

Automatic Geometry-based Estimation of the Locus Coeruleus Region on T₁-Weighted Magnetic Resonance Images

Iman Aganj^{1,2*}, Jocelyn Mora¹, Bruce Fischl^{1,2}, Jean C. Augustinack^{1,2}

¹ Athinoula A. Martinos Center for Biomedical Imaging, Radiology Department, Massachusetts General Hospital, Boston, MA 02129, USA

² Radiology Department, Harvard Medical School, Boston, MA 02115, USA

*** Correspondence:**

Corresponding Author

iaganj@mgh.harvard.edu

Keywords: locus coeruleus (LC), image segmentation, magnetic resonance imaging (MRI), expected label value (ELV), U-Net.

Abstract

The locus coeruleus (LC) is a key brain structure implicated in cognitive function and neurodegenerative disease. Automatic segmentation of the LC is a crucial step in quantitative non-invasive analysis of the LC in large MRI cohorts. Most publicly available imaging databases for training automatic LC segmentation models take advantage of specialized contrast-enhancing (e.g., neuromelanin-sensitive) MRI. Segmentation models developed with such image contrasts, however, are not readily applicable to existing datasets with conventional MRI sequences. In this work, we evaluate the feasibility of using non-contrast neuroanatomical information to geometrically approximate the LC region from standard 3-Tesla T₁-weighted images of 20 subjects from the Human Connectome Project (HCP). We employ this dataset to train and internally/externally evaluate two automatic localization methods, the Expected Label Value and the U-Net. We also test the hypothesis that using the *phase* image as input can improve the robustness of out-of-sample segmentation. We then apply our trained models to a larger subset of HCP, while exploratorily correlating LC imaging variables and structural connectivity with demographic and clinical data. This report contributes and provides an evaluation of two computational methods estimating neural structure.

1 Introduction

The locus coeruleus (LC) is a small elongated hyperpigmented nucleus in the rostral pontine brainstem (1). It synthesizes most of the brain's norepinephrine (2) and is involved in various cognitive functions (3). The LC undergoes neuron loss in the early stages of many neurodegenerative diseases (4-6), such as Alzheimer's disease (7-9) and Parkinson's disease (10, 11) through the accumulation of tau pathology (12, 13) and α -synuclein (14), respectively. Non-invasive assessment of the LC integrity *in vivo*, namely via magnetic resonance imaging (MRI), helps to elucidate how LC degeneration relates to the progression and symptoms of neurodegenerative diseases (4-6). Patterns of structural connectivity of the LC to other brain regions – quantified via diffusion MRI (dMRI) (15, 16) – may further inform us about the pathology distribution in the brain, particularly in the context of Alzheimer's disease, where some hypothesize that tau protein may transmit neuron to neuron from the LC to other areas (17).

Quantitative analysis of the LC from MRI requires knowledge of the LC location. Manual annotation of the LC in a large dataset not only necessitates significant expert effort, but yields a precision limited by moderate inter- and intra-rater variability (18). Automatic LC localization (19-22), which is not yet widely available in conventional neuroimaging toolboxes, is therefore highly desirable, as it can facilitate large-scale imaging studies that would have the power to detect subtle changes in the LC in health and disease.

To enhance the contrast of the LC in the MR image, the high concentration of neuromelanin in the LC (23) has been exploited. To that end, several neuromelanin-sensitive MRI sequences (24) have been successfully employed, including the T_1 -weighted (T_1W) Turbo Spin Echo (TSE) (18, 20, 25-27) and the magnetization transfer (13, 22, 28-31) sequences. The enhanced LC contrast on images acquired with such sequences allows for manual delineation of the LC and the creation of datasets that include gold-standard LC labels. Using a dataset like this for training, a supervised automatic segmentation algorithm can segment the LC on a new similar-contrast image.

Neuromelanin-sensitive MRI, however, typically has a high specific absorption rate (5, 32, 33), and may also be suboptimal for younger adults due to their lower neuromelanin levels (6, 30, 34). Consequently, such a sequence is often not included in large open-access MRI databases of healthy or diseased populations. As for standard MR images (e.g., T_1W images included in almost all MRI databases), the boundaries of the LC cannot be delineated on these images due to the lack of contrast; therefore, the location of the LC can only be approximated at best relative to its surrounding structures using prior neuroanatomical information. Such geometrical localization of the LC might still be useful for some subsequent analyses not requiring high label accuracy. For instance, deep neural networks trained on silver-standard labels have shown to be capable of producing more reliable segmentation results than the labels they were trained on (35).

