actions of the Descending Noradrenergic and Serotonergic Systems in the Spinal Cord". In: *Journal of Pharmacological Sciences* 101.2, pp. 107–117.

Zhang, H. et al. (July 16, 2012). "NODDI: Practical in Vivo Neurite Orientation Dispersion and Density Imaging of the Human Brain". In: *NeuroImage* 61.4, pp. 1000–1016.

Zhang, X. et al. (Aug. 1999). "Physiological Studies of Spinohypothalamic Tract Neurons in the Lumbar Enlargement of Monkeys". In: *Journal of Neurophysiology* 82.2, pp. 1054–1058.

Zhang, Y., M. Brady, and S. Smith (Jan. 2001). "Segmentation of Brain MR Images through a Hidden Markov Random Field Model and the Expectation-Maximization Algorithm". In: *IEEE transactions on medical imaging* 20.1, pp. 45–57.

Zou, Q. et al. (Dec. 2013). "Intrinsic Resting-State Activity Predicts Working Memory Brain Activation and Behavioral Performance". In: *Human Brain Mapping* 34.12, pp. 3204–3215.