Born Connected: Do Infants Already Have Adult-Like Multi-Scale Connectivity Networks?

Prerana Bajracharya †⸸*, Shiva Mirzaeian †, Zening Fu †, Vince Calhoun †, Sarah Shultz §, Armin Iraji †⸸*

† Tri-Institutional Center for Translational Research in Neuroimaging and Data Science (Georgia State University, Georgia Institute of Technology, Emory University), Atlanta, GA 30303, USA
⸸ Department of Computer Science, Georgia State University, Atlanta, GA 30303, USA
§ Division of Autism & Related Disabilities, Department of Pediatrics, Emory University School of Medicine, Atlanta, GA 30322, USA

** Corresponding author Contact Information: Department of Computer Science, Georgia State University; P.O. Box 5010; Atlanta, GA 30302-5010; phone: (404) 413-6217; email: pbajracharya1@gsu.edu; airaji@gsu.edu*

Abstract

The human brain undergoes remarkable development with the first six postnatal months witnessing the most dramatic structural and functional changes, making this period critical for in-depth research. rsfMRI studies have identified intrinsic connectivity networks (ICNs), including the default mode network, in infants. Although early formation of these networks has been suggested, the specific identification and number of ICNs reported in infants vary widely, leading to inconclusive findings. In adults, ICNs have provided valuable insights into brain function, spanning various mental states and disorders. A recent study analyzed data from over 100,000 subjects and generated a template of 105 multi-scale ICNs enhancing replicability and generalizability across studies. Yet, the presence of these ICNs in infants has not been investigated.

This study addresses this significant gap by evaluating the presence of these multi-scale ICNs in infants, offering critical insight into the early stages of brain development and establishing a baseline for longitudinal studies. To accomplish this goal, we employ two sets of analyses. First, we employ a fully data-driven approach to investigate the presence of these ICNs from infant data itself. Towards this aim, we also introduce burst independent component analysis (bICA), which provides reliable and unbiased network identification. The results reveal the presence of these multi-scale ICNs in infants, showing a high correlation with the template (rho > 0.5), highlighting the potential for longitudinal continuity in such studies.

We next demonstrate that reference-informed ICA-based techniques can reliably estimate these ICNs in infants, highlighting the feasibility of leveraging the NeuroMark framework for robust brain network extraction. This approach not only enhances cross-study comparisons across lifespans but also facilitates the study of brain changes across different age ranges.

In summary, our study highlights the novel discovery that the infant brain already possesses ICNs that are widely observed in older cohorts.

Keywords: infants, rsfMRI, brain maturation, multiscale, intrinsic connectivity network (ICN), burst independent component analysis (bICA), NeuroMark framework, NeuroMark 2.2 template

1 Introduction

1.1 Resting-state functional MRI and Intrinsic Connectivity Networks

Resting-state functional magnetic resonance imaging (rsfMRI) has revolutionized our understanding of the human brain's functional organization. Unlike task-based fMRI, which focuses on identifying brain activity related to specific experimental tasks, rsfMRI focuses on capturing and studying the brain's intrinsic activity (B. Biswal et al., 1995; B. B. Biswal, 2012). This intrinsic activity is characterized by spontaneous fluctuations in blood-oxygenation-level-dependent (BOLD) signals, suggested to reflect changes in regional neural activity and metabolism (B. Biswal et al., 1995). One of the key advantages of rsfMRI is its ability to unveil intrinsic connectivity networks (ICNs) within the brain (Calhoun et al., 2001; Damoiseaux et al., 2006). These ICNs represent coordinated patterns of activity among spatially distributed brain regions that exhibit synchronized fluctuations in BOLD signal (Calhoun et al., 2001; Iraji et al., 2023). ICNs have significantly improved our understanding of various cognitive processes, including sensory perception, motor control, memory, and attention (B. Biswal et al., 1995; Damoiseaux et al., 2006; De Luca et al., 2006; Fransson, 2006; Hunter et al., 2006; Laird et al., 2011; Nir et al., 2006). Changes or disruptions in ICNs have been associated with various neurological and psychiatric disorders, including autism spectrum disorder, schizophrenia, Alzheimer's disease, and depression (Badhwar et al., 2017; Bi et al., 2018; Chang et al., 2018; Du et al., 2020; Fu et al., 2020; Iraji et al., 2015, 2019, 2022, 2023; Jensen, Calhoun, et al., 2024; Uddin et al., 2013).

To enhance the generalizability and reliability of findings across studies, researchers have recently delineated 105 multi-scale ICNs (NeuroMark_fMRI_2.2_modelorder-multi) from a large dataset of over 100k rsfMRI subjects (Iraji et al., 2023) , which can serve as a robust template for comparative analysis across multiple datasets and studies (Andrés-Camazón et al., 2024; Iraji et al., 2023). However, further research is needed to investigate the presence of these ICNs in infants.

1.2 Intrinsic Connectivity Networks (ICNs) in Infants

The first six postnatal months are characterized by significant brain maturation, supported by concurrent changes in brain anatomy (Huttenlocher & Dabholkar, 1997; Knickmeyer et al., 2008; Bethlehem et al., 2022) and functional networks. Research on infant functional network development and early childhood has emerged in the last two decades, initially focusing on preterm newborns (Fransson et al., 2007), followed by studies on full-term infants (Liu et al., 2008).

Liu et al. identified two networks, a unilateral sensorimotor network in either the right or left hemisphere, associated with hand movements, and a midline sensorimotor network, linked to the foot, leg, and supplementary motor areas (Liu et al., 2008). Studies indicate that the basic architecture of proto-networks and the proto-default mode network (DMN) are present at birth and develop dramatically throughout the first year of life (Gao, Alcauter, Elton, et al., 2015). Gao and colleagues (2015) mapped the maturation trajectories of nine main resting-state networks—Medial occipital network (V1), Occipital pole network (V2), Lateral visual/parietal network (V3), DMN, sensorimotor network (SM), auditory/language network (AN), salience network (SA), and the left and right frontoparietal networks (FP)—in 65 infants born at 35 to 42 weeks gestation with longitudinal scans taken every 3 months till one year of age. Their findings suggest that the sensorimotor and auditory networks are largely established at birth, while the visual network is the second earliest ICN to develop in the newborn brain. The dorsal attention network (DAN) and DMN undergo continuous and significant maturation and can only achieve adult-like network architecture by the end of the first year. The salience and bilateral frontoparietal networks

exhibit an even slower rate of development compared to the others (Gao, Alcauter, Elton, et al., 2015). In a recent study using the developing Human Connectome Project (dHCP) dataset Eyre et al. identified 11 ICNs in 24 term-born infants scanned at 43.5–44.5 weeks postmenstrual age or 3.5 to 4.5 weeks of corrected age (Eyre et al., 2021). These ICNs were divided into two categories: primary networks and association networks. The five primary networks—associated with motor and sensory functions—include the medial motor, lateral motor, somatosensory, auditory, and visual networks (Eyre et al., 2021). The remaining six networks, classified as association networks, are motor association, temporoparietal, posterior parietal, frontoparietal, prefrontal, and visual association (Eyre et al., 2021). In another study of neonates aged 0 to 5 weeks, Molloy and Saygin (2022) attempted to identify adult-like neural networks. Their findings revealed the presence of only five ICNs—sensorimotor, visual, default mode, ventral attention, and dorsal attention (Molloy & Saygin, 2022)—compared to the seven ICNs typically observed in adults (Thomas Yeo et al., 2011). Notably, the limbic and control networks were absent in the neonatal scans. Another study with equivalent age to dHCP in term and preterm infants reported the presence of auditory, sensorimotor, visual, and frontoparietal regions (França et al., 2024).

In addition to inconsistencies in identifying specific brain networks, there is significant variability in the number of networks reported across studies (Eyre et al., 2021; Gao, Alcauter, Elton, et al., 2015; Liu et al., 2008; Molloy & Saygin, 2022; Seraji et al., 2024). For example, while some research identifies as few as five networks in neonates (Molloy & Saygin, 2022), others detect up to eleven distinct networks (Eyre et al., 2021). This variability complicates efforts to integrate findings and construct a cohesive understanding of brain reorganization throughout development. Moreover, the selective focus on particular networks, such as the language network (Emerson et al., 2016) or the default mode network (Gao et al., 2009, 2013), often overlooks other essential networks, resulting in a fragmented understanding of network evolution. These challenges highlight the need for more comprehensive approaches, as advocated by recent calls for standardization (Iraji et al., 2023; Jensen, Turner, et al., 2024; Uddin et al., 2019), to better capture the full spectrum of brain network development in early life, enabling cross-study comparisons and facilitating longitudinal research.

We propose to evaluate the presence of reliable and canonical multi-scale ICNs (Iraji et al., 2023) in full-term, typically developing infants during their first six postnatal months, as guided by the framework established in (Iraji et al., 2023; Jensen, Turner, et al., 2024). By focusing on this critical early developmental window, we intend not only to characterize the ICNs in infants but to provide a comprehensive and holistic understanding of the extent to which adult-like ICNs are present in the first 180 days after birth, employing both fully data-driven (in Section 1.3.1) and guided approaches (in Section 1.3.2), addressing key gaps in the existing literature.

1.3 Extracting Intrinsic Connectivity Networks (ICNs)

Independent component analysis (ICA) is the primary tool to extract ICNs from rsfMRI (Mckeown et al., 1998). It is a data-driven multivariate method that estimates ICNs as temporally coherent spatial patterns that do not require prior information about the brain's functional architecture and pre-selected brain regions (Calhoun et al., 2001; Calhoun & Adalı, 2012; Calhoun & de Lacy, 2017). By decomposing the fMRI signal into statistically independent components, ICA can isolate functionally distinct ICNs in the brain. Additionally, ICA excels in separating meaningful brain activity from various sources of noise, including physiological artifacts such as cardiac and respiratory signals, thereby enhancing the reliability of the identified ICNs (Mckeown et al., 1998). ICA has been extensively applied in large cohorts and across diverse imaging modalities, demonstrating its capability to uncover brain functional networks and connectivity features where conventional regression-based methods often fall short (Calhoun et al., 2009; Calhoun & Adali, 2006; McKeown et al., 2003). Recent advances in infant neuroimaging have utilized ICA to identify networks during the first year of life that closely resemble

those seen in adults, in fMRI studies (Eyre et al., 2021; Fransson, 2006; Gao, Alcauter, Smith, et al., 2015; Gao et al., 2009; Liu et al., 2008). These findings highlight the potential of ICA to provide deeper insights into early brain development.

In this study, we leveraged the advantages of ICA by applying two variations to extract ICNs from infant rsfMRI data: blind ICA and semi-blind ICA. The blind ICA generates ICNs that are fully driven by infant rsfMRI data. On the other hand, semi-blind ICA leverages prior templates to guide the ICN identification facilitating the generalizability of findings and comparability across studies.

1.3.1 Burst Independent Component Analysis (bICA): A blind ICA

Blind ICA or simply ICA is a fully data-driven method used in neuroimaging to identify ICNs rsfMRI data. It works by decomposing the observed data into statistically independent components, some of which represent distinct neural-related networks, aka ICNs along with artifacts and noise that are identified using post hoc analysis based on temporal or spatial criteria (Calhoun & Adali, 2006). Unlike approaches such as seed-based or ROI-based analyses where a pre-defined region needs to be specified, no information is assumed about the time courses or spatial maps in blind ICA and ICNs are directly estimated from the data itself thus reducing the potential for bias in the analysis, and interpretation of the data (Beckmann & Smith, 2004; Calhoun et al., 2001; Neha & Gandhi, 2016).

Model order in ICA specifies the number of independent components (ICs) estimated from the data. Researchers have struggled to determine the optimal model order in fMRI data analysis (Abou-Elseoud et al., 2010; Beckmann, 2012; Kuang et al., 2018; Iraji et al., 2021). Setting the model order too low can lead to incomplete separation of independent signals (Allen et al., 2012), while, increasing the model order beyond the true dimensionality may result in the fragmentation of sources into multiple components (Abou-Elseoud et al., 2010). Several approaches have been proposed to address this issue, including information-theoretic criteria (ITC) (Y.-O. Li et al., 2007) such as the minimum description length (MDL) criterion (Rissanen, 1978), the Akaike information criterion (AIC) (Akaike, 1998), and the Bayesian information criterion (BIC) (Rissanen, 1978). Despite these efforts, discrepancies in the optimal number of components still arise depending on the estimation method, leaving the model-order selection problem unresolved.