Several LC atlases are publicly available (16, 18, 26, 27, 30, 31, 36, 37), which can be employed to automatize the localization of the LC via atlas alignment and label propagation. The original datasets used to create these public atlases, however, are not generally available. As a result, users cannot apply other supervised segmentation methods to localize the LC, such as those based on modern deep neural networks (38). A public database of approximate LC region masks accompanied with corresponding MR images with standard (e.g., T_1W , rather than neuromelanin-sensitive) contrasts is thus desirable. Such a database could help the research community to use existing or new methods to develop automatic LC localization tools that are applicable to many available databases, thereby facilitating large-scale retrospective and prospective analyses involving the LC.

Our contributions in this work are as follows:

- We first manually approximate the LC region on the 3T T_1W images of 20 subjects of the open-access *Human Connectome Project (HCP)* (39), sharing the geometrically annotated masks with the research community (see Section 8). To our knowledge, other publicly available datasets with manual LC labels instead contain 3T-TSE/7T T_1W images (33) or functional MR images (40). Our geometrical estimation is based on dimensional (instead of contrast) information and emphasizes the sensitivity of the detection, thereby resulting in LC masks slightly larger than the LC (i.e., encompassing the LC and some surrounding area).
- We then train two automatic segmentation methods of Expected Label Value (ELV) (41) and U-Net (38) on the abovementioned dataset, evaluating the LC localization ability internally as well as on an external dataset (33).

- Inspired by a previous observation (41), we test the hypothesis that using the phase image (i.e., discarding the magnitude Fourier data) would improve the performance of the above models on external datasets (data from different sources).
- We finally apply a trained model to one hundred HCP subjects and analyze the volume, image intensity, and dMRI structural connectivity of the LC masks, correlating them with non-MRI variables.

In the following, we will describe our methods (Section 2), provide our results (Section 3), and discuss them (Section 4).

2 Method

2.1 Manual LC Region Estimation

The human LC is a thin and long column of neurons that extends through multiple levels of the brainstem (1, 5). Located in the rostral pons, the LC is on average 14.5 mm long and 2-2.5 mm wide (1). Due to the lack of LC contrast on T₁W images, we instead used these dimensional landmarks collectively to approximate the LC location: 3 mm lateral from the midline, 1 mm rostral to the fourth ventricle, and 16-20 mm above the pontomedullary junction.

We used Freeview of FreeSurfer (42) to manually create masks of the areas containing each of the left and right LCs on preprocessed 3T T₁W MPRAGE images (T1w_acpc_dc_restore_brain.nii.gz) of the first 20 subjects of the “100 Unrelated Subjects” group of the HCP (39), which had the isotropic voxel size of (0.7 mm)³. Our localization approach prioritizes sensitivity to specificity (i.e., includes more voxels than the LC alone), producing masks that are somewhat inflated compared to the actual LC boundaries.

2.2 Automatic LC Region Estimation

2.2.1 Approaches

We trained two supervised image segmentation methods, both implemented in MATLAB, on our 20-subject dataset to automatically approximate the presumptive left or right LC areas (separately), as described below. We binarized the output soft mask and retained the largest connected component.

We used the ELV supervised segmentation (41) (see Section 8 for toolbox) as our first method, which creates a fuzzy map from a combination of labels suggested by all atlas-to-image transformations, weighted by a measure of transformation validity (without explicit deformable registration). The ELV method inherently uses *phase* images as input (obtained by computing the Fourier transform of the image, discarding the magnitude data, and computing the inverse Fourier transform), which we call “ELV [phase]”. The map can also be modulated by an image intensity prior (41), i.e. “ELV [phase + image]”, to benefit from the image intensity information initially excluded from the phase data.

The second method we used was the U-Net neural network architecture (38), which is a convolutional neural network consisting of a contracting path to capture context, a symmetric expanding path for precise localization, and cross connections. We employed a U-Net with 2 down-sampling layers and 16 initial filters (at the first convolutional layer), with the Dice coefficient as the objective function. We used the Adam optimizer to train the network for 20 epochs on 3D sample patches of size 132×132×132 with a mini-batch size of 8. We initialized the learning rate as 0.002 and dropped it by

95% every 5 epochs. The test subject's LC region was then predicted by averaging the label scores of overlapping patches (stride 10).

2.2.2 Validation

For performance evaluation, we first internally assessed the automatic localization via leave-one-out cross-validation (i.e., trained on 19 images and tested on the remaining image, repeating it for 20 test images). We compared the automatically generated mask with the manual one using the Dice similarity coefficient as the evaluation metric.

For external (out-of-sample) validation, we then applied our models (that had been pretrained on the 20 HCP subjects) to 12 7T T₁W images from the previously unseen dataset by Tona *et al* (33), which had LC labels manually delineated from 3T T₁W TSE images. The images from the latter dataset had the voxel size of $0.70 \times 0.64 \times 0.64$ mm³, thereby requiring resampling to match the HCP resolution of (0.7 mm)³. Since we trained our models on brain-masked HCP data, we extracted the brain in the new dataset using the SPM12 software package (43).

