

UC Irvine

UC Irvine Previously Published Works

Title

Segmenting hypothalamic subunits in human newborn magnetic resonance imaging data.

Permalink

<https://escholarship.org/uc/item/1h20702x>

Journal

Human Brain Mapping, 45(2)

Authors

Wang, Yun

Graham, Alice

Fair, Damien

et al.

Publication Date

2024-02-01


DOI

10.1002/hbm.26582

Peer reviewed

RESEARCH ARTICLE

Segmenting hypothalamic subunits in human newborn magnetic resonance imaging data

Jerod M. Rasmussen^{1,2}  | Yun Wang^{3,4} | Alice M. Graham⁵ | Damien A. Fair⁶ | Jonathan Posner^{3,4} | Thomas G. O'Connor⁷ | Hyagriv N. Simhan⁸ | Elizabeth Yen⁹ | Neel Madan¹⁰ | Sonja Entringer^{1,2,11} | Pathik D. Wadhwa^{1,2,12,13,14} | Claudia Buss^{1,2,11} | on behalf of program collaborators for Environmental influences on Child Health Outcomes

¹Development, Health and Disease Research Program, University of California, Irvine, California, USA

²Department of Pediatrics, University of California, Irvine, California, USA

³Department of Psychiatry and Behavioral Sciences, Duke University, Durham, North Carolina, USA

⁴New York State Psychiatric Institute, New York, New York, USA

⁵Department of Behavioral Neuroscience, Oregon Health & Science University, Portland, Oregon, USA

⁶Masonic Institute for the Developing Brain, University of Minnesota, Minneapolis, Minnesota, USA

⁷Departments of Psychiatry, Psychology, Neuroscience and Obstetrics and Gynecology, University of Rochester Medical Center, Rochester, New York, USA

⁸Department of Obstetrics and Gynecology, University of Pittsburgh, Pittsburgh, Pennsylvania, USA

⁹Department of Pediatrics, Tufts Medical Center, Boston, Massachusetts, USA

¹⁰Department of Radiology, Tufts Medical Center, Boston, Massachusetts, USA

¹¹Department of Medical Psychology, Charité – Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany

¹²Department of Psychiatry and Human Behavior, University of California, Irvine, California, USA

¹³Department of Obstetrics and Gynecology, University of California, Irvine, California, USA

¹⁴Department of Epidemiology, University of California, Irvine, California, USA

Correspondence

Jerod M. Rasmussen, UC Irvine Development, Health and Disease Research Program, University of California, Irvine, School of Medicine, 3117 Gillespie Neuroscience Research Facility (GNRF), 837 Health Sciences Road, Irvine, CA 92697, USA.
Email: rasmussj@hs.uci.edu

Funding information

National Institute of Mental Health, Grant/Award Numbers: MH091351, MH119510, MH121070; National Institute of Diabetes and Digestive and Kidney Diseases,

Abstract

Preclinical evidence suggests that inter-individual variation in the structure of the hypothalamus at birth is associated with variation in the intrauterine environment, with downstream implications for future disease susceptibility. However, scientific advancement in humans is limited by a lack of validated methods for the automatic segmentation of the newborn hypothalamus. $N = 215$ healthy full-term infants with paired T1-/T2-weighted MR images across four sites were considered for primary analyses (mean postmenstrual age = 44.3 ± 3.5 weeks, $n_{\text{male}}/n_{\text{female}} = 110/106$). The outputs of FreeSurfer's hypothalamic subunit segmentation tools designed for adults

Abbreviations: ANTs, advanced normalization tools; BCP, baby connectome project; DSC, dice similarity coefficient; HTH, hypothalamus; ICV, intracranial volume; MNI, Montreal Neurological Institute; ROI, region of interest; SD, standard deviation; segATLAS, atlas-based segmentation; segFS, FreeSurfer segmentation; segMAN, manual-edit segmentation; segMAS, multi-atlas segmentation; TE, echo time; TR, repetition time.

See Acknowledgments for full listing of collaborators.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2024 The Authors. *Human Brain Mapping* published by Wiley Periodicals LLC.

Grant/Award Number: DK118578; National Institutes of Health, Grant/Award Numbers: U24OD023319, U24OD023382, U2COD023375, UG3OD023349, UH3OD023328; Eunice Kennedy Shriver National Institute of Child Health and Human Development, Grant/Award Numbers: HD060628, HD100593, HD103912, HD0952535

(segFS) were compared against those of a novel registration-based pipeline developed here (segATLAS) and against manually edited segmentations (segMAN) as reference. Comparisons were made using Dice Similarity Coefficients (DSCs) and through expected associations with postmenstrual age at scan. In addition, we aimed to demonstrate the validity of the segATLAS pipeline by testing for the stability of inter-individual variation in hypothalamic volume across the first year of life ($n = 41$ longitudinal datasets available). SegFS and segATLAS segmentations demonstrated a wide spread in agreement (mean DSC = 0.65 ± 0.14 SD; range = {0.03–0.80}). SegATLAS volumes were more highly correlated with postmenstrual age at scan than segFS volumes ($n = 215$ infants; $R_{\text{segATLAS}}^2 = 65\%$ vs. $R_{\text{segFS}}^2 = 40\%$), and segATLAS volumes demonstrated a higher degree of agreement with segMAN reference segmentations at the whole hypothalamus (segATLAS DSC = 0.89 ± 0.06 SD; segFS DSC = 0.68 ± 0.14 SD) and subunit levels (segATLAS DSC = 0.80 ± 0.16 SD; segFS DSC = 0.40 ± 0.26 SD). In addition, segATLAS (but not segFS) volumes demonstrated stability from near birth to ~ 1 years age ($n = 41$; $R^2 = 25\%$; $p < 10^{-3}$). These findings highlight segATLAS as a valid and publicly available (https://github.com/jerodras/neonate_hypothalamus_seg) pipeline for the segmentation of hypothalamic subunits using human newborn MRI up to 3 months of age collected at resolutions on the order of 1 mm isotropic. Because the hypothalamus is traditionally understudied due to a lack of high-quality segmentation tools during the early life period, and because the hypothalamus is of high biological relevance to human growth and development, this tool may stimulate developmental and clinical research by providing new insight into the unique role of the hypothalamus and its subunits in shaping trajectories of early life health and disease.

KEYWORDS

growth, hypothalamus, infant, MRI, newborn, segmentation, subunit

Practitioner Points

- Registration-based segmentation outperforms current adult-oriented tools.
- Hypothalamic volume is associated with age, independent of intracranial volume.
- Hypothalamic volume stability from 0 to 1 years supports reliability.

1 | INTRODUCTION

Despite making up less than 1% of the brain by volume (Bethlehem et al., 2022; Kijonka et al., 2020; Spindler & Thiel, 2022), the human hypothalamus is critical in the control of several physiological functions necessary for survival, including breathing (Fukushi et al., 2019), sleeping (Mignot et al., 2002), thermoregulation (Van Tienhoven et al., 1979), and energy balance (Timper & Brüning, 2017). The hypothalamus acts by sensing internal bodily states through direct connections to the periphery (i.e., circumventricular organs), and in turn, by responding through the production of hormones responsible for human behaviors including those involving diet (Coll et al., 2007), stress (O'Connor et al., 2000), and mood (Bao & Swaab, 2019). The assertion that the hypothalamus is broadly important to human health is further supported by an accumulating body of evidence that inter-

individual variation in the structure and function of the hypothalamus is related to several complex disorders and diseases including Prader-Willi (Brown et al., 2022), obesity (Thomas et al., 2019), schizophrenia (Koolschijn et al., 2008), autism (Caria et al., 2020; Kurth et al., 2011), and dementia (Bocchetta et al., 2015).