Recently, Iraji et al. proposed a novel approach called multi-scale ICA (msICA), which generates multi-scale ICNs (Iraji et al., 2022). Unlike previous methods that focus on determining a single optimum model order, msICA incorporates information across multiple spatial model orders to obtain an optimal estimation of ICNs, accommodating diverse sizes. By integrating ICNs from multiple scales, this technique aims to create a comprehensive representation of brain connectivity. This approach was applied to a large dataset of over 100,000 rsfMRI subject data to develop a universal template (Iraji et al., 2023). Leveraging ICA with multiple model orders not only eliminates the challenges associated with model order selection, notoriously infamous as one of the major challenges of blind ICA (Abou-Elseoud et al., 2010; Beckmann, 2012; Iraji et al., 2019, 2022, 2023), but also improves the stability of ICA (Iraji et al., 2019, 2022, 2023; H. Li et al., 2018; Meng et al., 2022, 2023). It also produces multi-scale components that capture functional segregation across various levels of granularity in the brain. This enables the identification of functional sources and their interactions in greater detail.

Inspired by groundbreaking advances in multi-model-order analysis, we introduce burst independent component analysis (bICA), a pioneering approach poised to revolutionize the estimation of ICNs. Unlike conventional methods that depend on a single model order or the limited spatial scales of msICA, bICA is engineered to harness the full potential of multi-scale analysis. It systematically explores an expansive range of model orders to capture different spatial scales of brain systems due to a wider and finely-grained range of model

orders, offering a more exhaustive and precise estimation of brain functional networks that were previously elusive. The strategic flexibility is advantageous given the varying complexity across brain systems, as different regions or systems may not conform to the same spatial scale within the brain's functional hierarchy (Iraji et al., 2019, 2022, 2023; H. Li et al., 2018; Meng et al., 2022, 2023), enabling the delineation of both large-scale and micro-scale networks, providing a comprehensive and holistic view of brain functional architecture (Iraji et al., 2019, 2022, 2023; Meng et al., 2022, 2023). By obviating the need for pre-defined model orders, bICA minimizes variability and enhances the reliability and stability of ICN identification. This methodological breakthrough not only ensures consistency across different studies and populations, increasing the generalizability of blind ICA but also uncovers new and potentially transformative networks. With its sophisticated, incremental multi-scale analysis, bICA promises to elevate neuroimaging to new heights, offering profound insights into the brain's complex connectivity and advancing our understanding of its functional organization.

We employ bICA to estimate ICNs from the infant dataset and compare them with the NeuroMark 2.2 template, to confirm the feasibility of the method and the existence of adult-like or matured ICNs. To validate the accuracy of our estimated ICNs and ensure they are not unduly influenced by the method itself, we conduct simulation-based validation. We generate a null dataset with ICNs intentionally selected to exhibit minimal similarity to the target ICNs, ensuring the absence of the target ICN's characteristics. We then apply bICA to this null dataset, estimate the components, and evaluate the similarity between the obtained components and the template. This validation process confirms that our method accurately identifies ICNs and ensures that the estimated signals correspond to those present within the data.

1.3.2 NeuroMark Framework: A reference-informed approach

While blind ICA is fully data-driven and generates unbiased components, it requires post hoc analysis to match the components of interest, which is not always ideal. This process can involve redundant computations, demand significant memory for estimating unnecessary components, and often degrade the quality of recovered components, especially in high-dimensional applications (Lu & Rajapakse, 2006). Researchers have proposed that the full potential of ICA could be realized by incorporating prior information during component estimation (Suzuki et al., 2001). Reference-informed ICA methods, such as semi-blind ICA (Lin et al., 2010), refine the extraction process by incorporating spatial priors to guide decomposition. This approach enhances the stability and interpretability of the resulting ICNs, allowing for more accurate subject-level mapping of ICNs (Du & Fan, 2013). An advanced form of this method, Multi-Objective Optimization ICA with Reference (MOO-ICAR) (Du et al., 2020), simultaneously maximizes the independence and spatial specificity of components. MOO-ICAR dynamically adjusts the decomposition based on the data characteristics and the defined objectives, thereby eliminating the need for post hoc analysis present in blind ICA (Du & Fan, 2013). This method combines the strengths of data-driven ICA with the robustness of template-based approaches, ensuring more precise and reliable identification of ICNs across diverse subjects and conditions.

However, implementing a reference-based approach like semi-blind ICA necessitates a reliable template. Recognizing the need for enhanced templates with greater generalizability and granularity, Iraji et al. developed a comprehensive template, known as the multi-scale NeuroMark 2.2 template (Iraji et al., 2023). This template represents a substantial improvement over its predecessor, comprising 105 canonical multi-scale ICNs handpicked from 900 multi-scale ICNs after repeated independent analyses and spatial similarity assessments (Iraji et al., 2023). The multi-scale NeuroMark 2.2 template was derived from rsfMRI data of over 100,000 individuals, including healthy as well as pathological subjects, providing refined and advanced multi-scale ICNs (Iraji et al., 2023). These multi-scale ICNs elucidate functional sources and their interactions across multiple spatial scales, capturing functional segregation at different levels of granularity across the brain (Iraji et al., 2019, 2022, 2023). The 105 ICNs are classified into seven distinct domains based on anatomical and functional characteristics which

are further divided into subdomains (Jensen, Turner, et al., 2024). The spatial maps are illustrated in Figure 1 and the details division of ICNs are listed in Table 1. This template represents a cornerstone in elucidating the functional connectivity patterns of the brain, distinguished by its capacity to pinpoint the most robust and reproducible ICNs across diverse populations, datasets, research endeavors, and pathological conditions (Iraji et al., 2023).

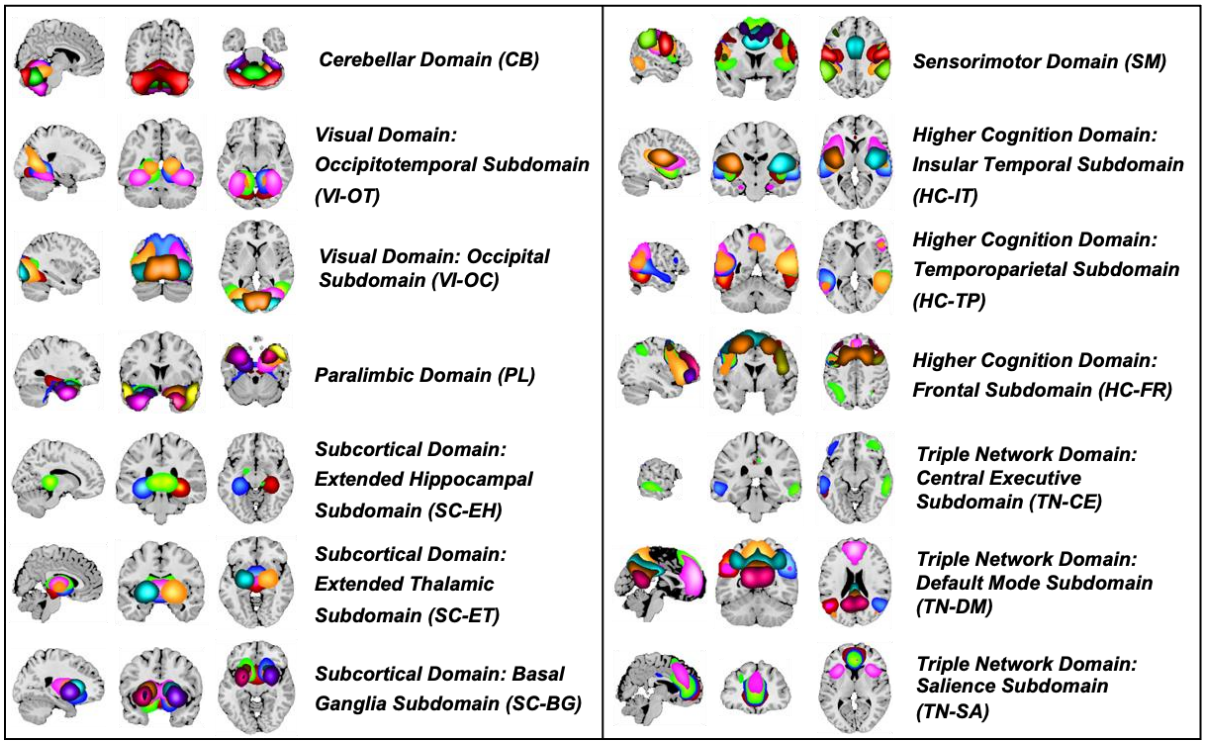

Figure 1 105 ICNs from NeuroMark 2.2 template (NeuroMark_fMRI_2.2_modelorder-multi). Presented as composite maps of 7 domains and 14 subdomains. Cerebellar Domain, Visual Domain (includes Occipitotemporal and Occipital subdomains), Paralimbic Domain, Subcortical Domain (includes Extended Hippocampal, Extended Thalamic and Basal Ganglia subdomains), Sensorimotor Domain, Higher Cognition Domain (includes Insular Temporal, Temporoparietal and Frontal subdomains) and Triple Network Domain (includes Central Executive, Default Mode, and Salience subdomains).

In this study, we utilized the NeuroMark framework, employing MOO-ICAR as the reference-informed method and the multi-scale NeuroMark 2.2 (NeuroMark_fMRI_2.2_modelorder-multi) template - (Iraji et al., 2023) as the reference, which serves as a soft spatial constraint. This approach allows for the precise identification of ICNs in infant data while accounting for the possibility of their absence. Current studies on functional network development in younger and older age groups are usually conducted independently of each other, leading to a lack of a coherent overview. This separation in research focus results in fragmented insights, preventing a comprehensive understanding of the continuum of brain development and aging. By using the directive ICNs from a reference-informed approach, we enhance the accuracy and relevance of the identified networks and can confidently compare them with canonical ICNs. This comparison facilitates a more comprehensive understanding of the status of functional networks in infants, bridging the gap between developmental and aging studies, and offering a more integrated perspective on brain connectivity across the lifespan.

Similar to bICA, we validate that the estimated ICNs are unbiased by applying MOO-ICAR with the target ICN as the reference on a null dataset. This confirms that ICNs cannot be extracted unless present in the data, reinforcing the robustness of our approach. Additionally, we evaluate whether the estimated ICNs retain distinct network-specific information at the subject level in the infant dataset.

Table 1 Domain and Sub-domain labels of ICNs included in NeuroMark 2.2 template (NeuroMark_fMRI_2.2_modelorder-multi)

Domain Label	Domain abbreviation	Sub-Domain Label	Sub-Domain abbreviation	# of ICNs in the sub-domains
Cerebellar Domain	CB	Cerebellar Domain	CB	13
Visual Domain	VI	Occipitotemporal Subdomain	OT	5
	VI-OC	Occipital Subdomain	OC	7
Paralimbic Domain	PL	Paralimbic Domain	PL	11
Subcortical Domain	SC	Extended Hippocampal Subdomain	EH	3
		Extended Thalamic Subdomain	ET	6
		Basal Ganglia Subdomain	BG	9
Sensorimotor Domain	SM	Sensorimotor Domain	SM	14
Higher Cognition Domain	HC	Insular Temporal Subdomain	IT	7
		Temporoparietal Subdomain	TP	5
		Frontal Subdomain	FR	10
Triple Network Domain	TN	Central Executive Subdomain	CE	3
		Default Mode Subdomain	DM	8
		Salience Subdomain	SA	4

To summarize, in this paper, we introduced the bICA approach as blind ICA approach to enhance ICN identification and employed it on an infant dataset. We also employed and investigated the feasibility of using NeuroMark framework, utilizing the MOO-ICAR with the NeuroMark 2.2 template, ensuring the precise identification of ICNs in infants. Validation on a simulated null dataset confirmed the robustness of our findings. These advanced methods, bICA and MOO-ICAR, provide a more detailed and nuanced characterization of multi-scale ICNs in infants, allowing us to detect the early presence of adult-like ICNs with greater precision. This approach deepens our insight into the architecture of early brain networks and lays the groundwork for understanding how deviations from typical development may lead to neurodevelopmental disorders.