Note that our geometrical LC approximation produces a generous area containing the LC, and, as such, comparing it to the specific label of the actual LC in external validation results in a suboptimal Dice score. For instance, if our localized area has a volume α times larger than that of the LC label, then the Dice will be no higher than $2/(1 + \alpha)$. Nonetheless, this comparison can still help to assess how much our automatic inflated LC neighborhood overlaps with the LC.

2.2.3 Phase Image as Input

We have previously observed that ELV [phase + image] outperformed ELV [phase] in internal cross-validation, but not in external out-of-sample validation (41). We hypothesized that, being less sensitive to inter-dataset variation in image intensity, the phase image might be more robust to domain shift, resulting in better model performance than the image itself would in external validation. We tested this by comparing the two abovementioned ELV variations, as well as comparing the original U-Net (“U-Net [image]”) to a variation of it that received the phase image as input (“U-Net [phase]”) and a two-channel-input variation that received both the image and the phase as input (“U-Net [image + phase]”).

Prompted by the different image intensity distributions of our two datasets, we also experimented with normalizing the input image by its intensity standard deviation during both training and testing (denoting the normalized image with “nmz”), which led to additional variations such as “ELV [phase + nmz]”, “U-Net [nmz]”, and “U-Net [nmz + phase]”. For the latter two U-Net variations, the input layer also performs patch-wise zero-meaning and normalization.

2.3 Analysis of HCP Data

We applied the default implementation of the U-Net to the “100 Unrelated Subjects” group of the HCP. We used the resulting left and right LC masks to compute their volumes as well as the mean T₁W and T₂W image intensities inside them. We then propagated the masks to the dMRI space, and similarly computed the mean fractional anisotropy (FA) and the mean apparent diffusion coefficient (ADC), resulting in a total of 10 LC (regional) imaging variables.

Next, we performed an exploratory analysis to correlate the imaging variables with 504 non-MRI variables (demographics, medical history, family history, dementia/cognitive exam scores, personality/emotion tests, motor/sensory tests, task performance, etc.). We Bonferroni-corrected the Pearson's correlation p -values for multiple comparisons through multiplication by the numbers of

imaging and non-MRI variables, i.e., $p_B = p \times 10 \times 504$. We visually inspected relationships with $p_B < 0.05$ to exclude any spurious correlations due to outliers (e.g., avoiding situations with most data points clustered together with no obvious relationship), reporting the surviving significant correlations.

Finally, we examined the associations between the connectivity pattern of LC and non-MRI variables by quantifying dMRI-derived structural connectivity of the LC area to the rest of the brain. We used our open-source toolbox (see Section 8) to reconstruct the diffusion orientation distribution function in constant solid angle (44), perform Hough-transform global probabilistic tractography (45), compute a symmetric structural connectivity matrix between the two LC areas and 85 other brain regions segmented by FreeSurfer (42), and augment the raw matrices with indirect connections (46). We have previously described this pipeline in detail (47).

3 Results

3.1 Manual LC Region Estimation

We have provided our geometrically annotated (enlarged) LC masks for the 20 HCP subjects to the public (Section 8). As expected, we did not observe any LC contrast on T₁W images to guide the manual delineation of the LC region. Figure 1 shows the approximated LC masks (in green/yellow) for a representative subject.

The volume of the left and right LC regions had a cross-subject mean \pm standard error of the mean (SEM) of 19.5 ± 0.8 mm³ and 19.8 ± 0.8 mm³, respectively. We also measured this for the manual labels in the external (Tona *et al* (33)) dataset; the cross-subject average volumes of the left and right LC labels were 6.9 ± 0.7 mm³ and 7.4 ± 0.8 mm³, respectively. Two-sided paired *t*-tests between the left and right volumes did not reveal a statistically significant difference between them in either dataset.

3.2 Automatic LC Region Estimation

We assessed our variations of the ELV and the U-Net methods (Section 2.2) via leave-one-out cross-validation on the 20-subject subset of the HCP as well as on the external dataset by Tona *et al* (33). Table 1 summarizes the median, mean, and SEM of the Dice scores.

By assessing the left and right LC together, the highest cross-validation Dice scores were achieved by ELV [phase + image] in terms of the mean (0.63 ± 0.01) and by U-Net [image + phase] in terms of the median (0.64). Figure 1 shows the labels generated by both methods for the representative subject with median ELV Dice score. For either ELV or U-Net, the best cross-validation results involved the use of the non-normalized image.

The external validation Dice scores were considerably lower, with the best mean (0.24 ± 0.03) and median (0.27) scores obtained by U-Net [nmz + phase]. Image normalization (nmz) improved the external validation Dice in all cases. The best input for either ELV or U-Net included the phase image, i.e., ELV [phase] and U-Net [nmz + phase].