The importance of the hypothalamus to early life growth (Sutton et al., 2021) and nutrient partitioning (Lam et al., 2021) is well-established. In addition, recent evidence across biological scales (e.g., from cells [Bouret, 2009] to brain circuits [Bouret & Simerly, 2006]) and species (e.g., rodents; Dearden & Ozanne, 2015; Lippert & Brüning, 2021; and humans Rasmussen et al., 2021; Rasmussen et al., 2022) suggests that the origins of complex disorders can be traced, in part, back to developmental processes occurring during the intrauterine period of life. In this context, the developing hypothalamus may respond to “suboptimal” conditions by producing

structural and functional changes that persist across the life span (i.e., the concept of fetal, or developmental, programming of health and disease risk). As recently argued (Rasmussen et al., 2021), this supports the importance of characterizing the structure and function of the hypothalamus at birth, a developmental period that reflects the cumulative sum of influences by in utero exposures during fetal brain development, yet precedes influence by the postnatal environment under which the hypothalamus further adapts. However, despite the abundant preclinical evidence supporting the role of the hypothalamus in the intergenerational transfer of disease risk (Balsevich et al., 2016; Bouret, 2009; Dearden & Ozanne, 2015; Rivet et al., 2022), very little in vivo newborn human hypothalamus data are currently available. One key hurdle in this respect is the relative paucity of methods available for the automatic segmentation of the newborn hypothalamus in modalities capable of in vivo characterization (e.g., MRI). This work aims to precisely address this through the identification of an automated newborn hypothalamus segmentation pipeline.

Even in adults, until very recently, the hypothalamus has not been included in most MRI segmentation software packages. Instead, packages often include a region-of-interest (ROI) that covers the ventral diencephalon (Lebedeva et al., 2017) at a scale that is far more inclusive than the hypothalamus, let alone the nuclei that constitute its cellular and functional heterogeneity (Neudorfer et al., 2020). In recognition of this knowledge gap, FreeSurfer recently introduced a validated deep-learning based method for the reliable segmentation of the *adult* hypothalamus and its five regional subunits (anterior inferior, anterior superior, posterior, inferior tubular, and superior tubular) (Billot et al., 2020). One example of the potential impact of hypothalamus subunit segmentation lies in the context of obesity where it is well-established that highly specific nuclei/subunits like the arcuate nucleus/inferior tubular region regulate appetite (Jais & Brüning, 2021). The importance of acknowledging the heterogeneity of the hypothalamus through subunit segmentation also generalizes well to the study of other complex disorders, for example, those dependent on the hypothalamic–pituitary–adrenal axis (e.g., schizophrenia [Goldstein et al., 2007], mood disorders [Bao & Swaab, 2019], anxiety [Fischer, 2021], and dementia [Bocchetta et al., 2015]) wherein the paraventricular nucleus/superior tubular region plays a dominant role in neuroendocrine processes. While FreeSurfer have now taken steps to address this shortcoming in hypothalamus segmentation for adults, it is unknown if these tools are applicable to the early life period when accurate segmentation is challenged by reversed MR image contrast due to myelination dynamics.

The aim of the current work was to identify a hypothalamus segmentation pipeline for use in neonatal structural MRI. Towards this, and in addition to be accurate and precise, we considered the following four design criteria. First, the subunit segmentation should be consistent with that found in adults (i.e., FreeSurfer) to facilitate the longitudinal characterization of hypothalamic growth and development. Second, we aimed to focus on reliability through full automation to address inter-rater subjectivity in voxel assignment and reduce the labor involved with manual segmentation. Third, the pipeline ought to

be simple, open, and robust to the image acquisition strategy to promote its adoption by the research community. And fourth, we considered the validity of the pipeline outputs by demonstrating inter-individual stability over time.

We undertook a study leveraging four cohorts with paired newborn T1- and T2-weighted MR images. We began via the “naïve” application of adult-trained FreeSurfer deep-learning networks (referred to as segFS from here on out) to newborn brain images. Here, we use the term naïve to acknowledge that the application of a neural network (segFS) trained on adults is unlikely to succeed in the context of neonatal MRI due to profound differences in morphology and image contrast. We then compared the segFS pipeline to a separate automatic registration-based approach (segATLAS) designed to be more robust to infant image characteristics, yet compatible with the FreeSurfer output (e.g., consistent in indexing convention and boundary definition). Based on an observed disagreement between segFS and segATLAS segmentations, and in the absence of a “ground truth,” we then performed semi-automatic segmentation via manual correction to build out a reference segmentation (segMAN). Finally, having identified a candidate pipeline, we then tested the stability of the derived phenotype (hypothalamus volume, $n = 41$ longitudinal pairs, single cohort) across the first year of life. Collectively, based on the provided evidence, we introduce a novel, validated, and publicly available (https://github.com/jerodras/neonate_hypothalamus_seg) pipeline for the segmentation of hypothalamic subunits using newborn MRI.

2 | MATERIALS AND METHODS

2.1 | Study population

Data from four sites ($N = 215$) were used for primary analyses (e.g., establishing and testing segmentation pipelines), while data from two additional sites ($N = 59$) were visually confirmed for accuracy, as a supplementary analysis. The four primary sites contributed an initial set of 219 total T1/T2-weighted image pairs drawn from a possible of 365 pairs ($n = 48$ removed due to missing T2-weighted image, $n = 43$ removed due to missing T1-weighted image, and $n = 55$ removed for poor image quality as determined by the contributing sites). After receiving the initial set of 219 images, four additional pairs were removed based on the image quality (see Supplementary Section S1: *Quality Control Rejected Participants Were Clear Outliers*, for example images), resulting in a final sample size of $N = 215$ T1-/T2-weighted image pairs available for evaluation. Basic demographics (age and sex) relevant to the analyses are provided below (Table 1). Site-specific study exclusion criteria are outlined below.

2.1.1 | Site 1 specifics

Mother–child dyads were part of a prospective cohort study at the University of California, Irvine's Development, Health and Disease

TABLE 1 Basic infant demographics.

	Site 1 (n = 101)	Site 2 (n = 39)	Site 3 (n = 50)	Site 4 (n = 25)
Infant sex (n _{M/F})	53/48	20/19	25/25	12/13
Gestational age at birth (weeks [SD]) ^a	39.2 (1.5)	39.0 (1.2)	39.7 (1.1)	38.6 (2.1)
Postnatal age at scan (weeks [SD]) ^b	3.7 (1.8)	4.0 (2.0)	6.8 (2.4)	9.2 (5.1)
Postmenstrual age at scan (weeks [SD]) ^c	42.9 (2.0)	43.0 (2.6)	46.5 (2.7)	47.8 (5.5)

Note: Infants were born to mothers recruited for the study of pregnancy conditions in a normative sample representative of the local population demographic (i.e., healthy pregnancies).

^aGestational age at birth at site 3 was larger on average than those at sites 2 and 4 ($p < .05$).

^bSites differed in postnatal age at scan as follows: site 1 = site 2 < site 3 < site 4 ($p < .05$).

^cInfants from sites 3 and 4 were older on average than infants from sites 1 and 2 ($p < .05$).

Research Program (subsequently referred to as “Site 1”). The Site 1 study was designed to investigate the association between maternal conditions during pregnancy and offspring development. Pregnant women attending antenatal care at clinics affiliated with the UCI Medical Center in Orange County, California, were enrolled in the study. The present analysis considered $n = 101$ offspring of mothers enrolled between March 2011 and December 2013 whose children underwent a brain MRI scan shortly after birth and at 12 months of age ($n = 41$ longitudinal pairs). Exclusion criteria were as follows: mother less than 18 years of age; non-singleton/intrauterine pregnancy; diabetes; maternal use of psychotropic medications or systemic corticosteroids during pregnancy; infant birth before 34 weeks of gestation; and infant congenital, genetic, or neurologic disorder. The Institutional Review Board of the University of California, Irvine, approved all study procedures, and all parents provided written, informed consent (UCI IRB: #2009-7251).

2.1.2 | Site 2 specifics

Mother–child dyads were part of a prospective cohort study at the University of Pittsburgh Medical Center (UPMC, subsequently referred to as “Site 2”) that is part of the Environmental Influences on Child Health Outcomes (ECHO) Program. The study enrolled pregnant women attending antenatal care at clinics were affiliated with the UPMC Magee-Women’s Hospital in the Pittsburgh metropolitan area. The present analysis included $n = 39$ offspring of mothers enrolled between November 2019 and May 2021 whose children underwent a brain MRI scan shortly after birth. Inclusion criteria were women self-reporting as Black, White, or bi-racial (with an endorsement of black or white); ≤ 16 weeks gestation at recruitment; and fluency in English. Exclusion criteria were as follows: mother less than 18 years of age; non-singleton/intrauterine pregnancy; infant birth before 34 weeks of gestation; presence of any conditions that may dysregulate endocrine, immune, or oxidative state or systemic inflammation, such as autoimmune disorders requiring chronic systemic steroid or immunomodulator use; and presence of congenital malformations or fetal chromosomal abnormalities. The Institutional Review Board of the University of Pittsburgh approved all study

procedures, and all parents provided written, informed consent (IRB: #19080274).