2 Methods

2.1 Dataset

Infant data were obtained from prospective longitudinal studies of infants at low- and elevated familial genetic likelihood for autism spectrum disorder, conducted at the Marcus Autism Center in Atlanta, Georgia, USA. Prior to participation, informed written consent was acquired from the parents of all participants, and ethical approval for all study procedures was granted by the Emory University Institutional Review Board. The rsfMRI data was acquired using a 32-channel head coil, with scans conducted during natural sleep. Data was collected using a non-uniform longitudinal sampling design, with data collected from each infant at up to 3 randomized time points between birth and 6 months (see Figure 1 in Supplemental). Detailed demographic information is provided in

Table 2. For all analyses, infant age at each scan was corrected for gestational age at birth, calculated as the postnatal age minus 7*(40 - gestational age in weeks). This corrected age was used to define the six-month postnatal period, corresponding to an age range of 0 to 180 days.

Table 2 Demographics of participant sample. For information categories with a participant number less than the total sample, the N is specified next to the category title.

	(N = 71)
Infant Sex	30f, 41m
Gestational Age at Birth, mean (SD)	39.1wks (1.27)
Race (N=64)	
Black	7.8%
White	87.5%
More than one race	4.7%
Maternal Education (N=60)	
High School	1.7%
Trade/Vocational Training	1.7%
College Courses	5.0%
Associate's Degree	3.3%
College Degree	36.7%
Graduate Degree	51.6%
Household Income (N=61)	
< $40000	4.9%
$40000-$80000	19.7%
$80001-$100000	23.0%
$100001-$150000	26.2%
> $150001	26.2%

2.2 Preprocessing

We performed standard fMRI preprocessing using the FMRIB Software Library (FSL v6.0, https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/) and the Statistical Parametric Mapping (SPM12, http://www.fil.ion.ucl.ac.uk/spm/) toolboxes under the MATLAB 2020b environment. Firstly, the initial ten dummy scans with substantial signal changes were discarded. Following that, we corrected the distortion in the images using SBRef data with phase encoding blips in reverse. After distortion correction, we performed slice timing to correct for these slice-dependent delays. Head motion correction was performed followed by the slice timing to realign all the scans to the reference scan. To normalize the infant data to the standard adult Montreal Neurological Institute (MNI) space, we applied a two-step normalization procedure. We first warped the UNC-BCP 4D Infant Brain Template (https://www.nitrc.org/projects/uncbcp_4d_atlas/) (Howell et al., 2019) into the adult MNI space using the EPI template as the reference. Here, to mitigate introducing an age-specific bias into normalization, we chose the template with the age of 4 months which is the median of the age range of our infant dataset. After having the UNC-BCP template in the adult MNI space, we normalized the infant data using it as the reference. Finally, the normalized fMRI data were spatially smoothed using a Gaussian kernel featuring a 6 mm full width at half maximum (FWHM).

2.2.1 Burst Independent Component Analysis (bICA)

The proposed burst independent component analysis (bICA) method introduces a novel blind ICA approach designed to generate ICNs at multiple spatial scales, effectively mitigating the need to select a single optimal model order. The analysis pipeline, as illustrated in Figure 2, applies bICA to 126 fMRI scans using sequential model orders ranging from 2 to 225, with a step size of 1.

The pre-processed resting-state fMRI (rsfMRI) data underwent a two-stage Principal Component Analysis (PCA) (Calhoun et al., 2001) to effectively reduce dimensionality and optimize memory usage. The first stage

involves a subject-specific spatial PCA, which normalizes the data to ensure that each subject contributes comparably to the shared subspace. This is mathematically represented as

$$X_{m\,(L\times V)} = F^{-1}_{m\,(L\times K)}\, Y_{m\,(K\times V)} \tag{1}$$

where $X_{(C\times V)}$ is the reduced-dimensionality data for subject m, F^{-1} is the reduction matrix obtained from PCA, L is the reduced dimensionality of time points, K is the original number of time points, V represents the number of voxels, and Y_m is the preprocessed data matrix for subject m. This PCA step reduces the dimensionality of the input time points from K to L, effectively capturing the most significant spatial patterns while reducing the computational load for subsequent analyses. After performing the subject-specific PCA, the principal components derived for each subject were concatenated along the time dimension, creating a combined dataset that retains the key features of each subject's brain activity, facilitating group-level analysis. Subsequently, a group-level spatial PCA was applied to the concatenated matrix of subject-level principal components to further reduce the dimensionality and identify patterns common across all subjects. This process is described by

$$X_{(C\times V)} = G^{-1}{}_{(C\times LM)} \begin{bmatrix} F^{-1}_{1} Y_i \\ \vdots \\ F^{-1}_{M} Y_i \end{bmatrix}_{(LM\times V)} \tag{2}$$

Here G^{-1} is the group-level reduction matrix, C is the number of components determined by the chosen model order, M is the total number of subjects, and LM is the product of the reduced dimensionality L and the number of subjects M. This group-level PCA step provides a common subspace that reflects the shared structure across all subjects, preparing the data for the next step of ICA. The group-level principal components derived from the previous step were then used as input for spatial ICA to evaluate group independent components. This is mathematically expressed as

$$X_{(C\times V)} = A_{(C\times C)}\, S_{(C\times V)} \tag{3}$$

In this equation, A represents the mixing matrix that combines the independent components, and S represents the estimated independent sources. The objective of ICA is to estimate an unmixing matrix, W that can transform the mixed signals back into independent sources, described by

$$Y_{(C\times V)} = W_{(C\times C)}\, X_{(C\times V)} \tag{4}$$

where Y is the output representing a good approximation of the independent latent sources S. The Infomax principle, central to ICA, guides the optimization of the unmixing matrix W. According to this principle, the mutual information between the input signals X and the output signals Y should be maximized, which is mathematically represented as

$$I(X,Y) = H(Y) - H(Y|X) \tag{5}$$

Here, $H(Y)$ is the entropy of the output signals, representing the amount of uncertainty or information content, and $H(Y|X)$ is the conditional entropy, minimized in this context. Maximizing the mutual information $I(X,Y)$ is equivalent to maximizing the entropy $H(Y)$ of the output signals, encouraging the components to be as statistically independent as possible, described by

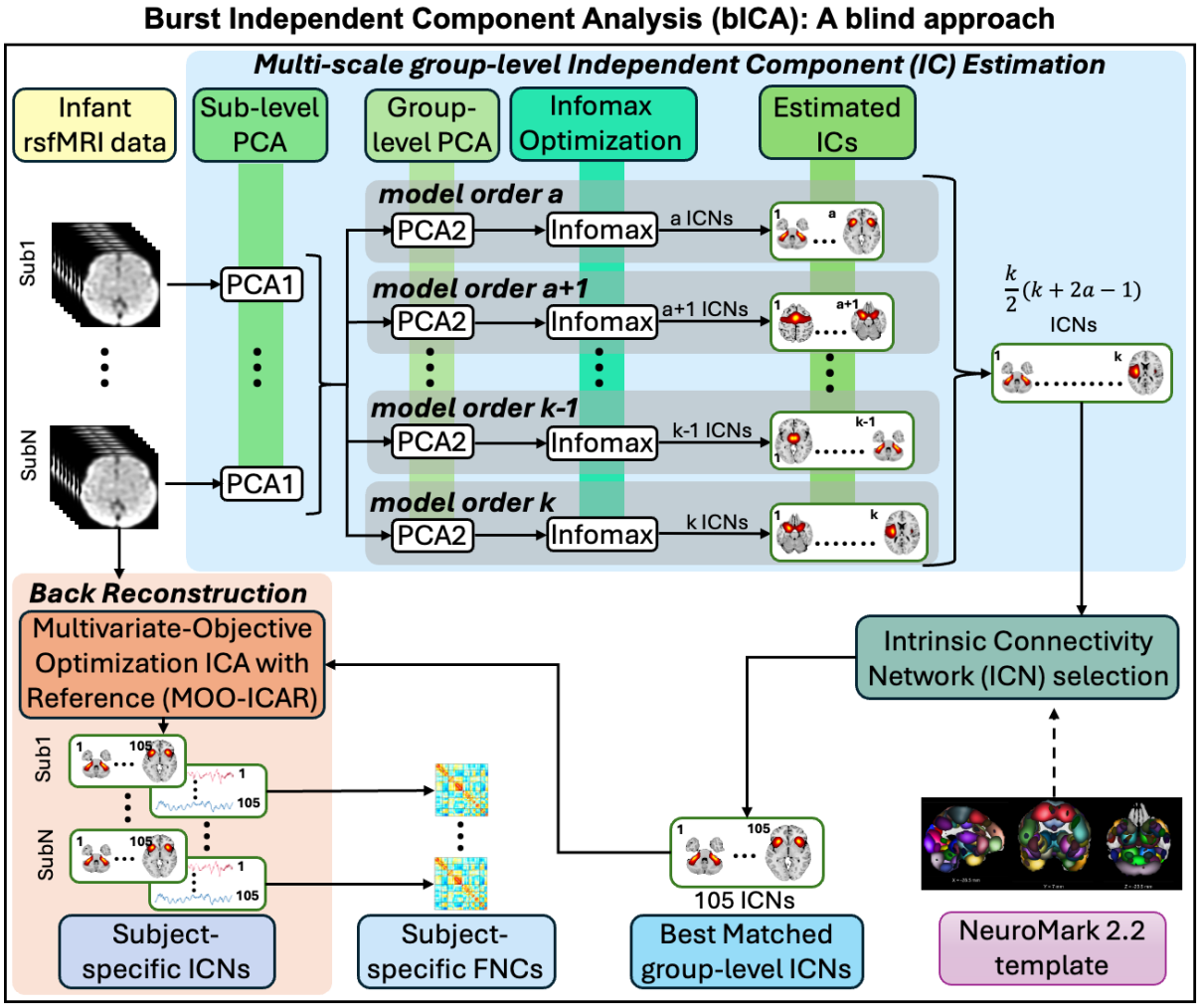

Figure 2 A pipeline for Burst independent component analysis (bICA), a blind approach involving estimation of Independent Components (ICs), ICN selection, and back-reconstruction. The method initiates with subject-level PCA, followed by group-level PCA and Infomax optimization, iterated across multiple model orders to generate a comprehensive set of multi-scale Independent Components (ICs), from which the top 105 ICs were selected based on Pearson's correlation with the NeuroMark 2.2 template. These ICNs served as references for back-reconstruction, specifically MOO-ICAR, to generate subject-level spatial maps and time courses. Subsequently, the time courses were used to generate Functional Network Connectivity (FNC).

$$H(Y) = H(y_1, y_2, \dots, y_c) \tag{6}$$

Assuming that the components are independent, which aligns with the ICA goal, the entropy of the output signals can be rewritten as:

$$H(Y) = \sum_{c=1}^{C} H(y_c) \tag{7}$$

The Infomax ICA algorithm iteratively adjusts W to maximize the entropy of the outputs Y, thereby enhancing the separation of mixed signals into independent components. By optimizing W using the Infomax criterion, the ICA algorithm effectively separates the mixed signals into their group-level ICs. To refine this process, the algorithm iterates through group-level PCA and Infomax ICA for each model order C, where C varies sequentially, $(C = a, a+1, , \dots, , k-1, k)$. For the purpose of our analysis, we used the values of C from 2 to 225 covering a broad range of model orders, from minimalistic to the highest order used in recent studies (Iraji et al., 2023). This

iterative process allows for the exploration of various levels of data granularity, ensuring a thorough analysis across multiple scales. The aggregated group-level ICs were matched with the multi-scale NeuroMark 2.2 template to identify the most similar among all the ICs using Pearson's correlation. The component matching led to the identification of the top 105 best-matched ICNs from infant dataset, providing a robust set of ICNs that align closely with the NeuroMark 2.2 template.