Left and right LC Dice scores were significantly correlated with each other in most cross-validation results by both methods (ELV [phase + image]: $r = 0.53$, $p = 0.02$; U-Net [image]: $r = 0.61$, $p = 0.004$) and in the external validation results by ELV (ELV [phase]: $r = 0.88$, $p = 0.0001$).

3.3 Findings from HCP

Next, we trained a U-Net (with the default implementation, which receives the image and normalizes the patches at its first layer) on the 20 subjects and applied it to 100 HCP subjects. After correlating non-MRI variables with imaging variables (Section 2.3), only some correlations with the mean FA survived the Bonferroni correction, all of which passed the visual inspection. Table 2 lists these significant relationships, mainly with body weight and memory, and Figure 2 shows two examples. Adjusting for the intracranial volume (ICV) improved the correlation significance with the memory task accuracy, but reduced that with the body weight.

We lastly computed the strength of connectivity between each LC area and 86 other brain regions (85 + contralateral LC). After Bonferroni correction for all possible structural connections to the LC and for all non-MRI variables, none of the few significant correlations between the two passed the visual inspection (see Section 2.3).

4 Discussion

We have presented a new dataset of high-resolution (isotropic 0.7 mm voxel) LC areas manually approximated for 20 subjects of the HCP. This dataset is publicly available to the research community (Section 8), allowing the development of supervised tools applicable to standard 3T T₁W MPRAGE (rather than neuromelanin-sensitive) MRI for quantitative analysis using large existing or future MRI databases. Given the lack of LC contrast on standard T₁W images, we used dimensional information instead to estimate the region that approximately included the LC. We emphasized sensitivity for this task and created masks that were slightly larger than and contained the LC (which, if desired, could be shrunk in post-processing via the *erosion* operation). The masks had a bilateral mean volume of $19.7 \pm 0.6 \text{ mm}^3$, expectedly larger than the LC volume reported in the literature, such as 6.6 mm^3 (20), 7.2 mm^3 (that we computed from the dataset) (33), 9.5 mm^3 (18), 12.8 mm^3 (48), and 16.7 mm^3 (49).

Our internal evaluation of (the optimal variations of) the ELV and U-Net automatic segmentation approaches on our data resulted in a similar cross-validation Dice score of 0.63~0.64 for both methods. In comparison, the LC label Dice scores reported in the literature for inter-rater reliability are 0.50 (20), 0.54~0.64 (18), 0.64 (33), and 0.68 (19), for scan-rescan reliability are 0.24~0.48 (21) and 0.63 (50), and for automatic segmentation are 0.40 (20), 0.54~0.64 (21), and 0.60~0.71 (19).

Our external validation resulted in lower mean Dice scores (ELV: 0.09, U-Net: 0.24). Several reasons could account for this. Given that the average volume of our presumptive LC areas was $\alpha = 2.7$ times larger than that of the LC labels in the test (Tona *et al* (33)) dataset, the Dice score between the two was capped at 0.54 (see Section 2.2.2). The cap was possibly even lower due to inter-rater variability between the two datasets, especially since our LC areas were based on dimensional information whereas the labels in the external dataset were delineated based on the LC contrast seen on neuromelanin-sensitive (3T T₁W TSE) images. Another factor contributing to the lower Dice value may have been the domain shift, particularly caused by the different MRI field strengths of the input training (3T) and test (7T) images, which are known to reveal different MRI tissue properties (51).

We tested the hypothesis that using phase images could enhance out-of-sample segmentation. In line with our previous finding (41), the phase-only ELV map yielded a higher external-validation Dice value *without* additional modulation with the image intensity prior. The U-Net also achieved the highest accuracy in both internal and external validation when the input included the phase image. The different field strengths and acquisition protocols of the training (HCP) and test (Tona *et al* (33)) datasets may have caused inconsistencies in their image intensities. The phase image improved our

external validation results by ignoring the Fourier-domain magnitude information, which possibly alleviated such inter-database inconsistencies to some extent.

The Dice scores corresponding to the approximated left and right LC regions were often significantly correlated with each other, perhaps due to their correlation with the image quality and variance of the test subject.

Following image segmentation of 100 unrelated HCP subjects and an exploratory analysis, we found significant correlations with the mean FA inside the LC region, mainly negative correlations with body weight and a positive correlation with working memory. In a similar analysis, LC connectivity was found not to be significantly correlated with non-MRI variables, possibly due to the homogeneity and narrow age range (22 – 36 years old) of the healthy HCP cohort (47). Our stringent Bonferroni correction for all compared variables and LC connections may additionally have led to type II errors (false negatives). In most related HCP studies, the LC connectivity has been measured to predefined ipsilateral target regions pertinent to disease, such as the transentorhinal cortex (15) and limbic regions (16). In contrast, we ran whole-brain tractography to explore all regions' potential connectivity to the LC, which is especially important considering the LC's extensive axonal branching innervating diverse remote areas throughout the brain (3).