2.1.3 | Site 3 specifics

Mother–child dyads were part of a prospective cohort study at the University of Rochester Medical Center (subsequently referred to as “Site 3”) that is part of the Environmental Influences on Child Health Outcomes (ECHO) Project. The study enrolled pregnant women attending antenatal care at clinics affiliated with the Medical Center in Rochester, New York. The present analysis included $n = 50$ offspring of mothers enrolled between December 2015 and December 2020 whose children underwent a brain MRI scan shortly after birth. Exclusion criteria were as follows: mother less than 18 years of age; non-singleton/intrauterine pregnancy; Type 1 diabetes; maternal use of psychotropic medications or hormonal medications during pregnancy; infant birth before 37 weeks of gestation; and infant congenital, genetic, or neurologic disorder. The Institutional Review Board of the University of Rochester approved all study procedures, and all parents provided written, informed consent (RSRB: #00058456).

2.1.4 | Site 4 specifics

Mother–child dyads were part of a prospective cohort study at the New York State Psychiatric Institute (subsequently referred to as “Site 4”). Site 4 research activities were conducted in the context of the Boricua Youth Study—Environmental Influences on Child Health Outcomes (BYS-ECHO) Project. The study enrolled pregnant women who themselves were part of the original BYS (Bird et al., 2006) or whose partner was enrolled in the BYS. The present analysis included $n = 25$ offspring of mothers who underwent a brain MRI scan shortly after birth. Of these, $n = 4$ were missing gestational age at birth and postnatal age at scan in the available database but were included for analysis as postmenstrual age at scan was available. Infants were excluded from the study for MRI contraindications or significant medical concerns as determined by a study physician. The Institutional Review Board of the New York State Psychiatric

TABLE 2 Study acquisition parameters.

Site	Sample (n)	Vendor	Resolution (mm)	TR (ms)	TE (ms)	TI (ms)	Flip angle (°)
1	101	Siemens	1.0 iso	2400	3.16	1200	8
2	39	Siemens	0.8 iso	2400	2.22	1000	8
3	50	Siemens	0.8 iso	2400	2.22	1000	8
4	25	GE	0.9 iso	8.24	3.25	400	13

Note: T1-weighted imaging parameters are described.

Abbreviations: TE, echo time; TR, repetition time.

Institute approved all study procedures, and all parents provided written, informed consent (IRB: #7377).

2.2 | Data collection procedures

2.2.1 | Demographics

Demographic data required for analyses (gestational age at birth and infant sex) were collected via a combination of medical record abstraction, questionnaire, and participant interview. Postmenstrual age at scan was defined as the sum of gestational age at birth and chronological age at scan. Gestational age at birth was determined by the best obstetric estimate with a combination of the last menstrual period and early uterine size and was confirmed by obstetric ultrasonographic biometry before 15 weeks using standard clinical criteria at all sites. Infant sex was abstracted from the medical record, and infant scan age was recorded on the day of scan.

2.2.2 | MRI data collection

All MRI scans (including serial scans at 1 year of age at Site 1) were acquired during natural sleep after feeding and soothing. MRI data collection procedures broadly overlapped across sites, in that they were performed at a common field strength (3 Tesla) but differed in the vendor and acquisition strategy, as described below (Table 2).

2.3 | Hypothalamus segmentation approaches

2.3.1 | Naïve application of FreeSurfer hypothalamus segmentation (segFS)

The segFS pipeline (<https://surfer.nmr.mgh.harvard.edu/fswiki/HypothalamicSubunits>) (Billot et al., 2020) was applied to the newborn T1-weighted images despite the tools having been previously designed for, and validated in, adults. Sites not natively collected at 1 mm isotropic spatial resolution were resampled to 1 mm isotropic spatial resolution in accordance with standard segFS procedures in order to match the resolution at which the networks were trained. Input images were not prospectively corrected for signal intensity heterogeneity as segFS was designed to handle image intensity

heterogeneity via data augmentation in training. Output segmentations were stored in the FreeSurfer compatible indexing schema consisting of the five lateralized (10 total) subunits previously validated at high resolution (Makris et al., 2013). In all analyses, subunits were bilaterally averaged for increased statistical power and the absence of well-founded a priori hypotheses about lateralization.

2.3.2 | Atlas-based registration for hypothalamus segmentation (segATLAS)

In order to assess the accuracy of the segFS approach naïvely applied to infants, we aimed to register a segFS atlas created in adult Montreal Neurological Institute (MNI) template-space to the individual native T1-weighted space (Figure 1) via a novel pipeline that leveraged recently published high-quality age-appropriate templates (Baby Connectome Project, BCP) (Chen et al., 2022) and robust infant brain extraction methods (ANTsPyNet, <https://github.com/ANTsX/ANTsPyNet>) (Tustison et al., 2021). First, the segFS pipeline described above was applied to the widely used 1-mm isotropic MNI template and visually confirmed to be accurate. Next, native T1-weighted images were brain-extracted using ANTsPyNet using T1-weighted/T2-weighted imaging pairs, consistent with the BCP age-specific templates. Having confirmed an accurate segmentation of the hypothalamus in MNI-space and robust brain extraction, we non-linearly registered the MNI template to the native T1-weighted space via an age-appropriate middle registration point (Baby Connectome Project, BCP) using Advanced Normalization Tools' (ANTs, <http://stnava.github.io/ANTs/>) *antsRegistrationSyN.sh* function. Following visual confirmation of accurate registration with an emphasis on the region surrounding the hypothalamus, the non-linear transforms were concatenated and applied to the segFS-derived atlas using the “GenericLabel[Linear]” option in ANTs' *antsApplyTransforms* function to result in a native space segFS-compatible hypothalamic subunit segmentation.

2.3.3 | Semi-automatic hypothalamus segmentation via manual correction (segMAN)

Ground-truth segmentations using a semi-automatic approach were created to further assess segmentation performance. Specifically, manual corrections to binary hypothalamus assignment were made on

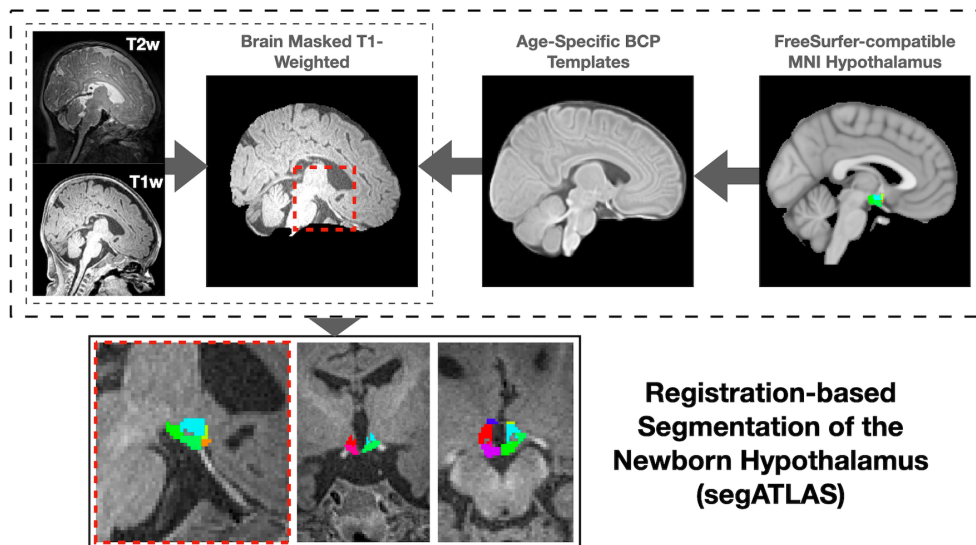


FIGURE 1 Atlas-based registration pipeline (segATLAS). A registration pipeline was designed to leverage high quality age-appropriate atlases in combination with a robust brain extraction procedure. Collectively, the proposed pipeline uses an age-appropriate middle registration point to warp a FreeSurfer compatible definition of hypothalamic subunits into native infant space for segmentation. BCP, Baby Connectome Project; MNI, Montreal Neurological Institute.

top of the intersection between segATLAS segmentation and an alternative multi-atlas-based registration approach (segMAS, see Supplementary Section S2, *Multi-Atlas-based Registration for Hypothalamus Segmentation*). The rationale for including a second registration approach (segMAS) was to reduce bias toward the segATLAS definition of the hypothalamus in the starting/initialization point for manual corrections. In brief, for each individual, segMAS identifies five other individuals matched for postmenstrual age and site with “successful” segFS segmentations, non-linearly registers (*antsRegistrationSyN.sh*) the set of five and their respective segFS segmentations to the individual, and then uses a majority vote algorithm to provide a hypothalamus segmentation largely independent of the segATLAS definition. Notably, because segMAS requires a two-stage process to identify candidate segmentations for atlas construction, it was not considered a viable method for widespread use. The segATLAS and segMAS segmentations were then combined into an intersection (segATLAS and segMAS) and union (segATLAS or segMAS) set as an underlay for manual correction. Manual corrections were made in accordance with the pre-defined borders used by Billot et al. (2020) to define segFS regions and further outlined in Makris et al. (2013). All $n = 215$ datasets in the primary analyses were manually corrected in this manner. A more complete description of the segMAS pipeline and manual correction validation and protocol is provided as Supplementary Material (Section S3: *Validating Manual and Registration-based Segmentation Protocol*).