Finally, to obtain subject-level ICN, the back reconstruction step was conducted using spatially constrained ICA that employs a multi-objective function optimization approach - Multi-Objective Optimization Independent Component Analysis with Reference (MOO-ICAR) , which is detailed in Section 2.2.2 (Du & Fan, 2013). This method reconstructs subject-specific networks based on the 105 selected group-level ICNs, generating both spatial maps and time courses for each of the 126 subjects. The resulting time courses were then utilized to evaluate functional network connectivity, offering insights into the functional organization of the infant brain.

2.2.2 NeuroMark Framework: A Reference-Informed Approach

In this section, we introduce a reference-informed approach using the NeuroMark Framework. Specifically, we employed the Multi-Objective Optimization Independent Component Analysis with Reference (MOO-ICAR), a spatially constrained ICA method for estimating ICNs, utilizing the NeuroMark 2.2 template as a reference. This approach allows us to evaluate both the accuracy of ICN estimation in infants and the practical feasibility of NeuroMark Framework. We implemented MOO-ICAR using the GIFT toolbox (Iraji et al., 2021) using the multi-scale NeuroMark 2.2 template (available at https://trendscenter.org/data/) as the spatial prior, facilitating the estimation of corresponding subject-specific ICNs. The details of this methodology are illustrated in Figure 3.

MOO-ICAR is a spatially constrained ICA that employs a predefined template or reference to facilitate the identification of components, such that it aligns with the spatial characteristics of the reference. This approach aims to identify corresponding subject-specific ICNs across different datasets or populations based on a reference, usually obtained from a large dataset or a priori knowledge (Du et al., 2020; Iraji et al., 2023). Mathematically, MOO-ICAR maximizes two objective functions to estimate each subject-specific network corresponding to a given template (Du et al., 2020). The first objective functions, $J(S_I^k)$ aims to optimize the independence of the networks and the second objective function, $F(S_I^k)$ aims to optimize the correspondence between each subject-specific network and the template. Equation 1 summarizes how the Ith network can be estimated for the kth subject using the network template S_I as guidance.

$$max \begin{cases} J(S_I^k) = \{E[G(S_I^k)] - E[G(v)]\}^2 \\ F(S_I^k) = E[S_I S_I^k] \end{cases} \quad s.t. \left|\left|w_I^k\right|\right| = 1 \tag{8}$$

In this formulation, S_I represents the Ith network of the template and $S_I^k = (w_I^k)^T . X^k$ signifies the estimated I^{th} network of the kth subject. X^k is the whitened fMRI data matrix of the kth subject and w_I^k represents the unmixing column vector that needs to be solved within the optimization functions. The first function $J(S_I^k)$ uses negentropy to optimize the independence measure of S_I^k. Here, v denotes a Gaussian variable with mean zero and unit variance. The function $G(\bullet)$ is a nonquadratic function, and $E[\bullet]$ denotes the expectation of the variable. The second function $F(S_I^k)$ measured the correspondence between the template network (S_I) and subject network (S_I^k). The two objective functions are merged using a linear weighted sum with weights set at 0.5 to solve the optimization problem. Therefore, MOO-ICAR yields N subject-specific networks corresponding to N network templates, along with the relevant time courses for each scan.

Using this approach, we estimated 105 subject-specific multi-scale ICNs for each scan, including their spatial maps and time courses. To assess the correspondence between the estimated ICNs and the NeuroMark 2.2

template, we computed Pearson's correlation coefficients using the average spatial maps derived from the estimated ICNs. We then evaluated the unique information contained within each ICN estimated by the NeuroMark Framework for each subject. For each target ICN from the NeuroMark 2.2 template, we identified a comparison ICN that had the second highest spatial correlation with the target ICN, with the estimated ICN itself being the highest and the network of interest. Using these ICNs, we then extracted their equivalents from the infant dataset. We computed Pearson's correlation with the target ICN for the network of interest and the comparison ICNs from the dataset. This analysis allowed us to evaluate the preservation of network-specific information across subjects. The time courses were further used to generate Functional Network Connectivity (FNC) by correlating time courses obtained for each ICN.

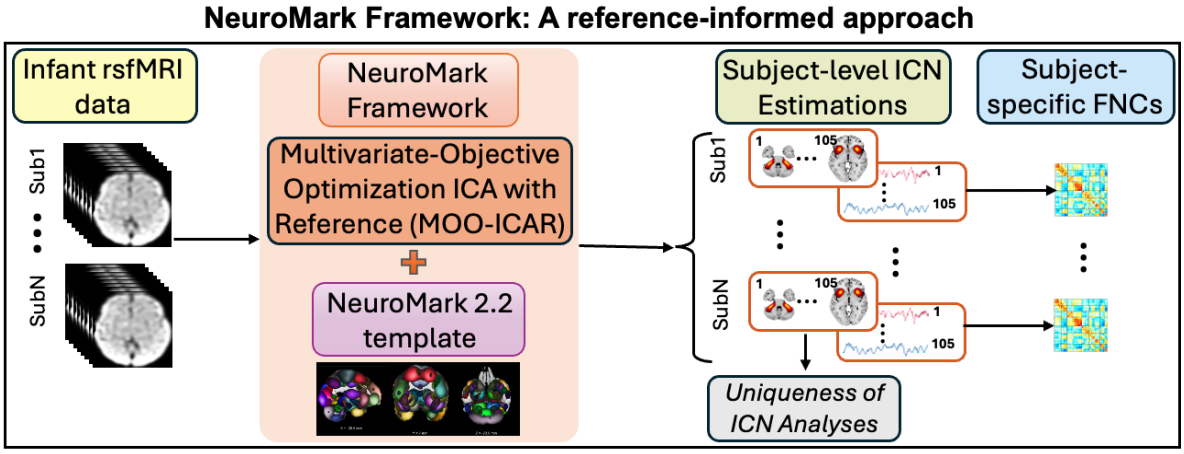

Figure 3 A pipeline for NeuroMark Framework, a reference-informed approach that uses MOO-ICAR method along with NeuroMark 2.2 Template as a` reference to generate precisely 105 subject-level ICN estimations from the infant dataset. The spatial maps are further used to analyze the unique information in each ICN, and the time courses are used to evaluate Functional Network Connectivity (FNC).

2.3 Reliability Test: Null Dataset Analysis

To ensure the reliability of the extracted ICNs, we generated a null dataset and performed extensive analyses using both blind ICA and semi-blind ICA on the simulated null dataset. ICA is based on the assumption that data can be linearly decomposed into spatially independent components, which we refer to as spatial maps (SMs), and their time courses (TCs) (Mckeown et al., 1998). Following ICA principles, we employed a spatiotemporally separable data simulation model to generate realistic null data that closely resembles actual rsfMRI data but deliberately excludes any target or desired ICN component. For each target ICN, null data were created by multiplying 20 ICNs with the lowest spatial correlation to the target by randomly generated time courses. This process was repeated multiple times, with each iteration representing a simulated subject. The process can be mathematically represented as in Equation 2.

$$Xnull_{i\,(V\times T)}^{n} = SM_{i\,(V\times C)}^{n}\, TC_{i\,(C\times T)}^{n} \tag{9}$$

Here, $Xnull_{i}^{n}$ is the null fMRI of dimension voxel (V) and time (T) simulated for i^{th} network and n^{th} subject. SM represents the spatial maps with the dimension voxel (V) by components (C) and TC represents the time course maps with the dimension time (T) by components (C).

The bICA was employed on the null data to assess the similarity of extracted components with the target ICN. The group-level ICs were obtained from the null dataset and subject-level ICs were extracted; a procedure identical to the one followed in the infant dataset. We compared the spatial similarities of estimated components

obtained from the actual data and null data with the template to evaluate the statistical significance of the ICNs estimated from the infant dataset. We also implemented the NeuroMark framework, MOO-ICAR using the target ICN from the template as the reference and evaluated whether MOO-ICAR produces similar components in null data, determining if it forces the generation of components similar to the references even when they are not present in the data.

3 Results

3.1 Burst Independent Component Analysis (bICA)

3.1.1 Estimation of reliable ICNs

The bICA analysis identified all 105 ICNs with high spatial similarity to the multi-scale NeuroMark 2.2 template. Figure 4 presents these 105 best-matched ICNs from the group-level bICA as composite maps, organized into 7 domains and 14 subdomains.

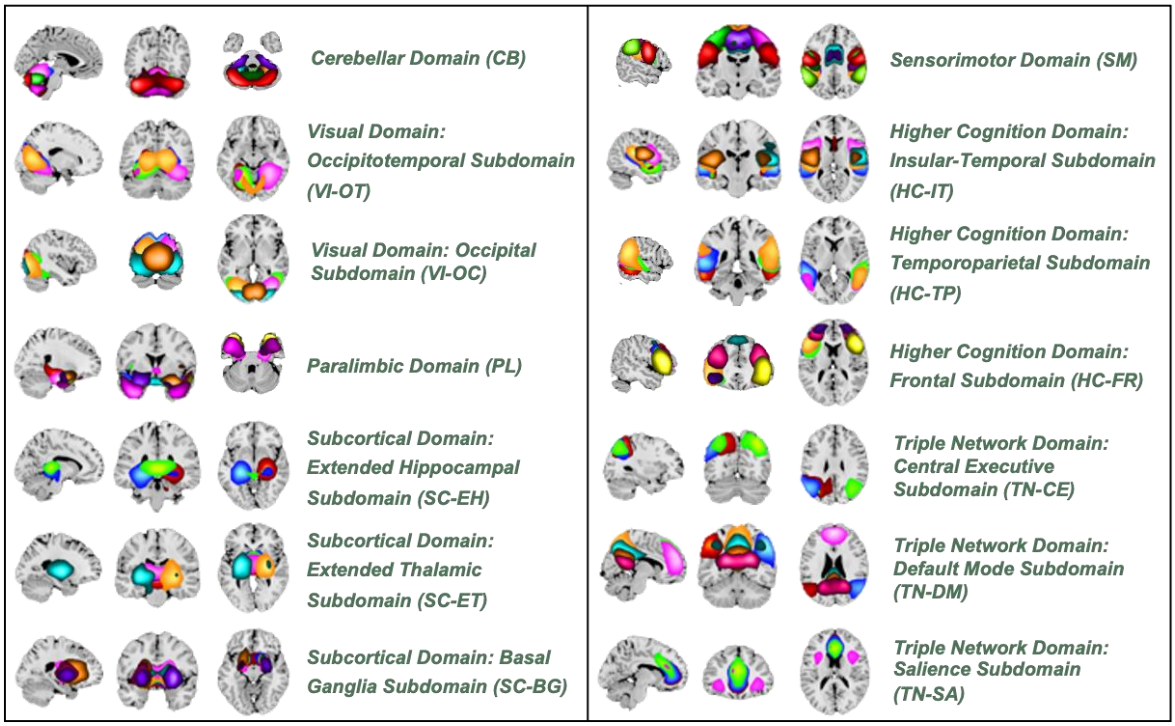

Figure 4 Best Matched 105 ICNs estimated using group-level bICA from the infant dataset. Presented as composite maps of 7 domains and 14 subdomains. Cerebellar Domain, Visual Domain (includes Occipitotemporal and Occipital subdomains), Paralimbic Domain, Subcortical Domain (includes Extended Hippocampal, Extended Thalamic and Basal Ganglia subdomains), Sensorimotor Domain, Higher Cognition Domain (includes Insular Temporal, Temporoparietal and Frontal subdomains) and Triple Network Domain (includes Central Executive, Default Mode, and Salience subdomains).

Figure 5a details the spatial similarity of each ICN with the template. Notably, four ICNs exhibited remarkably high similarity to the template ICNs, with correlation coefficients (rho) exceeding 0.9—three of these ICNs belonged to the sensorimotor (SM) domain and one to the high cognitive-insular temporal (HC-IT) domain. Additionally, 32 out of the 105 ICNs had rho values greater than 0.8, while the remaining 69 ICNs all exceeded 0.5. The correlation coefficients ranged from 0.5456 to 0.9587, with an average of 0.7462 ± 0.0962, all surpassing previously established thresholds (Smith et al., 2009) also mentioned in NeuroMark 1.0 (Du et al., 2020). Figure 5b illustrates the distribution of these correlation values, while Figure 5c displays the functional connectivity for

these ICNs, providing a comprehensive view of their spatial and functional characteristics. The strength of the FNCs ranged from –0.48 to 0.99, reflecting the variability in connectivity patterns.