This report has some limitations. First, the LC masks were not based on anatomical contrast and the location is approximate. Second, the LC is difficult to model given its small size. Third, partial voluming effects can introduce noticeable error in the imaging quantity derived from the LC. Should a non-MRI variable and the error in an imaging variable happen to be related to each other (e.g., by both being correlated to a third factor such as the ICV), a spurious relationship might appear in the correlation analysis. Thorough investigation of the automatically misclassified voxels to identify any potential bias is a subject of our ongoing research. For the above reasons, we do not imply our LC masks to be anatomically accurate LC labels, and caution their use where specific and accurate labels are required.

Quantitative MRI studies of the LC – requiring automatic LC region estimation – have the potential to generate imaging biomarkers for early diagnosis of neurodegenerative diseases and patient stratification (6). Making our LC neighborhood localization algorithm more specific and integrating it into FreeSurfer (42) are subjects of our future work.

5 Conflict of Interest

B. Fischl has a financial interest in CorticoMetrics, a company whose medical pursuits focus on brain imaging and measurement technologies. His interests were reviewed and are managed by Massachusetts General Hospital and Mass General Brigham in accordance with their conflict-of-interest policies. The other authors have nothing to disclose.

6 Author Contributions

IA: experiment design, data analysis, writing of manuscript; JM: literature search, data curation; BF: experiment design, feedback on manuscript; JCA: manual approximation of the LC region, manuscript editing.

7 Funding

Support for this research was provided by the National Institutes of Health (NIH), specifically the National Institute on Aging (NIA; R56AG068261, RF1AG068261).

Additional support was provided in part by the NIA (RF1AG082223, R21AG082082, R56AG064027, R01AG064027, R01AG008122, R01AG016495, R01AG070988), the BRAIN Initiative Cell Census/Atlas Network (U01MH117023, UM1MH130981), the Brain Initiative Brain Connects consortium (U01NS132181, UM1NS132358), the National Institute for Biomedical Imaging and Bioengineering (P41EB015896, R01EB023281, R01EB006758, R21EB018907, R01EB019956, P41EB030006), the National Institute of Mental Health (R01MH121885, RF1MH123195), the National Institute for Neurological Disorders and Stroke (R01NS0525851, R21NS072652, R01NS070963, R01NS083534, U01NS086625, U24NS10059103, U24NS135561, R01NS105820), the NIH Blueprint for Neuroscience Research (U01MH093765), part of the multi-institutional Human Connectome Project, and the Michael J. Fox Foundation for Parkinson's Research (MJFF-021226).

Computational resources were provided by the Massachusetts Life Sciences Center.

The HCP WU-Minn Consortium (U54MH091657) was funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

8 Data Availability Statement

Upon acceptance of the paper, our manually annotated (enlarged) LC masks will be publicly available at: <https://www.nitrc.org/projects/lc20>

Magnetic resonance images were provided by:

- the Human Connectome Project (HCP, RRID:SCR_006942) (39), WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil): <https://www.humanconnectome.org/study/hcp-young-adult>
- Klodiana-Daphne Tona *et al* (33): <https://doi.org/10.34894/PMQHZD>

The following toolboxes were used for data processing and analysis:

- Our MATLAB (RRID:SCR_001622) toolboxes for:
 - the expected label value (ELV) supervised image segmentation (41): <https://www.nitrc.org/projects/elv>
 - the reconstruction of the orientation distribution function in constant solid angle (44), Hough-transform tractography (45), and connectivity matrix computation and augmentation (46): <http://www.nitrc.org/projects/csaodf-hough>
- FreeSurfer (RRID:SCR_001847) (42): <https://freesurfer.net>
- SPM12 (RRID:SCR_007037) (43): <https://www.fil.ion.ucl.ac.uk/spm/software/spm12>