2.4 | Hypothalamus segmentation output assessment overview

The three main segmentation approaches (segFS, segATLAS, and segMAN) were assessed in a systematic manner reflecting the study design aimed at identifying a FreeSurfer consistent segmentation that is fully automated, robust, and accurate. We began by comparing the segFS versus segATLAS segmentations to determine the suitability of

adult-trained networks applied to newborn data. This comparison was done qualitatively (visual inspection) and quantitatively via the Dice Similarity Coefficient (DSC, MATLAB's *dice* function) and the distance measures average distance (dA, average minimum distance between surfaces) and the Hausdorff distance (dH, maximum distance between surfaces). The DSC characterizes the amount of overlap between segmentations and is formally defined as twice the number of elements in intersection divided by the sum of the number of elements in each set. A DSC of 0 indicates no overlap, while a DSC of 1 indicates perfect overlap. Previous reports of median DSCs in the context of the adult hypothalamus range from values of 0.42 (expert inter-rater DSCs for small subunits) to 0.89 (expert intra-rater DSC for the whole hypothalamus) (Billot et al., 2020). After observing disagreement between segFS and segATLAS segmentation approaches, we chose to further interrogate these differences by regressing the DSCs against postmenstrual age at scan to check for age-associations (i.e., assess whether disagreement decreases as the brain matures) and quantifying differences at the voxel level by the percentage of participants in which voxels were identified in one set but not the other (segATLAS but not segFS, and vice versa). Finally, both (segATLAS and segFS) approaches were compared against the manually corrected segmentation (segMAN) definitions using DSCs and by extracting hypothalamus volume to compare differences.

2.5 | Applied validation: Longitudinal stability analysis

Longitudinal stability of the inter-individual variation in derived volumes was assessed to help establish the validity and reliability of the segmentation. Longitudinal stability validation was performed via registration-based segmentation done precisely as above (segATLAS) using the BCP 12-month template as the middle registration point. Volumes were then residualized for postmenstrual age at scan. Finally, newborn and 1-year-old derived volumes (adjusted for postmenstrual

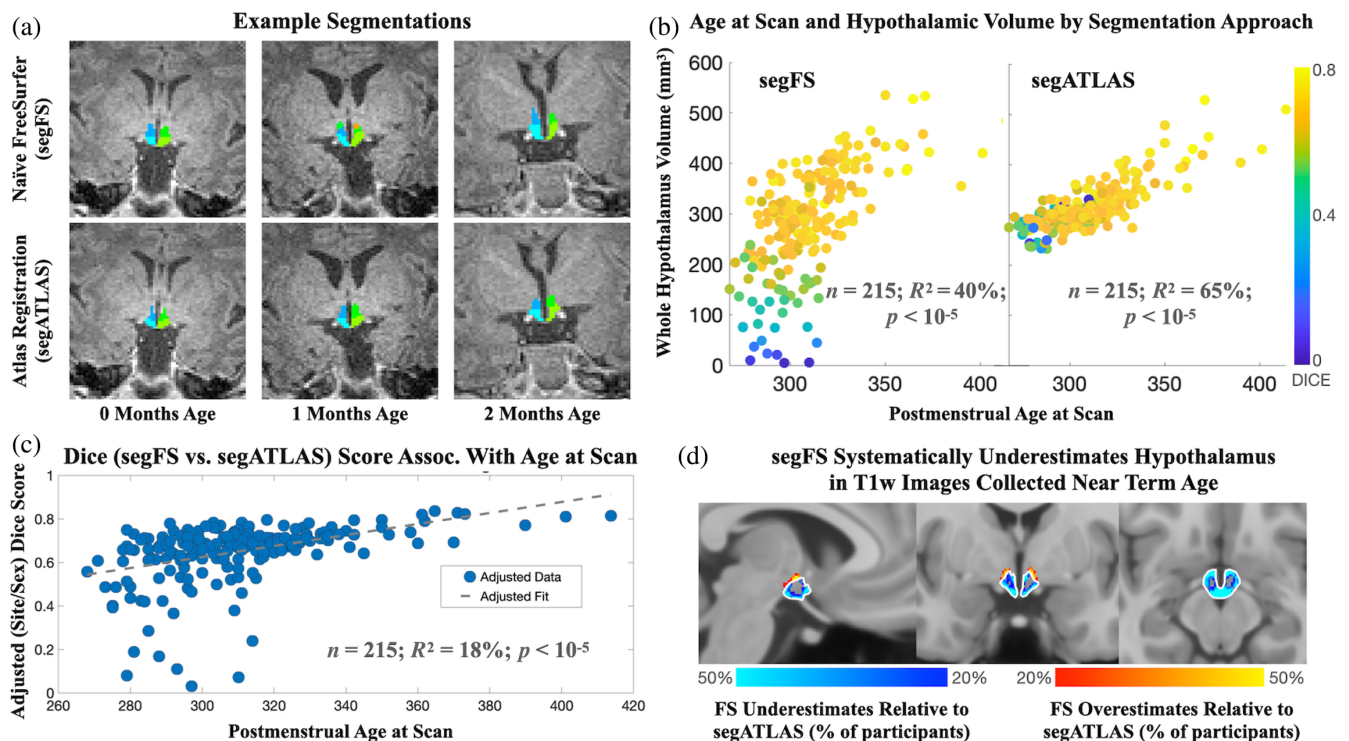


FIGURE 2 Summary of comparison between Naïve FreeSurfer (segFS) and atlas-based registration (segATLAS) approaches to segmentation. Segmentation approaches (segFS and segATLAS) varied in their degree of agreement (example of good agreement, a). SegFS had several instances of gross underestimation (e.g., data points below 200mm³, b), and segATLAS volume was more strongly associated with postmenstrual age at scan than segFS (b). Segmentations in disagreement (segFS vs. segATLAS) tended to be younger (c) and underestimated by segFS (d, white border indicates whole-hypothalamus border in template space).

age at scan) were regressed against one another, and stability was reported as the coefficient of determination (R^2). Similar procedures were performed for segFS defined volumes, and previously validated measures of semi-automatically segmented intracranial volume (ICV) and amygdala (Graham et al., 2017) were reported for contextual comparison. Of note, while we do not assert that these measures are directly comparable due to differences in measurement and growth and make no a priori hypotheses regarding the degree of relative stability between these structures, we do believe that their comparison provides some context as to the relative reliability/stability of the hypothalamic segmentations performed here.

2.6 | Code and data sharing plans

In accordance with the Material Design Analysis Reporting (MDAR) guidelines promoting access to underlying data and code, de-identified and unprocessed MRI data are available where applicable. Specifically, Site 1 data are publicly accessible through the National Institute of Mental Health Data Archive Collection #1890 (https://nda.nih.gov/edit_collection.html?id=1890). Unprocessed data from sites 2, 3, and 4 are available through investigator approval and the execution of the necessary Data Use Agreements (DUAs). The code and pipeline described here are fully available and supported via a GitHub repository (https://github.com/jerodras/neonate_hypothalamus_seg).