3.1.2 Null Data Validation

Figure 5d exemplifies a clear comparison of the spatial maps for ICN 64 from the NeuroMark 2.2 template, the infant dataset, and the null dataset obtained using bICA. The spatial map obtained from the infant dataset shows a markedly higher similarity to the template than that from the null dataset, highlighting the distinctiveness of the ICNs in the infant data. The average spatial correlation between the target component from the template and the component obtained using bICA from the null data was mean ± std = 0.0056 ± 1.56e-04, which is significantly lower than the correlation observed in the actual infant data (mean ± std = 0.5710 ± 0.0140), as depicted in Figure 5e.

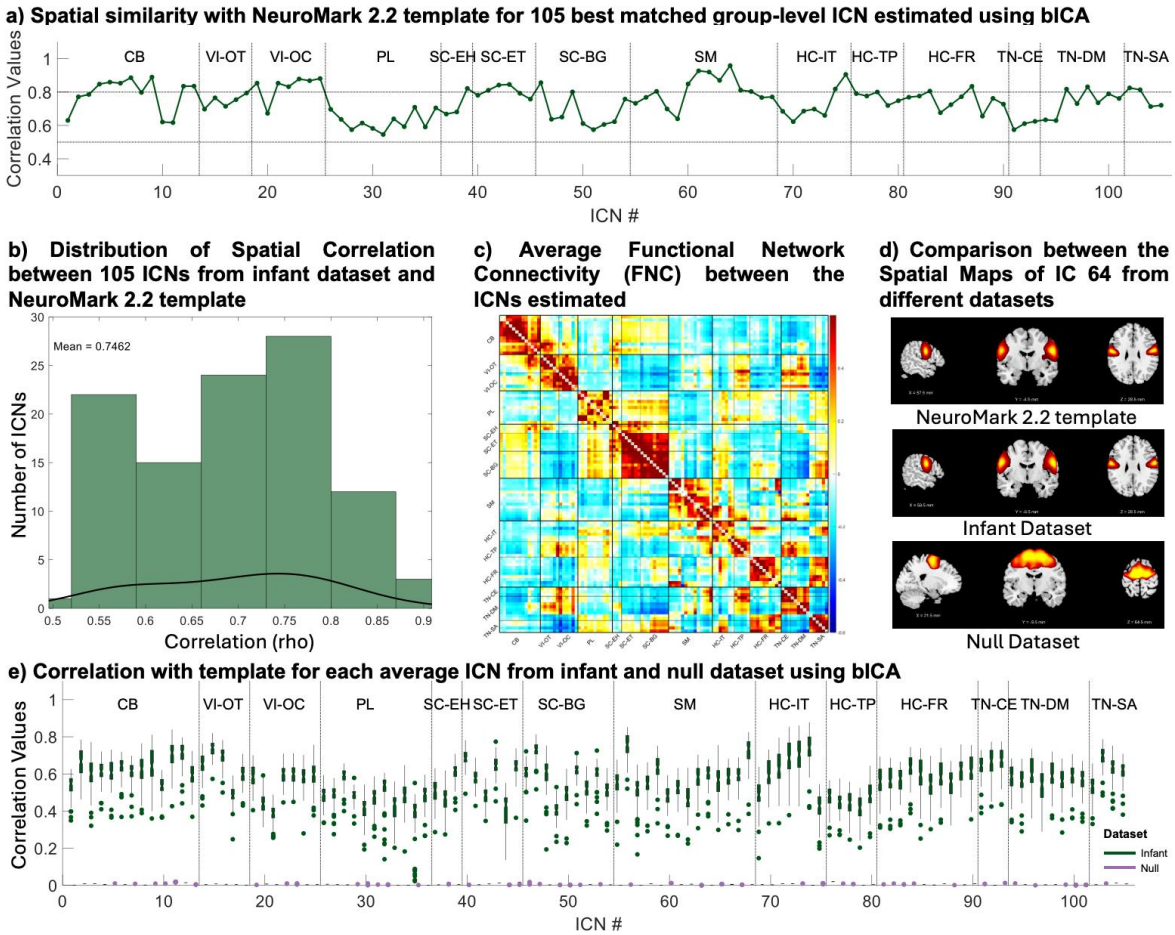

Figure 5 Analysis of 105 ICNs estimated from the Infant Dataset from bICA and matched with the NeuroMark 2.2 Template a) Spatial similarity (rho values) for each of the 105 best-matched group-level ICNs estimated using bICA, compared with the NeuroMark 2.2 template. This panel illustrates the individual correlation values indicating how well each ICN aligns with the template b) Distribution of spatial correlation (rho values) with the template for the 105 ICNs from the infant dataset. This histogram shows the range and frequency of similarity values, providing an overview of the alignment between the estimated ICNs and the template. c) Average Functional Network Connectivity (FNC) between the ICNs derived from the infant dataset. This panel presents the mean connectivity patterns among the identified ICNs, offering insights into their functional relationships. d) Comparison of spatial maps for ICN 64 from different datasets: NeuroMark 2.2 template, the infant dataset, and a null dataset. This example highlights the similarities and differences in ICNs obtained from the actual data versus the null dataset. e) Subject-level ICNs reconstructed using the 105 best-matched ICNs. This panel shows the correlation with the template for subject-level ICNs obtained using bICA from both the infant and null datasets, demonstrating the reliability of the ICNs in the infant dataset.

Furthermore, t-tests comparing the spatial similarity of subject-level spatial maps from the infant and null datasets with the template yielded highly significant results, with t-values ranging from 26.4620 to 124.5382 and p-values less than 3.42E-63. These findings strongly support the validity of the ICNs identified in the infant rsfMRI.

3.2 NeuroMark Framework

3.2.1 Estimation of reliable ICNs

The NeuroMark Framework analysis conducted on the same 126 scans resulted in the generation of subject-level spatial maps for each scan. Figure 6 offers a composite view of the average 105 ICNs spatial maps derived from the subject-specific ICNs generated using NeuroMark Framework from the infant dataset, organized by subdomains. The correlation coefficients between the estimated ICNs and the corresponding networks from the multi-scale NeuroMark 2.2 template demonstrated a high degree of spatial similarity, with all values exceeding rho > 0.83 (mean = 0.9138 ± 0.0236). The highest spatial correlation observed was rho = 0.96, as illustrated in Figure 7a and Figure 7b. These findings underscore the robustness and accuracy of the ICNs identified through MOO-ICAR, reinforcing the alignment between the infant data and the established template. The average FNC between the ICNs estimated using NeuroMark Framework is shown in Figure 7c with their strengths ranging from -0.48 to 0.96.

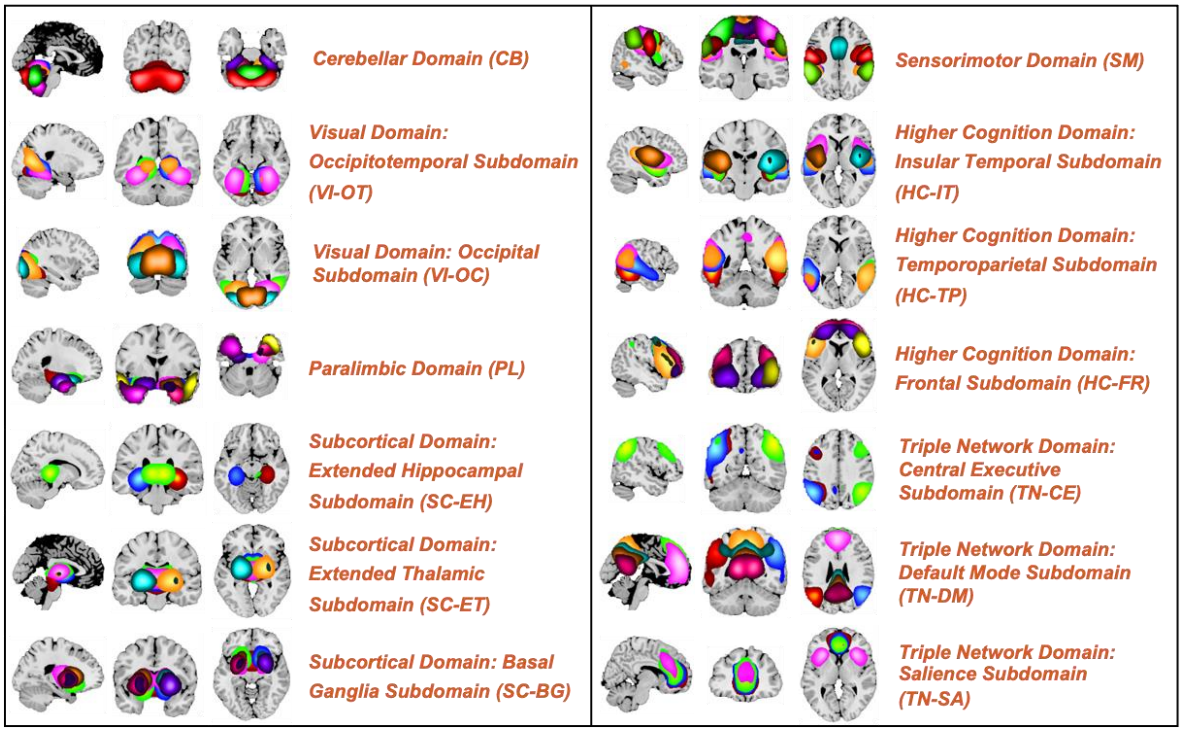

Figure 6 Average 105 ICNs estimated using subject-level MOO-ICAR (NeuroMark Framework) from the infant dataset. Presented as composite maps of 7 domains and 14 subdomains. Cerebellar Domain, Visual Domain (includes Occipitotemporal and Occipital subdomains), Paralimbic Domain, Subcortical Domain (includes Extended Hippocampal, Extended Thalamic and Basal Ganglia subdomains), Sensorimotor Domain, Higher Cognition Domain (includes Insular Temporal, Temporoparietal and Frontal subdomains) and Triple Network Domain (includes Central Executive, Default Mode, and Salience subdomains).

3.2.1 Subject-Level Analysis and Null Data Validation

Figure 7d illustrates the spatial maps for ICN 64 as an example for visual comparison, obtained from the NeuroMark 2.2 template, the infant dataset, and the null dataset using MOO-ICAR. This comparison highlights

the distinct difference between the ICNs derived from the infant data and those from the null dataset, with the infant ICNs showing much greater similarity to the template. Figure 7e further illustrates Pearson's correlation coefficient for all ICNs obtained from the NeuroMark Framework for both the infant and the null dataset obtained on a subject-level. The spatial similarity for the actual infant data (mean ± std = 0.6737 ± 0.0082) is much higher than those estimated from the null dataset (mean ± std = 0.0781 ± 0.0311). At the subject level, t-tests were conducted for each component, comparing the spatial similarity of the spatial maps from the infant and null datasets with the template. The t-test results revealed a maximum t-value of 88.4138, a minimum t-value of 40.5421, and p-values less than 7.11E-91. These findings strongly indicate that the ICNs estimated using the NeuroMark Framework on the infant dataset are significantly different from those derived from the null dataset, thereby confirming the reliability and validity of the ICN estimation process. Notably, we observed networks as early as the first week after birth, with high spatial similarity to the templates (mean ± std: 0.6286 ± 0.0478).