9 References

1. Fernandes P, Regala J, Correia F, Gonçalves-Ferreira AJ. The human locus coeruleus 3-D stereotactic anatomy. *Surgical and Radiologic Anatomy*. 2012;34(10):879-85.
2. Aston-Jones G, Cohen JD. AN INTEGRATIVE THEORY OF LOCUS COERULEUS-NOREPINEPHRINE FUNCTION: Adaptive Gain and Optimal Performance. *Annual Review of Neuroscience*. 2005;28(1):403-50.
3. Sara SJ. The locus coeruleus and noradrenergic modulation of cognition. *Nature Reviews Neuroscience*. 2009;10(3):211-23.
4. Galgani A, Lombardo F, Della Latta D, Martini N, Bonuccelli U, Fornai F, et al. Locus Coeruleus Magnetic Resonance Imaging in Neurological Diseases. *Current Neurology and Neuroscience Reports*. 2020;21(1):2.
5. Liu KY, Marijatta F, Hämmerer D, Acosta-Cabronero J, Düzel E, Howard RJ. Magnetic resonance imaging of the human locus coeruleus: A systematic review. *Neuroscience & Biobehavioral Reviews*. 2017;83:325-55.
6. Betts MJ, Kirilina E, Otaduy MCG, Ivanov D, Acosta-Cabronero J, Callaghan MF, et al. Locus coeruleus imaging as a biomarker for noradrenergic dysfunction in neurodegenerative diseases. *Brain*. 2019;142(9):2558-71.
7. Grudzien A, Shaw P, Weintraub S, Bigio E, Mash DC, Mesulam MM. Locus coeruleus neurofibrillary degeneration in aging, mild cognitive impairment and early Alzheimer's disease. *Neurobiology of Aging*. 2007;28(3):327-35.
8. Giorgi FS, Ryskalin L, Ruffoli R, Biagioni F, Limanaqi F, Ferrucci M, et al. The Neuroanatomy of the Reticular Nucleus Locus Coeruleus in Alzheimer's Disease. *Frontiers in Neuroanatomy*. 2017;11(80).
9. Kelly SC, Nelson PT, Counts SE. The locus coeruleus: A link between cerebrovascular and neuronal pathology in preclinical Alzheimer's disease. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*. 2017;13(7):P917.
10. Gesi M, Soldani P, Giorgi FS, Santinami A, Bonaccorsi I, Fornai F. The role of the locus coeruleus in the development of Parkinson's disease. *Neuroscience & Biobehavioral Reviews*. 2000;24(6):655-68.
11. Vermeiren Y, De Deyn PP. Targeting the norepinephrinergic system in Parkinson's disease and related disorders: The locus coeruleus story. *Neurochemistry International*. 2017;102:22-32.
12. Braak H, Thal DR, Ghebremedhin E, Del Tredici K. Stages of the Pathologic Process in Alzheimer Disease: Age Categories From 1 to 100 Years. *Journal of Neuropathology & Experimental Neurology*. 2011;70(11):960-9.
13. Jacobs HIL, Becker JA, Kwong K, Munera D, Ramirez-Gomez L, Engels-Domínguez N, et al. Waning locus coeruleus integrity precedes cortical tau accrual in preclinical autosomal dominant Alzheimer's disease. *Alzheimer's & Dementia*. 2023;19(1):169-80.
14. Braak H, Tredici KD, Rüb U, de Vos RAI, Jansen Steur ENH, Braak E. Staging of brain pathology related to sporadic Parkinson's disease. *Neurobiology of Aging*. 2003;24(2):197-211.
15. Sun W, Tang Y, Qiao Y, Ge X, Mather M, Ringman JM, et al. A probabilistic atlas of locus coeruleus pathways to transentorhinal cortex for connectome imaging in Alzheimer's disease. *NeuroImage*. 2020;223:117301.