3 | RESULTS

3.1 | Comparison between Naïve FreeSurfer (segFS) and atlas-based registration (segATLAS) approaches

SegFS and segATLAS were compared qualitatively (visual inspection) and quantitatively using DSCs as measures of agreement. Under visual inspection, whole hypothalamus segmentations were often consistent with one another (see example segmentations, Figure 2a), with the exception of several cases presenting as gross underestimation of the hypothalamus. In a supplementary analysis, visual inspection of two additional external holdout cohorts with differing acquisition approaches (vendor, resolution, and head coil) demonstrated no gross errors in segmentation and were consistent with segmentations contained in the primary dataset used here (see Supplementary Section-S4: *SegATLAS Pipeline Generalizes to Novel Cohorts*). In line with visual inspection, DSCs comparing segFS to segATLAS had a wide spread of agreement ranging from virtually no agreement to good agreement (mean DSC = 0.65 ± 0.14 SD; range = {0.03–0.80}). Whole hypothalamus volume determined by segATLAS was more strongly associated with postmenstrual age at scan (Figure 2b; $R^2 = 65\%$; $p < 10^{-5}$) than segFS ($R^2 = 40\%$; $p < 10^{-5}$), including when controlling for intracranial volume ($R^2_{\text{partial}} = 7.6\%$ and 6.2% respectively; both $p < 10^{-5}$). A main effect of sex was present (male 4.2% larger than female; $p < 10^{-3}$).

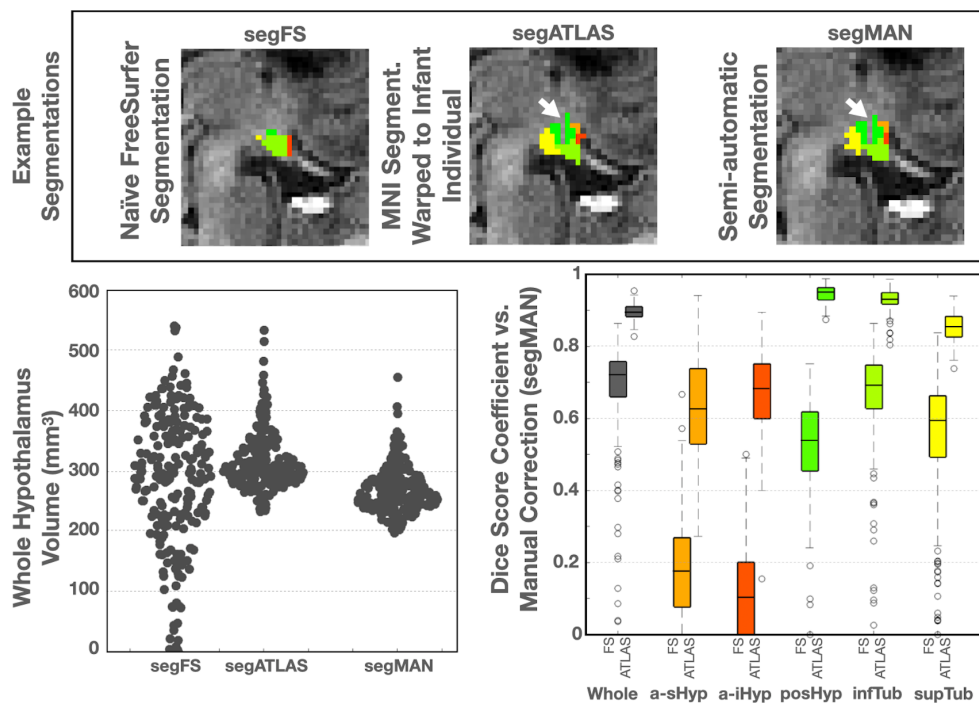


FIGURE 3 Comparison between manual corrections (segMAN) and automatic approaches to segmentation (segFS and segATLAS). Segmentation approaches were compared against manually edited standards. Example segmentations are shown in the top row (note the fornix is properly excluded from the segmentation, white arrow). Bottom left depicts whole hypothalamic volume across methods. SegFS volumes had a relatively widespread encompassing some implausibly small volumes. SegMAN volumes were smaller on average as they were biased (the intersection of segATLAS and segMAS as the initialization point). Bottom right depicts the DSC values from each of the three segmentation approaches in reference to segMAN segmentations.

using the segATLAS derived volumes but not the segFS derived volumes ($p > 0.1$); however, the main effect of sex was no longer present ($p > 0.1$) after adjusting for intracranial volume. On average, those hypothalamic segmentations that were in disagreement (low DSC) between segmentation methods were also younger (Figure 2c; $p < 10^{-5}$) and underestimated by segFS (Figure 2d).

3.2 | Comparison between manual edits (segMAN) and automatic segmentation (segATLAS and segFS) approaches

Manually edited segmentations (segMAN) were quantitatively (DSCs, volumetrics) compared with the two automatic segmentation approaches used here (segFS and segATLAS). SegATLAS had higher agreement than segFS (Figure 3, lower right). SegATLAS demonstrated acceptable (defined here as median DSC > 0.7) DSCs in whole and subunit hypothalamus segmentations, with the exception of the two small anterior subunits (anterior-inferior and anterior-superior hypothalamus). SegFS agreement with segMAN was acceptable (median DSC = 0.72) at the whole hypothalamus level on aggregate but with a relatively high failure rate (defined here as greater than 15% of the sample at a DSC < 0.6 threshold). At the segFS subunit level, the failure rates ranged from as low as 54% in the larger inferior-tubular region to 100% in the two small anterior subunits. In

contrast, the segATLAS segmentations had low failure rates ($< 5\%$) in the three larger subunits (superior-tubular, inferior-tubular, and posterior hypothalamus). However, in the two small anterior regions, segATLAS had failure rates of 56% (anterior-superior hypothalamus) and 65% (anterior-inferior hypothalamus). Average (dA) and Hausdorff (dH) distance measures support surface agreement with segATLAS over segFS, relative to the manually corrected segMAN set (Table 3). With respect to volumetrics, the segFS approach demonstrated a relatively high degree of variability compared to segATLAS and segMAN (Figure 3). On average, segMAN was significantly lower in volume and consistent with the manual editing approach whose basis was the intersection set of segATLAS and segMAS segmentations.

3.3 | Validation of atlas-based registration (segATLAS) approach using longitudinal stability

Based on the above considerations demonstrating that the segATLAS approach proposed here is accurate and robust while also satisfying the first three of the initial design criteria (FreeSurfer compatible, fully automated, open source), we sought to demonstrate the validity of the segATLAS segmentations in the context of longitudinal stability.

SegATLAS, but not segFS ($p > 0.1$), hypothalamic volume was longitudinally stable across the first year of life ($n = 41$; $R^2 = 25\%$; $p < 10^{-3}$, Supplementary Figure S10) in a subset of the above data

TABLE 3 Distance measures of hypothalamus segmentations.

	Whole	Ant-sup	Ant-inf	Posterior	Inf. tubular	Sup. tubular
dA(segFS)	0.61 ± 0.44	1.55 ± 0.93	1.46 ± 0.67	1.45 ± 0.76	1.24 ± 0.34	1.16 ± 0.32
dA(segATLAS)	0.13 ± 0.04	0.01 ± 0.01	0.02 ± 0.06	0.17 ± 0.07	0.07 ± 0.04	0.05 ± 0.02
dH(segFS)	2.25 ± 1.06	2.15 ± 1.52	2.39 ± 1.42	3.86 ± 2.15	2.86 ± 1.28	2.23 ± 0.97
dH(segATLAS)	1.13 ± 0.22	0.01 ± 0.10	0.21 ± 0.44	1.77 ± 1.69	1.25 ± 0.74	1.06 ± 0.23

Note: Distance measures reflect the average minimum distance between surfaces (dA) and maximum distance (dH). Semi-automatic segmentations serve as the reference surface for distance measure calculation. Units provided are in mm.

with repeated measures in the first year of life. This suggests that the segATLAS approach can quantify early life phenotypic inter-individual variation that is sufficiently larger than measurement noise for detection. For additional context, we found the stability of the derived volume to be comparable to that of amygdala volume ($R_{\text{Amygdala}}^2 = 23\%$, see also Supplementary Section S5: *Early Life Hypothalamic Volume Stability is Comparable to Amygdala Volume* for further details). With respect to subunit volumetrics, stability measures across the first year of life were low in the two smaller anterior subunits (anterior-inferior = 3%; anterior-superior = 3%) but comparable to the whole hypothalamus in the three larger subunits (posterior = 36%, inferior tubular = 18%, and superior tubular = 19%). While the hypothalamic volume stability effect sizes are smaller when compared to stability in intracranial volume ($R^2 = 55\%$), the stability of hypothalamus volume was significant ($n = 41$; $t = 4.2$; $p < 10^{-3}$) over and above that of intracranial volume (i.e., after also adjusting for newborn and 1-year old ICV).