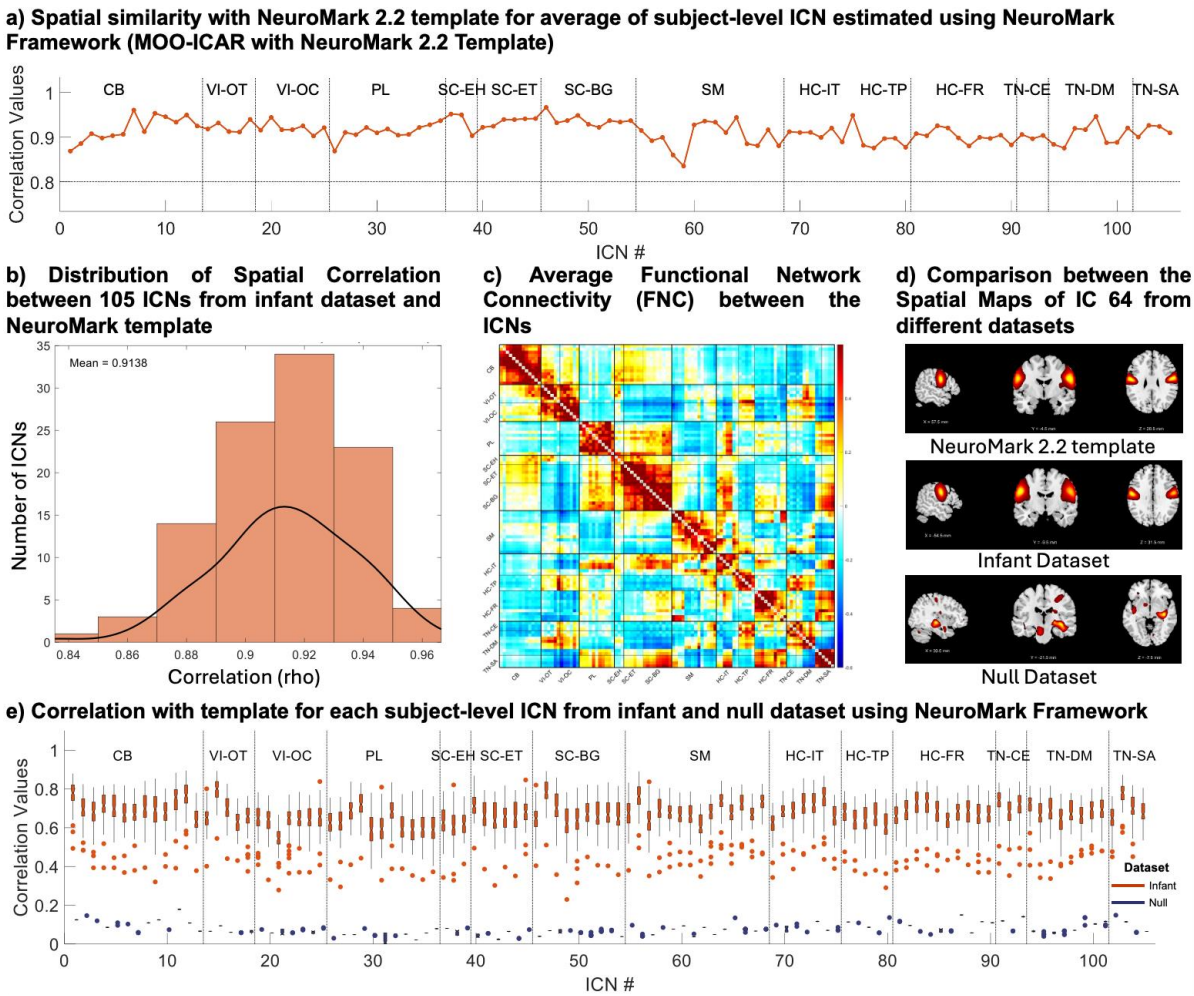

Figure 7 Analysis of 105 ICNs Estimated Using the NeuroMark Framework from the Infant Dataset a) Spatial similarity between the average estimated subject-level ICNs and the NeuroMark 2.2 template. This panel shows the mean correlation values for ICNs across subjects, highlighting their alignment with the template. b) Distribution of spatial correlation between the 105 ICNs from the infant dataset and the template. This histogram illustrates the range and frequency of similarity values, providing an overview of how well the estimated ICNs match the template. c) Average Functional Network Connectivity (FNC) among the ICNs estimated using NeuroMark Framework. This panel presents the mean connectivity patterns between the identified ICNs, offering insights into their functional relationships d) Comparison of spatial maps for ICN 64 across different datasets: NeuroMark 2.2 template, the infant dataset, and the null dataset. e) Correlation with the template for each subject-level ICN obtained using NeuroMark Framework from both the infant and null datasets.

3.2.2 *Network-Specific Information at the Subject-Level in the Infants*

In this section, we examined whether the ICNs could be reliably differentiated at the subject level. For each network of interest from the template, we identified an ICN from the remaining 104 components that showed the highest spatial similarity, which we designated as the comparison ICN. We then compared the network of interest with the comparison ICN for all subjects in the infant dataset. Our analysis revealed that the spatial similarity between the network of interest and the template consistently exceeded that of the comparison ICN, demonstrating that the networks of interest could be reliably distinguished across subjects. Additionally, as illustrated in Figure 8-top, the differences in correlation values between the networks of interest and the comparison ICNs were consistently positive across all ICNs, with an overall mean ± std of 0.2618 ± 0.1288. These results were statistically validated through t-tests, which rejected the null hypothesis, indicating that the mean correlation between networks of interest ICN from the infant dataset and the template significantly surpassed that of the comparison ICN (t-values: 15.0445 to 173.4301, $p < 1.54E-46$, Figure 8-bottom). These findings suggest that while the ICNs may exhibit multi-spatial characteristics, they maintain distinct and unique information from other ICNs obtained from the same subject, with significant differences evident in their correlations.

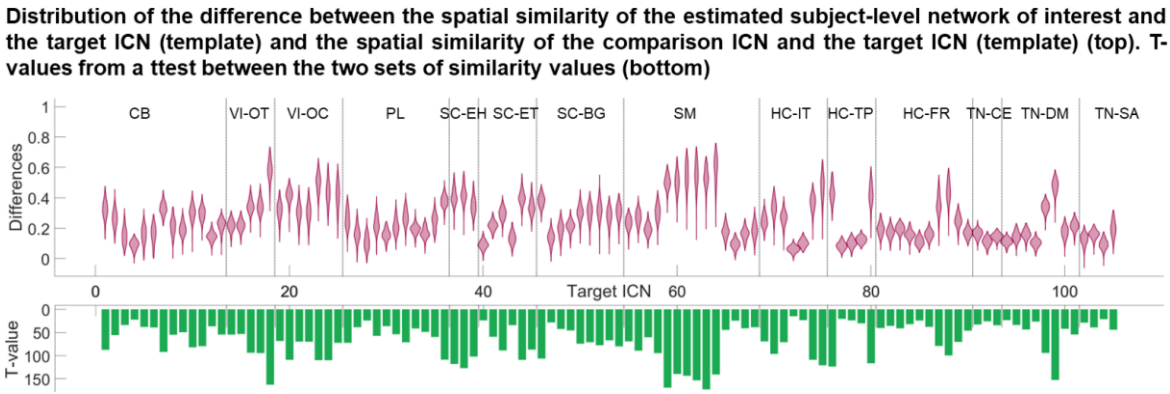

Figure 8 Distribution of the difference in spatial similarity between the estimated subject-level network of interest and the template, compared to the comparison ICN similarity with the template (top). The bottom panel presents t-values from a t-test between these two sets of similarity values, highlighting statistical significance in the differences observed.

4 Discussion

rsfMRI studies in infants remain nascent compared to those conducted in adults. While past studies have reported the presence of multiple ICNs in infants (Damaraju et al., 2014; Edwards et al., 2022; Eyre et al., 2021; França et al., 2024; Gao, Alcauter, Elton, et al., 2015; Gao et al., 2009, 2013; Molloy & Saygin, 2022), only a limited number of networks have been thoroughly examined in infants (Eyre et al., 2021; Gao, Alcauter, Elton, et al., 2015; Liu et al., 2008; Molloy & Saygin, 2022). Moreover, no study to date has provided a comprehensive mapping of infant networks comparable to the detailed maps seen in adult brain research. Here, we fill this gap by identifying 105 canonical multi-scale ICNs in infants using the NeuroMark 2.2 template (Iraji et al., 2023; Jensen, Turner, et al., 2024).

In this study, we introduced and used burst ICA and NeuroMark Framework, the latter integrated with the NeuroMark framework, to estimate these multi-scale ICNs. Our novel bICA approach identified a broad array of networks—25,380 ICNs in total—from which 105 were selected based on their high spatial similarity (correlation > 0.5) with adult ICNs. This threshold surpasses the 0.40 benchmark (Smith et al., 2009) used by NeuroMark 1.0 to estimate 53 meaningful ICNs (Du et al., 2020). The high correlation achieved through an unbiased approach confirms the existence of canonical multi-scale ICNs in infants as young as six months, demonstrating the potential for detailed network studies during this critical developmental window. Using a guided approach, the

NeuroMark Framework with MOO-ICAR, and the NeuroMark 2.2 template, precisely estimated 105 ICNs with a correlation > 0.8, consistent with the threshold used during the generation of the NeuroMark 2.2 template. This strong alignment of infant ICNs with a template derived primarily from older participants, ranging from adolescents to the elderly, underscores the early establishment of brain networks. The ICs obtained from the null dataset did not resemble the ICNs in the template, further affirming reliability. For the ICN estimations from the NeuroMark Framework, the within-between ICN analysis provided further insights into the consistency and specificity of ICNs. The mean correlation between the within-component and the template consistently surpassed that between-component correlation and the template. This observation confirms the distinctiveness of each estimated ICN within the subject.

Our findings offer critical insights into neurodevelopment, particularly the early formation of brain network architecture. The strong spatial correlations observed between infant ICNs, and the NeuroMark multi-scale template suggest that the structural basis of these networks is detectable as early as the first week of life. While prior research demonstrated complex networks, such as the DMN emerge in a nascent form and mature over the first year with the salience network developing by the second year (Gao, Alcauter, Smith, et al., 2015; Gao et al., 2009), our results indicate that ICNs—including the DMN and salience network—in infants aged 0 to 6 months already exhibit a high degree of similarity to their mature counterparts. The multi-scale ICNs we identified showed that DMN-related networks within the TN-DM subdomain have a high spatial similarity to the template (mean rho = 0.9047 ± 0.0245). Similarly, the TN-SN subdomain related to the salience network showed substantial similarity (mean rho = 0.9149 ± 0.0123). Both values were within a comparable range to primary networks such as the visual network, which were expected to demonstrate adult-like structure in neonates (Gao, Alcauter, Smith, et al., 2015; Gao et al., 2009) — VI-OT (mean rho = 0.9226 ± 0.0124) and VI-OC (mean rho = 0.9202 ± 0.0126). Additionally, the subcortical network, which had not been studied in infants until now, demonstrated the highest similarity among subdomains, with SC-BG (mean rho = 0.9380 ± 0.0128), SC-EH (mean rho = 0.9348 ± 0.0274), and SC-ET (mean rho = 0.9346 ± 0.0089). Furthermore, the subject-level correlation with the template shows significant variability in correlation values across individual ICNs (Figure 5e, Figure 7e), indicating dramatic network development while maintaining a high average correlation across subjects considering that each subject is of a different age. Furthermore, the comparable average correlation observed in other datasets suggests that by six months of age, most networks have reached a near-matured state. This variability highlights the need for further studies to examine how ICNs evolve with age, which will be a focus of future work.

In addition, our proposed methods address the challenge of selecting a single optimal model order by enabling the estimation of a wide range of networks. Although this study focused on networks closely aligned with the NeuroMark 2.2 template, it sets the stage for future research to investigate networks that may be unique to infants, which is beyond the scope of this paper. We also established the feasibility of applying the NeuroMark Framework and the generalizability of the NeuroMark 2.2 template to infant data. However, there are limitations to our approach. bICA is computationally intensive, making it less suitable for large datasets. While the NeuroMark Framework is better suited for studies where a reliable template is available, bICA remains valuable for identifying new components that may be specific to certain datasets, ages, or cohorts. The ability to identify and validate ICNs in infants has important implications for the early detection of developmental disorders. The methods demonstrated in this study could be used to monitor brain development in infants at increased likelihood for conditions such as autism spectrum disorder or ADHD, potentially leading to earlier and more targeted interventions.