16. Levinson S, Miller M, Iftekhar A, Justo M, Arriola D, Wei W, et al. A structural connectivity atlas of limbic brainstem nuclei. *Frontiers in Neuroimaging*. 2023;1.
17. Braak H, Del Tredici K. Alzheimer's pathogenesis: is there neuron-to-neuron propagation? *Acta Neuropathologica*. 2011;121(5):589-95.
18. Tona K-D, Keuken MC, de Rover M, Lakke E, Forstmann BU, Nieuwenhuis S, et al. In vivo visualization of the locus coeruleus in humans: quantifying the test-retest reliability. *Brain Struct Funct*. 2017;222(9):4203-17.
19. Dünwald M, Ernst P, Düzel E, Tönnies K, Betts MJ, Oeltze-Jafra S. Fully automated deep learning-based localization and segmentation of the locus coeruleus in aging and Parkinson's disease using neuromelanin-sensitive MRI. *International Journal of Computer Assisted Radiology and Surgery*. 2021;16(12):2129-35.
20. Ariz M, Abad RC, Castellanos G, Martínez M, Muñoz-Barrutia A, Fernández-Seara MA, et al. Dynamic Atlas-Based Segmentation and Quantification of Neuromelanin-Rich Brainstem Structures in Parkinson Disease. *IEEE Trans Med Imaging*. 2019;38(3):813-23.
21. Sibahi A, Gandhi R, Al-Haddad R, Therriault J, Pascoal T, Chamoun M, et al. Characterization of an automated method to segment the human locus coeruleus. *Human Brain Mapping*. 2023;44(9):3913-25.
22. Morris LS, Tan A, Smith DA, Grehl M, Han-Huang K, Naidich TP, et al. Sub-millimeter variation in human locus coeruleus is associated with dimensional measures of psychopathology: An in vivo ultra-high field 7-Tesla MRI study. *NeuroImage: Clinical*. 2020;25:102148.
23. Zucca FA, Bellei C, Giannelli S, Terreni MR, Gallorini M, Rizzio E, et al. Neuromelanin and iron in human locus coeruleus and substantia nigra during aging: consequences for neuronal vulnerability. *Journal of Neural Transmission*. 2006;113(6):757-67.
24. Sasaki M, Shibata E, Kudo K, Tohyama K. Neuromelanin-Sensitive MRI. *Clin Neuroradiol*. 2008;18(3):147-53.
25. Sasaki M, Shibata E, Tohyama K, Takahashi J, Otsuka K, Tsuchiya K, et al. Neuromelanin magnetic resonance imaging of locus coeruleus and substantia nigra in Parkinson's disease. *NeuroReport*. 2006;17(11):1215-8.
26. Keren NI, Lozar CT, Harris KC, Morgan PS, Eckert MA. In vivo mapping of the human locus coeruleus. *NeuroImage*. 2009;47(4):1261-7.
27. Dahl MJ, Mather M, Düzel S, Bodammer NC, Lindenberger U, Kühn S, et al. Rostral locus coeruleus integrity is associated with better memory performance in older adults. *Nature Human Behaviour*. 2019;3(11):1203-14.
28. Nakane T, Nishashi T, Kawai H, Naganawa S. Visualization of Neuromelanin in the Substantia Nigra and Locus Coeruleus at 1.5T Using a 3D-gradient Echo Sequence with Magnetization Transfer Contrast. *Magnetic Resonance in Medical Sciences*. 2008;7(4):205-10.
29. Priovoulos N, Jacobs HIL, Ivanov D, Uludağ K, Verhey FRJ, Poser BA. High-resolution in vivo imaging of human locus coeruleus by magnetization transfer MRI at 3T and 7T. *NeuroImage*. 2018;168:427-36.
30. Liu KY, Acosta-Cabronero J, Cardenas-Blanco A, Loane C, Berry AJ, Betts MJ, et al. In vivo visualization of age-related differences in the locus coeruleus. *Neurobiology of Aging*. 2019;74:101-11.

31. Ye R, Rua C, O'Callaghan C, Jones PS, Hezemans FH, Kaalund SS, et al. An in vivo probabilistic atlas of the human locus coeruleus at ultra-high field. *NeuroImage*. 2021;225:117487.
32. Chen X, Huddleston DE, Langley J, Ahn S, Barnum CJ, Factor SA, et al. Simultaneous imaging of locus coeruleus and substantia nigra with a quantitative neuromelanin MRI approach. *Magnetic Resonance Imaging*. 2014;32(10):1301-6.
33. Tona K-D, van Osch MJP, Nieuwenhuis S, Keuken MC. Quantifying the contrast of the human locus coeruleus in vivo at 7 Tesla MRI. *PLOS ONE*. 2019;14(2):e0209842.
34. Zecca L, Stroppolo A, Gatti A, Tampellini D, Toscani M, Gallorini M, et al. The role of iron and copper molecules in the neuronal vulnerability of locus coeruleus and substantia nigra during aging. *Proceedings of the National Academy of Sciences*. 2004;101(26):9843-8.
35. Henschel L, Conjeti S, Estrada S, Diers K, Fischl B, Reuter M. FastSurfer - A fast and accurate deep learning based neuroimaging pipeline. *NeuroImage*. 2020;219:117012.
36. Brooks JCW, Davies W-E, Pickering AE. Resolving the Brainstem Contributions to Attentional Analgesia. *The Journal of Neuroscience*. 2017;37(9):2279-91.
37. Betts MJ, Cardenas-Blanco A, Kanowski M, Jessen F, Düzel E. In vivo MRI assessment of the human locus coeruleus along its rostrocaudal extent in young and older adults. *NeuroImage*. 2017;163:150-9.
38. Ronneberger O, Fischer P, Brox T, editors. *U-Net: Convolutional Networks for Biomedical Image Segmentation* 2015; Cham: Springer International Publishing.
39. Van Essen DC, Smith SM, Barch DM, Behrens TEJ, Yacoub E, Ugurbil K. The WU-Minn Human Connectome Project: An overview. *NeuroImage*. 2013;80(Supplement C):62-79.
40. Krebs RM, Park HRP, Bombeke K, Boehler CN. Modulation of locus coeruleus activity by novel oddball stimuli. *Brain Imaging and Behavior*. 2018;12(2):577-84.
41. Aganj I, Fischl B. Multi-Atlas Image Soft Segmentation via Computation of the Expected Label Value. *IEEE Trans Med Imaging*. 2021;40(6):1702-10.
42. Fischl B. FreeSurfer. *NeuroImage*. 2012;62(2):774-81.
43. SPM12 2014 [Available from: <https://www.fil.ion.ucl.ac.uk/spm/software/spm12/>].
44. Aganj I, Lenglet C, Sapiro G, Yacoub E, Ugurbil K, Harel N. Reconstruction of the orientation distribution function in single- and multiple-shell q-ball imaging within constant solid angle. *Magnetic Resonance in Medicine*. 2010;64(2):554-66.
45. Aganj I, Lenglet C, Jahanshad N, Yacoub E, Harel N, Thompson PM, et al. A Hough transform global probabilistic approach to multiple-subject diffusion MRI tractography. *Med Image Anal*. 2011;15(4):414-25.
46. Aganj I, Prasad G, Srinivasan P, Yendiki A, Thompson PM, Fischl B, editors. *Structural brain network augmentation via Kirchhoff's laws*. Annual Meeting of the International Society for Magnetic Resonance in Medicine; 2014.
47. Aganj I, Mora J, Frau-Pascual A, Fischl B, Initiative ftAsDN. Exploratory correlation of the human structural connectome with non-MRI variables in Alzheimer's disease. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*. 2023;15(4):e12511.
48. Theofilas P, Ehrenberg AJ, Dunlop S, Di Lorenzo Alho AT, Nguy A, Leite REP, et al. Locus coeruleus volume and cell population changes during Alzheimer's disease progression: A stereological