4 | DISCUSSION

We report on a novel registration-based pipeline for segmentation of the newborn human hypothalamus. The identified pipeline (segATLAS) satisfies the initial design criteria by being compatible with existing tools (i.e., FreeSurfer), fully automated, simple to use and open-sourced, and valid. Specifically, we demonstrate that the segATLAS pipeline is currently superior to existing tools (segFS, FreeSurfer neural networks trained on adults) when applied to infant MRI data and consistent with manually edited segmentations of the newborn hypothalamus (segMAN). In addition, segATLAS segmentations demonstrate validity in that they are associated with postmenstrual age at scan, and relatively stable across the first year of life.

The performance of the segATLAS pipeline can be characterized by its overlap with manually edited segmentations (segMAN). In adults, agreement measures (DSC) between manual expert intra- and inter-rated whole hypothalamus segmentations have been reported as 0.89 and 0.75, respectively (Billot et al., 2020). In this context, the agreement between the atlas-based approach and manually edited segmentations (DSC = 0.89) was comparable to expert intra-rater agreement. Thus, these findings support the assertion that segATLAS segmentations recapitulate manually defined hypothalamus segmentations to a relatively sufficient degree.

We considered the strength of associations with postmenstrual age at scan and the stability of the phenotype across the first year of

life as additional measures of segATLAS performance. Compared to segFS, segATLAS volumes were more strongly associated with postmenstrual age at scan ($R_{\text{segATLAS}}^2 = 65\%$ vs. $R_{\text{segFS}}^2 = 40\%$). While the magnitude of this effect size is not particularly surprising as the infant brain is rapidly growing on a global scale, it is worth noting that the effect of postmenstrual age was significant ($R_{\text{segATLAS}}^2 = 7.6\%$), even when adjusting for intracranial volume, suggesting that there are growth dynamics in the hypothalamus that are independent of global growth and detectable using the proposed segmentation approach. This point is particularly salient when considering that the hypothalamus is known to be a highly relevant structure to early life growth (Lam et al., 2021; Sutton et al., 2021) and that, to our knowledge, no validated tools exist to properly automatically segment this important structure in the newborn brain. In addition to age effects, segATLAS hypothalamic volume demonstrated a relatively high degree of stability across the first year of life. Specifically, its degree of stability ($R_{\text{Hypothalamus}}^2 = 25\%$) was comparable to that of amygdala volume ($R_{\text{Amygdala}}^2 = 23\%$, see Supplementary Section S5: *Early Life Hypothalamic Volume Stability is Comparable to Amygdala Volume*), a brain phenotype/structure whose volume quantification has been well-validated (Morey et al., 2009) and shown to be of high biological relevance (Graham et al., 2017; Ramirez et al., 2020). Based on these considerations, we assert that the current pipeline holds considerable promise for clinical research on intra- and inter-individual differences in the hypothalamus and its role in health and development from birth.

It is worth noting that DSC and surface distance measures between segATLAS versus segMAN were likely biased. Because segMAN initialization was the joint set between segATLAS and segMAS, there is likely to be a bias toward each of these segmentation procedures. We assert that this was necessary in order to begin with FreeSurfer as a prior while also allowing manual intervention under a rigorous protocol to correct segmentations in accordance with previously published procedures (Goldstein et al., 2007) in an efficient and feasible manner. However, because lateral subunit borders are challenging to define in infant MR images, and because manual interventions were a priori decided to be conservative, it is likely that the smaller subunit segATLAS vs. segMAN DSCs and surface-based measures are especially inflated due to this bias and should not be interpreted as accurate. Indeed, consistent with recent literature (Billot et al., 2020; Ruzok et al., 2022), this work found the two smaller anterior subunit segmentations to be unreliable and volumetric measures should be used with caution. Future efforts aimed at improved

imaging methods (e.g., resolution/contrast) will help ensure more reliable discrimination of such structures.

The current manuscript describes a well-established registration-based approach to image segmentation, applied here to the newborn hypothalamus. While emerging neural network techniques show promise for improving speed and accuracy, their performance often diminishes when the neural network, generally trained on a limited dataset with labels, is applied to a new dataset. This challenge is further complicated in the context of infant MRI due to domain shifts including varied data collection strategies and rapid changes in infant brain size and signal contrast. Although transfer learning, a supervised domain adaptation method, can mitigate the issue of domain shift, it still requires expertise in brain annotation, which limits its widespread adoption. It is notable that several research groups are now advancing unsupervised domain adaptation methods (Shang et al., 2022; Zhang et al., 2023), which obviate the need for expert annotation of local data and, while not yet supported, are a promising avenue for fast, accurate, and generalizable segmentation of the newborn hypothalamus. Toward this, the segmentation protocols detailed in this study offer a framework for defining the newborn hypothalamus in a manner that is consistent with a definition commonly used in adolescents and adults. Thus, this work provides a reference point for establishing manual annotations that are relevant for training neural networks capable of segmenting the newborn hypothalamus while also being applicable to the longitudinal study of its development. Furthermore, the performance metrics reported here may provide critical performance benchmarks to guide the development of unsupervised domain adaptation methods capable of enhancing the efficiency and generalizability of hypothalamic segmentation in infants.

Limitations of this pipeline include the relatively large FreeSurfer subunit definitions, limited age-range, and algorithm speeds. First, the hypothalamic subunits used here, despite validation at higher resolutions (Makris et al., 2013), are limited in their biological application as they are large with respect to the size of specific neuronal populations. However, we aimed to remain consistent with the FreeSurfer definitions as we believe the benefit of being able to perform longitudinal measures more readily across the lifespan outweighs the benefit of specificity at current standard image resolution. In addition, it was observed here that the two smaller anterior subunits were not reliably segmented, further justifying the use of relatively large subunits. It is also worth noting that the pipeline designed here is highly adaptable in the sense that any hypothalamic parcellation scheme (e.g., a more finely parcellated hypothalamus; Neudorfer et al., 2020) can be substituted in place of the FreeSurfer scheme by simply replacing the source atlas in MNI space. While the smaller ROIs provided by high-quality delineation of the hypothalamus may not be large enough to provide reliable estimates of volume, they are likely to be of high value as individualized ROIs that can then be propagated into associated functional or diffusion data as seeds or ROIs. Second, in this work, we focused entirely on the infant hypothalamus segmentation as this was internally identified as a current limitation of the available software. While we did perform stability analyses using this pipeline at 12 months, we have not fully validated the pipeline segmentations

at 12 months. However, based on the stability of the phenotype, the generalizability of registration-based approaches, and qualitative in-house assessment of the segmentations at 1 year of age, we expect that this pipeline would be readily applicable beyond the newborn time point, particularly in cases where segFS does not provide accurate segmentations. Lastly, the registration approaches used here took approximately 45 min in duration per subject on a 2.4 GHz 8-Core Intel Core i9 processor. Thus, this pipeline is not currently a strong candidate for extremely large datasets without the use of high-performance clusters. Therefore, future efforts aimed at using deep-learning approaches to segmentation that can provide outputs on the order of seconds would be warranted, particularly as large-scale efforts at phenotyping the developing brain are currently underway.

In the following, we provide the user with specific recommendations and considerations for utilizing segATLAS to its highest potential. First, we recommend that segATLAS be used in infants less than 3 months of postnatal age. After this early life period, we observed increasingly converging segmentations between segATLAS and segFS out to roughly 6 months of age. Second, we recommend that both T1-weighted and T2-weighted inputs be provided when available in order to have the best possible starting point for registration to infant atlases. However, if necessary, T1-only data inputs can be used with apparently minimal impact on segmentation outputs (see Supplementary Section S6: *SegATLAS Segmentation Using Only T1-weighted Input For Masking* for details). Lastly, we recommend using the highest input resolution possible. Due to the hypothalamus' relatively small size, any degradation in resolution will result in meaningful quantization error (see Supplementary Section S7: *Input Image Resolution and the SegATLAS Pipeline* for details). However, because we expect a user base with a diverse set of use cases (e.g., volumetrics vs. individualized seeds for connectivity measures) and expected effect sizes (e.g., nuanced measures of shape vs. high impact clinical endpoints), the user may be best positioned to determine what is sufficient for their research question of interest.

5 | CONCLUSION

In conclusion, the findings from the current study support a hypothalamus segmentation pipeline that is capable of accurately and reliably phenotyping biologically relevant subunits within the human newborn hypothalamus using anatomical (T1- and T2-weighted) MR images collected at resolutions on the order of 1 mm isotropic. Further, we assert that this tool holds great promise in providing further insight into the unique role of the hypothalamus and its subunits in shaping trajectories of early life health and disease.