A notable finding in this study is the distinct patterns observed in the FNCs of infants, with connectivity strengths displaying a wide range across the ICNs. Both methods produced highly consistent FNCs, reinforcing the presence of robust brain activity. Comparing these infant FNCs with those from a large-scale UKB dataset

(Jensen, Turner, et al., 2024) comprising 39,342 participants, we observed a high correlation of 0.7166. This indicates a strong similarity between the datasets, showing comparable modular patterns. However, despite this similarity, distinct differences in connectivity patterns are evident, underscoring developmental distinctions in early brain network organization. These findings encourage further in-depth investigation of infant brain networks, a task beyond the scope of this study but an exciting direction for future research. Future studies should aim to replicate these findings in larger and more diverse samples. Researchers have developed other templates like NeuroMark 3.0 (https://trendscenter.org/data/) including multiple modalities and templates for different age groups (Fu et al., 2024) that can be used for future research. Longitudinal studies tracking brain network development from infancy through early childhood would provide deeper insights into the dynamics of neurodevelopment. Additionally, exploring the application of these methods in clinical populations could aid in early diagnosis and intervention strategies. Further research on functional network connectivity, particularly the interaction and dynamic relationships of these networks, would be particularly valuable.

5 Conclusion

In conclusion, this study significantly advances infant neuroimaging by identifying 105 unique, multi-scale ICNs in infants aged 0 to 6 months. Using a novel bICA method, we were able to detect robust, adult-like networks without relying on prior assumptions or templates, addressing the challenge of model order selection by generating a wide range of ICNs intrinsic to the infant dataset itself. This unbiased approach revealed the early foundational existence of networks, including the default mode, visual, and salience, suggesting that the fundamental architecture of brain networks may have been established earlier than previously thought. The NeuroMark Framework, with its NeuroMark 2.2 template derived from over 100,000 subjects, provided further insights into these results by estimating precise ICNs. Together, bICA and NeuroMark provided complementary insights underscoring the importance of studying ICNs across multiple scales. The results from this study provide a detailed understanding of early brain development and lay the groundwork for future research into the interactions and dynamic relationships between brain networks. This has significant implications for developing early diagnostic tools and interventions for developmental disorders, ultimately promoting healthy brain development from early life.

6 Acknowledgement

This work was supported by the National Institute of Mental Health, USA (R01MH118285), and the National Institute of Biomedical Imaging and Bioengineering, USA (R01EB027147).

7 References

Abou-Elseoud, A., Starck, T., Remes, J., Nikkinen, J., Tervonen, O., & Kiviniemi, V. (2010). The effect of model order selection in group PICA. *Human Brain Mapping*, *31*(8), 1207–1216. https://doi.org/10.1002/hbm.20929

Akaike, H. (1998). Information Theory and an Extension of the Maximum Likelihood Principle. In E. Parzen, K. Tanabe, & G. Kitagawa (Eds.), *Selected Papers of Hirotugu Akaike* (pp. 199–213). Springer New York. https://doi.org/10.1007/978-1-4612-1694-0_15

Allen, E. A., Erhardt, E. B., Wei, Y., Eichele, T., & Calhoun, V. D. (2012). Capturing inter-subject variability with group independent component analysis of fMRI data: A simulation study. *NeuroImage*, *59*(4), 4141–4159. https://doi.org/10.1016/j.neuroimage.2011.10.010

Andrés-Camazón, P., Diaz-Caneja, C. M., Ballem, R., Chen, J., Calhoun, V. D., & Iraji, A. (2024). Neurobiology-based Cognitive Biotypes Using Multi-scale Intrinsic Connectivity Networks in Psychotic Disorders. *MedRxiv*, 2024.05.14.24307341. https://doi.org/10.1101/2024.05.14.24307341

Badhwar, A., Tam, A., Dansereau, C., Orban, P., Hoffstaedter, F., & Bellec, P. (2017). Resting-state network dysfunction in Alzheimer's disease: A systematic review and meta-analysis. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, *8*, 73–85. https://doi.org/10.1016/j.dadm.2017.03.007

Beckmann, C. F. (2012). Modelling with independent components. *NeuroImage*, *62*(2), 891–901. https://doi.org/10.1016/j.neuroimage.2012.02.020

Beckmann, C. F., & Smith, S. M. (2004). Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Transactions on Medical Imaging*, *23*(2), 137–152. IEEE Transactions on Medical Imaging. https://doi.org/10.1109/TMI.2003.822821

Bethlehem, R. A. I., Seidlitz, J., White, S. R., Vogel, J. W., Anderson, K. M., Adamson, C., Adler, S., Alexopoulos, G. S., Anagnostou, E., Areces-Gonzalez, A., Astle, D. E., Auyeung, B., Ayub, M., Bae, J., Ball, G., Baron-Cohen, S., Beare, R., Bedford, S. A., Benegal, V., … Alexander-Bloch, A. F. (2022). Brain charts for the human lifespan. *Nature*, *604*(7906), 525–533. https://doi.org/10.1038/s41586-022-04554-y

Bi, X., Zhao, J., Xu, Q., Sun, Q., & Wang, Z. (2018). Abnormal Functional Connectivity of Resting State Network Detection Based on Linear ICA Analysis in Autism Spectrum Disorder. *Frontiers in Physiology*, *9*. https://doi.org/10.3389/fphys.2018.00475

Biswal, B. B. (2012). Resting state fMRI: A personal history. *NeuroImage*, *62*(2), 938–944. https://doi.org/10.1016/j.neuroimage.2012.01.090

Biswal, B., Zerrin Yetkin, F., Haughton, V. M., & Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar mri. *Magnetic Resonance in Medicine*, *34*(4), 537–541. https://doi.org/10.1002/mrm.1910340409

Calhoun, V. D., & Adali, T. (2006). Unmixing fMRI with independent component analysis. *IEEE Engineering in Medicine and Biology Magazine*, *25*(2), 79–90. IEEE Engineering in Medicine and Biology Magazine. https://doi.org/10.1109/MEMB.2006.1607672

Calhoun, V. D., Adali, T., Pearlson, G. D., & Pekar, J. J. (2001). A method for making group inferences from functional MRI data using independent component analysis. *Human Brain Mapping*, *14*(3), 140–151. https://doi.org/10.1002/hbm.1048

Calhoun, V. D., & Adalı, T. (2012). Multi-subject Independent Component Analysis of fMRI: A Decade of Intrinsic Networks, Default Mode, and Neurodiagnostic Discovery. *IEEE Reviews in Biomedical Engineering*, *5*, 60–73. https://doi.org/10.1109/RBME.2012.2211076

Calhoun, V. D., & de Lacy, N. (2017). Ten Key Observations on the Analysis of Resting-state Functional MR Imaging Data Using Independent Component Analysis. *Neuroimaging Clinics of North America*, *27*(4), 561–579. https://doi.org/10.1016/j.nic.2017.06.012

Calhoun, V. D., Liu, J., & Adalı, T. (2009). A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data. *NeuroImage*, *45*(1 Suppl), S163–S172. https://doi.org/10.1016/j.neuroimage.2008.10.057

Chang, Y.-T., Huang, C.-W., Chang, W.-N., Lee, J.-J., & Chang, C.-C. (2018). Altered Functional Network Affects Amyloid and Structural Covariance in Alzheimer's Disease. *BioMed Research International*, *2018*, 8565620. https://doi.org/10.1155/2018/8565620

Damaraju, E., Caprihan, A., Lowe, J. R., Allen, E. A., Calhoun, V. D., & Phillips, J. P. (2014). Functional connectivity in the developing brain: A longitudinal study from 4 to 9months of age. *NeuroImage*, *84*, 169–180. https://doi.org/10.1016/j.neuroimage.2013.08.038

Damoiseaux, J. S., Rombouts, S. A. R. B., Barkhof, F., Scheltens, P., Stam, C. J., Smith, S. M., & Beckmann, C. F. (2006). Consistent resting-state networks across healthy subjects. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(37), 13848–13853. https://doi.org/10.1073/pnas.0601417103

De Luca, M., Beckmann, C. F., De Stefano, N., Matthews, P. M., & Smith, S. M. (2006). FMRI resting state networks define distinct modes of long-distance interactions in the human brain. *NeuroImage*, *29*(4), 1359–1367. https://doi.org/10.1016/j.neuroimage.2005.08.035

Du, Y., & Fan, Y. (2013). Group information guided ICA for fMRI data analysis. *NeuroImage*, *69*, 157–197. https://doi.org/10.1016/j.neuroimage.2012.11.008

Du, Y., Fu, Z., Sui, J., Gao, S., Xing, Y., Lin, D., Salman, M., Abrol, A., Rahaman, M. A., Chen, J., Hong, L. E., Kochunov, P., Osuch, E. A., & Calhoun, V. D. (2020). NeuroMark: An automated and adaptive ICA

based pipeline to identify reproducible fMRI markers of brain disorders. *NeuroImage: Clinical*, *28*, 102375. https://doi.org/10.1016/j.nicl.2020.102375

Edwards, A. D., Rueckert, D., Smith, S. M., Abo Seada, S., Alansary, A., Almalbis, J., Allsop, J., Andersson, J., Arichi, T., Arulkumaran, S., Bastiani, M., Batalle, D., Baxter, L., Bozek, J., Braithwaite, E., Brandon, J., Carney, O., Chew, A., Christiaens, D., … Hajnal, J. V. (2022). The Developing Human Connectome Project Neonatal Data Release. *Frontiers in Neuroscience*, *16*. https://doi.org/10.3389/fnins.2022.886772

Emerson, R. W., Gao, W., & Lin, W. (2016). Longitudinal Study of the Emerging Functional Connectivity Asymmetry of Primary Language Regions during Infancy. *Journal of Neuroscience*, *36*(42), 10883–10892. https://doi.org/10.1523/JNEUROSCI.3980-15.2016

Eyre, M., Fitzgibbon, S. P., Ciarrusta, J., Cordero-Grande, L., Price, A. N., Poppe, T., Schuh, A., Hughes, E., O'Keeffe, C., Brandon, J., Cromb, D., Vecchiato, K., Andersson, J., Duff, E. P., Counsell, S. J., Smith, S. M., Rueckert, D., Hajnal, J. V., Arichi, T., … Edwards, A. D. (2021). The Developing Human Connectome Project: Typical and disrupted perinatal functional connectivity. *Brain*, *144*(7), 2199–2213. https://doi.org/10.1093/brain/awab118

França, L. G. S., Ciarrusta, J., Gale-Grant, O., Fenn-Moltu, S., Fitzgibbon, S., Chew, A., Falconer, S., Dimitrova, R., Cordero-Grande, L., Price, A. N., Hughes, E., O'Muircheartaigh, J., Duff, E., Tuulari, J. J., Deco, G., Counsell, S. J., Hajnal, J. V., Nosarti, C., Arichi, T., … Batalle, D. (2024). Neonatal brain dynamic functional connectivity in term and preterm infants and its association with early childhood neurodevelopment. *Nature Communications*, *15*(1), 16. https://doi.org/10.1038/s41467-023-44050-z

Fransson, P. (2006). How default is the default mode of brain function? Further evidence from intrinsic BOLD signal fluctuations. *Neuropsychologia*, *44*(14), 2836–2845. https://doi.org/10.1016/j.neuropsychologia.2006.06.017

Fransson, P., Skiöld, B., Horsch, S., Nordell, A., Blennow, M., Lagercrantz, H., & Åden, U. (2007). Resting-state networks in the infant brain. *Proceedings of the National Academy of Sciences*, *104*(39), 15531–15536. https://doi.org/10.1073/pnas.0704380104

Fu, Z., Batta, I., Wu, L., Abrol, A., Agcaoglu, O., Salman, M. S., Du, Y., Iraji, A., Shultz, S., Sui, J., & Calhoun, V. D. (2024). Searching Reproducible Brain Features using NeuroMark: Templates for Different Age Populations and Imaging Modalities. *NeuroImage*, *292*, 120617. https://doi.org/10.1016/j.neuroimage.2024.120617

Fu, Z., Iraji, A., Caprihan, A., Adair, J. C., Sui, J., Rosenberg, G. A., & Calhoun, V. D. (2020). In search of multimodal brain alterations in Alzheimer's and Binswanger's disease. *NeuroImage: Clinical*, *26*, 101937. https://doi.org/10.1016/j.nicl.2019.101937