AUTHOR CONTRIBUTIONS

J.M.R., Y.W., D.A.F., A.M.G., and C.B. contributed to the conceptualization of this project. J.M.R. and Y.W. contributed to the methodology used. J.M.R. and Y.W. contributed to the analysis. J.M.R., D.A.F., J.P., T.G.O., H.N.S., E.Y., N.M., S.E., P.D.W. and C.B. provided valuable resources, and all co-authors contributed to the writing process.

ACKNOWLEDGMENTS

The authors wish to thank our ECHO colleagues; the medical, nursing, and program staff; and the children and families participating in the ECHO cohorts. We also acknowledge the contribution of the following ECHO program collaborators: ECHO components—Coordinating Center: Duke Clinical Research Institute, Durham, NC: Smith PB, Newby LK; Data Analysis Center: Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD: Jacobson LP; Research Triangle Institute, Durham, NC: Catellier DJ; Person-Reported Outcomes Core: Northwestern University, Evanston, IL: Gershon R, Cella D. ECHO Awardees and Cohorts—New York State Psychiatric Institute, New York, NY: Duarte CS; New York State Psychiatric Institute, New York, NY and Columbia University Vagelos College of Physicians and Surgeons, New York, NY: Monk C; University of Puerto Rico, Rio Piedras, PR: Canino G.

FUNDING INFORMATION

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The research reported in this publication was supported by the environmental influences on Child Health Outcomes (ECHO) Program, Office of the Director, National Institutes of Health, under Award Numbers U2COD023375 (Coordinating Center), U24OD023382 (Data Analysis Center), U24OD023319 with co-funding from the Office of Behavioral and Social Science Research (PRO Core), National Institute of Child Health and Human Development (R01 HD060628, K12 HD0952535, K99 HD100593, and K99 HD103912); National Institute of Mental Health (R01 MH091351, R01 MH119510, and R01 MH121070); National Institute of Diabetes and Digestive and Kidney Diseases (R21 DK118578); National Institutes of Health (UG3OD023349); Environmental Influences on Child Health Outcomes (ECHO) Program: Opportunities and Infrastructure Fund under Grants (U2COD023375) EC0360 (Yun Wang) and EC0397 (Jerod M. Rasmussen), Grants UG3OD023349 (Thomas G. O'Connor), and UH3OD023328 (Jonathan Posner, Duarte).

CONFLICT OF INTEREST STATEMENT

The authors declare the following financial interests/personal relationships which may be considered potential competing interests: Damien Fair reports a relationship with Turing Medical Inc that includes employment and equity or stocks. Damien Fair has a patent issued to Framewise Integrated Real-Time Motion Monitoring (FIRMM) software. He is also a co-founder of Turing Medical Inc. The nature of this financial interest and the design of the study have been reviewed by two committees at the University of Minnesota. They have put in place a plan to help ensure that this research study is not affected by the financial interest.

DATA AVAILABILITY STATEMENT

In accordance with the Material Design Analysis Reporting (MDAR) guidelines promoting access to underlying data and code, de-identified and unprocessed MRI data are available where applicable. Specifically, Site 1 data are publicly accessible through the National Institute of Mental Health Data Archive Collection #1890 (https://nda.nih.gov/edit_collection.html?id=1890). Unprocessed data from

sites 2, 3 and 4 are available through investigator approval and the execution of the necessary Data Use Agreements. The code and pipeline described here is fully available and supported via a Github repository (https://github.com/jerodras/neonate_hypothalamus_seg).

ORCID

Jerod M. Rasmussen  <https://orcid.org/0000-0002-9400-7750>

REFERENCES

- Balsevich, G., Baumann, V., Uribe, A., Chen, A., & Schmidt, M. V. (2016). Prenatal exposure to maternal obesity alters anxiety and stress coping behaviors in aged mice. *Neuroendocrinology*, *103*(3–4), 354–368.
- Bao, A. M., & Swaab, D. F. (2019). The human hypothalamus in mood disorders: The HPA axis in the center. *IBRO Reports*, *6*, 45–53.
- 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. (2022). Brain charts for the human lifespan. *Nature*, *604*(7906), 525–533.
- Billot, B., Bocchetta, M., Todd, E., Dalca, A. V., Rohrer, J. D., & Iglesias, J. E. (2020). Automated segmentation of the hypothalamus and associated subunits in brain MRI. *NeuroImage*, *223*, 117287.
- Bird, H. R., Davies, M., Duarte, C. S., Shen, S. A., Loeber, R., & Canino, G. J. (2006). A study of disruptive behavior disorders in Puerto Rican youth: II. Baseline prevalence, comorbidity, and correlates in two sites. *Journal of the American Academy of Child and Adolescent Psychiatry*, *45*(9), 1042–1053.
- Bocchetta, M., Gordon, E., Manning, E., Barnes, J., Cash, D. M., Espak, M., Thomas, D. L., Modat, M., Rossor, M. N., Warren, J. D., Ourselin, S., Frisoni, G. B., & Rohrer, J. D. (2015). Detailed volumetric analysis of the hypothalamus in behavioral variant frontotemporal dementia. *Journal of Neurology*, *262*(12), 2635–2642.
- Bouret, S., & Simerly, R. (2006). Developmental programming of hypothalamic feeding circuits. *Clinical Genetics*, *70*(4), 295–301.
- Bouret, S. G. (2009). Early life origins of obesity: Role of hypothalamic programming. *Journal of Pediatric Gastroenterology and Nutrition*, *48*, S31–S38.
- Brown, S. S. G., Manning, K. E., Fletcher, P., & Holland, A. (2022). In vivo neuroimaging evidence of hypothalamic alteration in Prader–Willi syndrome. *Brain Communications*, *4*(5), fca229.
- Caria, A., Ciringione, L., & de Falco, S. (2020). Morphofunctional alterations of the hypothalamus and social behavior in autism spectrum disorders. *Brain Sciences*, *10*(7), 435.
- Chen, L., Wu, Z., Hu, D., Wang, Y., Zhao, F., Zhong, T., Lin, W., Wang, L., & Li, G. (2022). A 4D infant brain volumetric atlas based on the UNC/UMN baby connectome project (BCP) cohort. *NeuroImage*, *253*, 119097.
- Coll, A. P., Farooqi, I. S., & O'Rahilly, S. (2007). The hormonal control of food intake. *Cell*, *129*(2), 251–262.
- Dearden, L., & Ozanne, S. E. (2015). Early life origins of metabolic disease: Developmental programming of hypothalamic pathways controlling energy homeostasis. *Frontiers in Neuroendocrinology*, *39*, 3–16.
- Fischer, S. (2021). Chapter 9 – The hypothalamus in anxiety disorders. In D. F. Swaab, F. Kreier, P. J. Lucassen, A. Salehi, & R. M. Buijs (Eds.), *Handbook of clinical neurology* (pp. 149–160). Elsevier. (The Human Hypothalamus; vol. 180). <https://www.sciencedirect.com/science/article/pii/B9780128201077000094>.
- Fukushi, I., Yokota, S., & Okada, Y. (2019). The role of the hypothalamus in modulation of respiration. *Respiratory Physiology & Neurobiology*, *265*, 172–179.
- Goldstein, J. M., Seidman, L. J., Makris, N., Ahern, T., O'Brien, L. M., Caviness, V. S., Kennedy, D. N., Faraone, S. V., & Tsuang, M. T. (2007). Hypothalamic abnormalities in schizophrenia: Sex effects and genetic vulnerability. *Biological Psychiatry*, *61*(8), 935–945.
- Graham, A. M., Rasmussen, J. M., Rudolph, M. D., Heim, C. M., Gilmore, J. H., Styner, M., Potkin, S. G., Entringer, S., Wadhwa, P. D.,

- Fair, D. A., & Buss, C. (2017). <http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=28754515&retmode=ref&cmd=prlinks>). Maternal systemic interleukin-6 during pregnancy is associated with newborn amygdala phenotypes and subsequent behavior at 2 years of age. *Biological Psychiatry*, 83, 109–119.
- Jais, A., & Brüning, J. C. (2021). Arcuate nucleus-dependent regulation of metabolism – pathways to obesity and diabetes mellitus. *Endocrine Reviews*, 43, 314–328. <http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=34490882&retmode=ref&cmd=prlinks>
- Kijonka, M., Borys, D., Psiuk-Maksymowicz, K., Gorczewski, K., Wojcieszek, P., Kossowski, B., Marchewka, A., Swierniak, A., Sokol, M., & Bobek-Billewicz, B. (2020). Whole brain and cranial size adjustments in volumetric brain analyses of sex- and age-related trends. *Frontiers in Neuroscience*, 14, 278. <https://doi.org/10.3389/fnins.2020.00278>
- Koolschijn, P. C. M. P., van Haren, N. E. M., Hulshoff Pol, H. E., & Kahn, R. S. (2008). Hypothalamus volume in twin pairs discordant for schizophrenia. *European Neuropsychopharmacology*, 18(4), 312–315.
- Kurth, F., Narr, K. L., Woods, R. P., O'Neill, J., Alger, J. R., Caplan, R., McCracken, J. T., Toga, A. W., & Levitt, J. G. (2011). Diminished gray matter within the hypothalamus in autism disorder: A potential link to hormonal effects? *Biological Psychiatry*, 70(3), 278–282.
- Lam, B. Y. H., Williamson, A., Finer, S., Day, F. R., Tadross, J. A., Gonçalves Soares, A., Wade, K., Sweeney, P., Bedenbaugh, M. N., Porter, D. T., Melvin, A., Ellacott, K. L. J., Lippert, R. N., Buller, S., Rosmaninho-Salgado, J., Dowsett, G. K. C., Ridley, K. E., Xu, Z., Cimino, I., ... O'Rahilly, S. (2021). MC3R links nutritional state to childhood growth and the timing of puberty. *Nature*, 599(7885), 436–441.
- Lebedeva, A. K., Westman, E., Borza, T., Beyer, M. K., Engedal, K., Aarsland, D., Selbaek, G., & Haberg, A. K. (2017). MRI-based classification models in prediction of mild cognitive impairment and dementia in late-life depression. *Frontiers in Aging Neuroscience*, 9, 13.
- Lippert, R. N., & Brüning, J. C. (2021). Maternal metabolic programming of the developing central nervous system: Unified pathways to metabolic and psychiatric disorders. *Biological Psychiatry*, 91, 898–906. <http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=34330407&retmode=ref&cmd=prlinks>
- Makris, N., Swaab, D. F., van der Kouwe, A., Abbs, B., Boriel, D., Handa, R. J., Tobet, S., & Goldstein, J. M. (2013). Volumetric parcellation methodology of the human hypothalamus in neuroimaging: Normative data and sex differences. *NeuroImage*, 69, 1–10.
- Mignot, E., Taheri, S., & Nishino, S. (2002). Sleeping with the hypothalamus: Emerging therapeutic targets for sleep disorders. *Nature Neuroscience*, 5(11), 1071–1075.
- Morey, R. A., Petty, C. M., Xu, Y., Pannu Hayes, J., Wagner, H. R., Lewis, D. V., LaBar, K. S., Styner, M., & McCarthy, G. (2009). A comparison of automated segmentation and manual tracing for quantifying hippocampal and amygdala volumes. *NeuroImage*, 45(3), 855–866.
- Neudorfer, C., Germann, J., Elias, G. J. B., Gramer, R., Boutet, A., & Lozano, A. M. (2020). A high-resolution in vivo magnetic resonance imaging atlas of the human hypothalamic region. *Scientific Data*, 7(1), 305.
- O'Connor, T. M., O'Halloran, D. J., & Shanahan, F. (2000). The stress response and the hypothalamic-pituitary-adrenal axis: From molecule to melancholia. *QJM: An International Journal of Medicine*, 93(6), 323–333.
- Ramirez, J. S. B., Graham, A. M., Thompson, J. R., Zhu, J. Y., Sturgeon, D., Bagley, J. L., Thomas, E., Papadakis, S., Bah, M., Perrone, A., Earl, E., Miranda-Dominguez, O., Feczko, E., Fombonne, E. J., Amaral, D. G., Nigg, J. T., Sullivan, E. L., & Fair, D. A. (2020). Maternal interleukin-6 is associated with macaque offspring amygdala development and behavior. *Cerebral Cortex (New York)*, 30(3), 1573–1585.
- Rasmussen, J. M., Thompson, P. M., Entringer, S., Buss, C., & Wadhwa, P. D. (2021). Fetal programming of human energy homeostasis brain networks: Issues and considerations. *Obesity Reviews*, 23, e13392.
- Rasmussen, J. M., Thompson, P. M., Gyllenhammer, L. E., Lindsay, K. L., O'Connor, T. G., Koletzko, B., Entringer, S., Wadhwa, P. D., & Buss, C. (2022). Maternal free fatty acid concentration during pregnancy is associated with newborn hypothalamic microstructure in humans. *Obesity*, 30(7), 1462–1471.
- Rivet, T. R., Lalonde, C., & Tai, T. C. (2022). Gene dysregulation in the adult rat paraventricular nucleus and amygdala by prenatal exposure to dexamethasone. *Life*, 12(7), 1077.
- Ruzok, T., Schmitz-Koep, B., Menegaux, A., Eves, R., Daamen, M., Boecker, H., Rieger-Fackeldey, E., Priller, J., Zimmer, C., Bartmann, P., & Wolke, D. (2022). Lower hypothalamus subunit volumes link with impaired long-term body weight gain after preterm birth. *Front Endocrinol (Lausanne)*, 13, 1057566.
- Shang, Z., Turja, M. A., Feczko, E., Houghton, A., Rueter, A., Moore, L. A., Snider, K., Hendrickson, T., Reiners, P., Stoyell, S., Kardan, O., Rosenberg, M., Elison, J. T., Fair, D. A., & Styner, M. A. (2022). Learning strategies for contrast-agnostic segmentation via SynthSeg for infant MRI data. *Proceedings of Machine Learning Research*, 172, 1075–1084.
- Spindler, M., & Thiel, C. M. (2022). Quantitative magnetic resonance imaging for segmentation and white matter extraction of the hypothalamus. *Journal of Neuroscience Research*, 100(2), 564–577.
- Sutton, A. K., Goforth, P. B., Gonzalez, I. E., Dell'Orco, J., Pei, H., Myers, M. G., & Olson, D. P. (2021). Melanocortin 3 receptor-expressing neurons in the ventromedial hypothalamus promote glucose disposal. *Proceedings of the National Academy of Sciences of the United States of America*, 118(15), e2103090118.
- Thomas, K., Beyer, F., Lewe, G., Zhang, R., Schindler, S., Schönknecht, P., Stumvoll, M., Villringer, A., & Witte, A. V. (2019). Higher body mass index is linked to altered hypothalamic microstructure. *Scientific Reports*, 9(1), 1–11.
- Timper, K., & Brüning, J. C. (2017). Hypothalamic circuits regulating appetite and energy homeostasis: Pathways to obesity. *Disease Models & Mechanisms*, 10(6), 679–689.
- Tustison, N. J., Cook, P. A., Holbrook, A. J., Johnson, H. J., Muschelli, J., Devenyi, G. A., Duda, J. T., das, S. R., Cullen, N. C., Gillen, D. L., Yassa, M. A., Stone, J. R., Gee, J. C., & Avants, B. B. (2021). The ANTsX ecosystem for quantitative biological and medical imaging. *Scientific Reports*, 11(1), 9068.
- Van Tienhoven, A., Scott, N. R., & Hillman, P. E. (1979). The hypothalamus and thermoregulation: A review. *Poultry Science*, 58(6), 1633–1639.
- Zhang, X., Wu, Y., Guo, J., Rasmussen, J. M., O'Connor, T. G., Simhan, H. N., Entringer, S., Wadhwa, P. D., Buss, C., Duarte, C. S., & Jackowski, A. (2023). 3D masked autoencoding and pseudo-labeling for domain adaptive segmentation of heterogeneous infant brain MRI. *arXiv*. <http://arxiv.org/abs/2303.09373>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Rasmussen, J. M., Wang, Y., Graham, A. M., Fair, D. A., Posner, J., O'Connor, T. G., Simhan, H. N., Yen, E., Madan, N., Entringer, S., Wadhwa, P. D., Buss, C., & on behalf of program collaborators for Environmental influences on Child Health Outcomes (2024). Segmenting hypothalamic subunits in human newborn magnetic resonance imaging data. *Human Brain Mapping*, 45(2), e26582. <https://doi.org/10.1002/hbm.26582>