Gao, W., Alcauter, S., Elton, A., Hernandez-Castillo, C. R., Smith, J. K., Ramirez, J., & Lin, W. (2015). Functional Network Development During the First Year: Relative Sequence and Socioeconomic Correlations. *Cerebral Cortex*, *25*(9), 2919–2928. https://doi.org/10.1093/cercor/bhu088

Gao, W., Alcauter, S., Smith, J. K., Gilmore, J. H., & Lin, W. (2015). Development of human brain cortical network architecture during infancy. *Brain Structure and Function*, *220*(2), 1173–1186. https://doi.org/10.1007/s00429-014-0710-3

Gao, W., Gilmore, J. H., Shen, D., Smith, J. K., Zhu, H., & Lin, W. (2013). The synchronization within and interaction between the default and dorsal attention networks in early infancy. *Cerebral Cortex (New York, N.Y.: 1991)*, *23*(3), 594–603. https://doi.org/10.1093/cercor/bhs043

Gao, W., Zhu, H., Giovanello, K. S., Smith, J. K., Shen, D., Gilmore, J. H., & Lin, W. (2009). Evidence on the emergence of the brain's default network from 2-week-old to 2-year-old healthy pediatric subjects. *Proceedings of the National Academy of Sciences*, *106*(16), 6790–6795. https://doi.org/10.1073/pnas.0811221106

Howell, B. R., Styner, M. A., Gao, W., Yap, P.-T., Wang, L., Baluyot, K., Yacoub, E., Chen, G., Potts, T., Salzwedel, A., Li, G., Gilmore, J. H., Piven, J., Smith, J. K., Shen, D., Ugurbil, K., Zhu, H., Lin, W., & Elison, J. T. (2019). The UNC/UMN Baby Connectome Project (BCP): An overview of the study design and protocol development. *NeuroImage*, *185*, 891–905. https://doi.org/10.1016/j.neuroimage.2018.03.049

Hunter, M. D., Eickhoff, S. B., Miller, T. W. R., Farrow, T. F. D., Wilkinson, I. D., & Woodruff, P. W. R. (2006). Neural activity in speech-sensitive auditory cortex during silence. *Proceedings of the National Academy of Sciences*, *103*(1), 189–194. https://doi.org/10.1073/pnas.0506268103

Huttenlocher, P. R., & Dabholkar, A. S. (1997). Regional differences in synaptogenesis in human cerebral cortex. *Journal of Comparative Neurology*, *387*(2), 167–178. https://doi.org/10.1002/(SICI)1096-9861(19971020)387:2<167::AID-CNE1>3.0.CO;2-Z

Iraji, A., Benson, R. R., Welch, R. D., O'Neil, B. J., Woodard, J. L., Ayaz, S. I., Kulek, A., Mika, V., Medado, P., Soltanian-Zadeh, H., Liu, T., Haacke, E. M., & Kou, Z. (2015). Resting State Functional Connectivity

in Mild Traumatic Brain Injury at the Acute Stage: Independent Component and Seed-Based Analyses. *Journal of Neurotrauma*, *32*(14), 1031–1045. https://doi.org/10.1089/neu.2014.3610

Iraji, A., Faghiri, A., Fu, Z., Rachakonda, S., Kochunov, P., Belger, A., Ford, J. M., McEwen, S., Mathalon, D. H., Mueller, B. A., Pearlson, G. D., Potkin, S. G., Preda, A., Turner, J. A., Van Erp, T. G. M., & Calhoun, V. D. (2022). Multi-spatial-scale dynamic interactions between functional sources reveal sex-specific changes in schizophrenia. *Network Neuroscience*, *6*(2), 357–381. https://doi.org/10.1162/netn_a_00196

Iraji, A., Faghiri, A., Lewis, N., Fu, Z., DeRamus, T., Qi, S., Rachakonda, S., Du, Y., & Calhoun, V. D. (2019). Ultra-high-order ICA: An exploration of highly resolved data-driven representation of intrinsic connectivity networks (sparse ICNs). In Y. M. Lu, M. Papadakis, & D. Van De Ville (Eds.), *Wavelets and Sparsity XVIII* (p. 18). SPIE. https://doi.org/10.1117/12.2530106

Iraji, A., Faghiri, A., Lewis, N., Fu, Z., Rachakonda, S., & Calhoun, V. D. (2021). Tools of the trade: Estimating time-varying connectivity patterns from fMRI data. *Social Cognitive and Affective Neuroscience*, *16*(8), 849–874. https://doi.org/10.1093/scan/nsaa114

Iraji, A., Fu, Z., Faghiri, A., Duda, M., Chen, J., Rachakonda, S., DeRamus, T., Kochunov, P., Adhikari, B. M., Belger, A., Ford, J. M., Mathalon, D. H., Pearlson, G. D., Potkin, S. G., Preda, A., Turner, J. A., van Erp, T. G. M., Bustillo, J. R., Yang, K., … Calhoun, V. D. (2023). Identifying canonical and replicable multi-scale intrinsic connectivity networks in 100k+ resting-state fMRI datasets. *Human Brain Mapping*, *44*(17), 5729–5748. https://doi.org/10.1002/hbm.26472

Jensen, K. M., Calhoun, V. D., Fu, Z., Yang, K., Faria, A. V., Ishizuka, K., Sawa, A., Andrés-Camazón, P., Coffman, B. A., Seebold, D., Turner, J. A., Salisbury, D. F., & Iraji, A. (2024). A whole-brain neuromark resting-state fMRI analysis of first-episode and early psychosis: Evidence of aberrant cortical-subcortical-cerebellar functional circuitry. *NeuroImage: Clinical*, *41*, 103584. https://doi.org/10.1016/j.nicl.2024.103584

Jensen, K. M., Turner, J. A., Calhoun, V. D., & Iraji, A. (2024). *Addressing inconsistency in functional neuroimaging: A replicable data-driven multi-scale functional atlas for canonical brain networks* (p. 2024.09.09.612129). bioRxiv. https://doi.org/10.1101/2024.09.09.612129

Knickmeyer, R. C., Gouttard, S., Kang, C., Evans, D., Wilber, K., Smith, J. K., Hamer, R. M., Lin, W., Gerig, G., & Gilmore, J. H. (2008). A Structural MRI Study of Human Brain Development from Birth to 2 Years. *The Journal of Neuroscience*, *28*(47), 12176–12182. https://doi.org/10.1523/JNEUROSCI.3479-08.2008

Laird, A. R., Fox, P. M., Eickhoff, S. B., Turner, J. A., Ray, K. L., McKay, D. R., Glahn, D. C., Beckmann, C. F., Smith, S. M., & Fox, P. T. (2011). Behavioral Interpretations of Intrinsic Connectivity Networks. *Journal of Cognitive Neuroscience*, *23*(12), 4022–4037. https://doi.org/10.1162/jocn_a_00077

Li, H., Zhu, X., & Fan, Y. (2018). Identification of Multi-scale Hierarchical Brain Functional Networks Using Deep Matrix Factorization. In A. F. Frangi, J. A. Schnabel, C. Davatzikos, C. Alberola-López, & G. Fichtinger (Eds.), *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018* (pp. 223–231). Springer International Publishing. https://doi.org/10.1007/978-3-030-00931-1_26

Li, Y.-O., Adalı, T., & Calhoun, V. D. (2007). Estimating the number of independent components for functional magnetic resonance imaging data. *Human Brain Mapping*, *28*(11), 1251–1266. https://doi.org/10.1002/hbm.20359

Lin, Q., Liu, J., Zheng, Y., Liang, H., & Calhoun, V. D. (2010). Semiblind spatial ICA of fMRI using spatial constraints. *Human Brain Mapping*, *31*(7), 1076–1088. https://doi.org/10.1002/hbm.20919

Liu, W.-C., Flax, J. F., Guise, K. G., Sukul, V., & Benasich, A. A. (2008). Functional connectivity of the sensorimotor area in naturally sleeping infants. *Brain Research*, *1223*, 42–49. https://doi.org/10.1016/j.brainres.2008.05.054

Lu, W., & Rajapakse, J. C. (2006). ICA with Reference. *Neurocomputing*, *69*(16), 2244–2257. https://doi.org/10.1016/j.neucom.2005.06.021

McKeown, M. J., Hansen, L. K., & Sejnowsk, T. J. (2003). Independent component analysis of functional MRI: What is signal and what is noise? *Current Opinion in Neurobiology*, *13*(5), 620–629. https://doi.org/10.1016/j.conb.2003.09.012

Mckeown, M. J., Makeig, S., Brown, G. G., Jung, T.-P., Kindermann, S. S., Bell, A. J., & Sejnowski, T. J. (1998). Analysis of fMRI data by blind separation into independent spatial components. *Human Brain Mapping*, *6*(3), 160–188. https://doi.org/10.1002/(SICI)1097-0193(1998)6:3<160::AID-HBM5>3.0.CO;2-1

Meng, X., Iraji, A., Fu, Z., Kochunov, P., Belger, A., Ford, J. M., McEwen, S., Mathalon, D. H., Mueller, B. A., Pearlson, G., Potkin, S. G., Preda, A., Turner, J., Van Erp, T. G. M., Sui, J., & Calhoun, V. D. (2023). Multi-model order spatially constrained ICA reveals highly replicable group differences and consistent predictive results from resting data: A large N fMRI schizophrenia study. *NeuroImage: Clinical*, *38*, 103434. https://doi.org/10.1016/j.nicl.2023.103434

Meng, X., Iraji, A., Fu, Z., Kochunov, P., Belger, A., Ford, J., McEwen, S., Mathalon, D. H., Mueller, B. A., Pearlson, G., Potkin, S. G., Preda, A., Turner, J., van Erp, T., Sui, J., & Calhoun, V. D. (2022).

Multimodel Order Independent Component Analysis: A Data-Driven Method for Evaluating Brain Functional Network Connectivity Within and Between Multiple Spatial Scales. *Brain Connectivity*, *12*(7), 617–628. https://doi.org/10.1089/brain.2021.0079

Molloy, M. F., & Saygin, Z. M. (2022). Individual variability in functional organization of the neonatal brain. *NeuroImage*, *253*, 119101. https://doi.org/10.1016/j.neuroimage.2022.119101

Neha, & Gandhi, T. K. (2016). Resting state fMRI analysis using seed based and ICA methods. *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*, 2551–2554. https://ieeexplore.ieee.org/document/7724722

Nir, Y., Hasson, U., Levy, I., Yeshurun, Y., & Malach, R. (2006). Widespread functional connectivity and fMRI fluctuations in human visual cortex in the absence of visual stimulation. *NeuroImage*, *30*(4), 1313–1324. https://doi.org/10.1016/j.neuroimage.2005.11.018

Seraji, M., Shultz, S., Li, Q., Fu, Z., Calhoun, V. D., & Iraji, A. (2024). *Spatial Development of Brain Networks During The First Six Postnatal Months* (p. 2024.09.19.613960). bioRxiv. https://doi.org/10.1101/2024.09.19.613960

Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., Filippini, N., Watkins, K. E., Toro, R., Laird, A. R., & Beckmann, C. F. (2009). Correspondence of the brain's functional architecture during activation and rest. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(31), 13040–13045. https://doi.org/10.1073/pnas.0905267106

Suzuki, K., Kiryu, T., & Nakada, T. (2001). Fast and precise independent component analysis for high field fMRI time series tailored using prior information on spatiotemporal structure. *Human Brain Mapping*, *15*(1), 54–66. https://doi.org/10.1002/hbm.1061

Thomas Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zöllei, L., Polimeni, J. R., Fischl, B., Liu, H., & Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, *106*(3), 1125–1165. https://doi.org/10.1152/jn.00338.2011

Uddin, L. Q., Supekar, K., & Menon, V. (2013). Reconceptualizing functional brain connectivity in autism from a developmental perspective. *Frontiers in Human Neuroscience*, *7*, 458. https://doi.org/10.3389/fnhum.2013.00458

Uddin, L. Q., Yeo, B. T. T., & Spreng, R. N. (2019). Towards a Universal Taxonomy of Macro-scale Functional Human Brain Networks. *Brain Topography*, *32*(6), 926–942. https://doi.org/10.1007/s10548-019-00744-6