study in human postmortem brains with potential implication for early-stage biomarker discovery. *Alzheimer's & Dementia*. 2017;13(3):236-46.

49. Schwarz ST, Xing Y, Tomar P, Bajaj N, Auer DP. In Vivo Assessment of Brainstem Depigmentation in Parkinson Disease: Potential as a Severity Marker for Multicenter Studies. *Radiology*. 2017;283(3):789-98.

50. Langley J, Huddleston DE, Liu CJ, Hu X. Reproducibility of locus coeruleus and substantia nigra imaging with neuromelanin sensitive MRI. *Magnetic Resonance Materials in Physics, Biology and Medicine*. 2017;30(2):121-5.

51. van der Zwaag W, Schäfer A, Marques JP, Turner R, Trampel R. Recent applications of UHF-MRI in the study of human brain function and structure: a review. *NMR in Biomedicine*. 2016;29(9):1274-88.

Table 1. Median, mean, and standard error of the mean (SEM) of the Dice score, measuring the overlap of the automatic and manual LC regions.

Method	Input	Cross-validation Dice score (<i>median</i>) (<i>mean ± SEM</i>)		External validation Dice score (<i>median</i>) (<i>mean ± SEM</i>)	
		Left LC	Right LC	Left LC	Right LC
ELV	phase	0.61 0.61 ± 0.02	0.64 0.63 ± 0.02	0.02 0.08 ± 0.04	0.10 0.10 ± 0.03
	phase + image	0.61 0.61 ± 0.02	0.64 0.63 ± 0.02	0.00 0.00 ± 0.00	0.00 0.004 ± 0.004
	phase + nmz	0.60 0.61 ± 0.02	0.64 0.63 ± 0.02	0.00 0.07 ± 0.04	0.00 0.07 ± 0.03
U-Net	image	0.58 0.57 ± 0.03	0.64 0.60 ± 0.03	0.00 0.00 ± 0.00	0.00 0.00 ± 0.00
	nmz	0.66 0.61 ± 0.02	0.56 0.55 ± 0.04	0.14 0.15 ± 0.03	0.27 0.24 ± 0.03
	phase	0.63 0.61 ± 0.02	0.62 0.60 ± 0.02	0.00 0.03 ± 0.02	0.11 0.13 ± 0.04
	image + phase	0.62 0.60 ± 0.02	0.64 0.60 ± 0.04	0.00 0.00 ± 0.00	0.00 0.00 ± 0.00
	nmz + phase	0.62 0.61 ± 0.02	0.62 0.62 ± 0.02	0.24 0.24 ± 0.04	0.27 0.23 ± 0.03

Table 2. Pearson’s correlation coefficient (r) and Bonferroni-corrected p -value (p_B) of the significant correlations of the mean fractional anisotropy (FA) inside the computed LC area with non-MRI variables, without and with intracranial volume (ICV) adjustment.

Non-MRI Variable	No ICV adjustment		Adjusted for ICV	
	Left LC	Right LC	Left LC	Right LC
Body weight	$r = -0.52$ $p_B = 0.0002$	$r = -0.52$ $p_B = 0.0002$	$r = -0.44$ $p_B = 0.03$	$r = -0.44$ $p_B = 0.02$
Body mass index (BMI)	-	$r = -0.44$ $p_B = 0.02$	-	$r = -0.43$ $p_B = 0.04$
Body mass index (BMI) at heaviest time	-	$r = -0.43$ $p_B = 0.04$	-	$r = -0.43$ $p_B = 0.046$
Working memory task accuracy: 2-back place	-	$r = 0.44$ $p_B = 0.02$	-	$r = 0.46$ $p_B = 0.009$

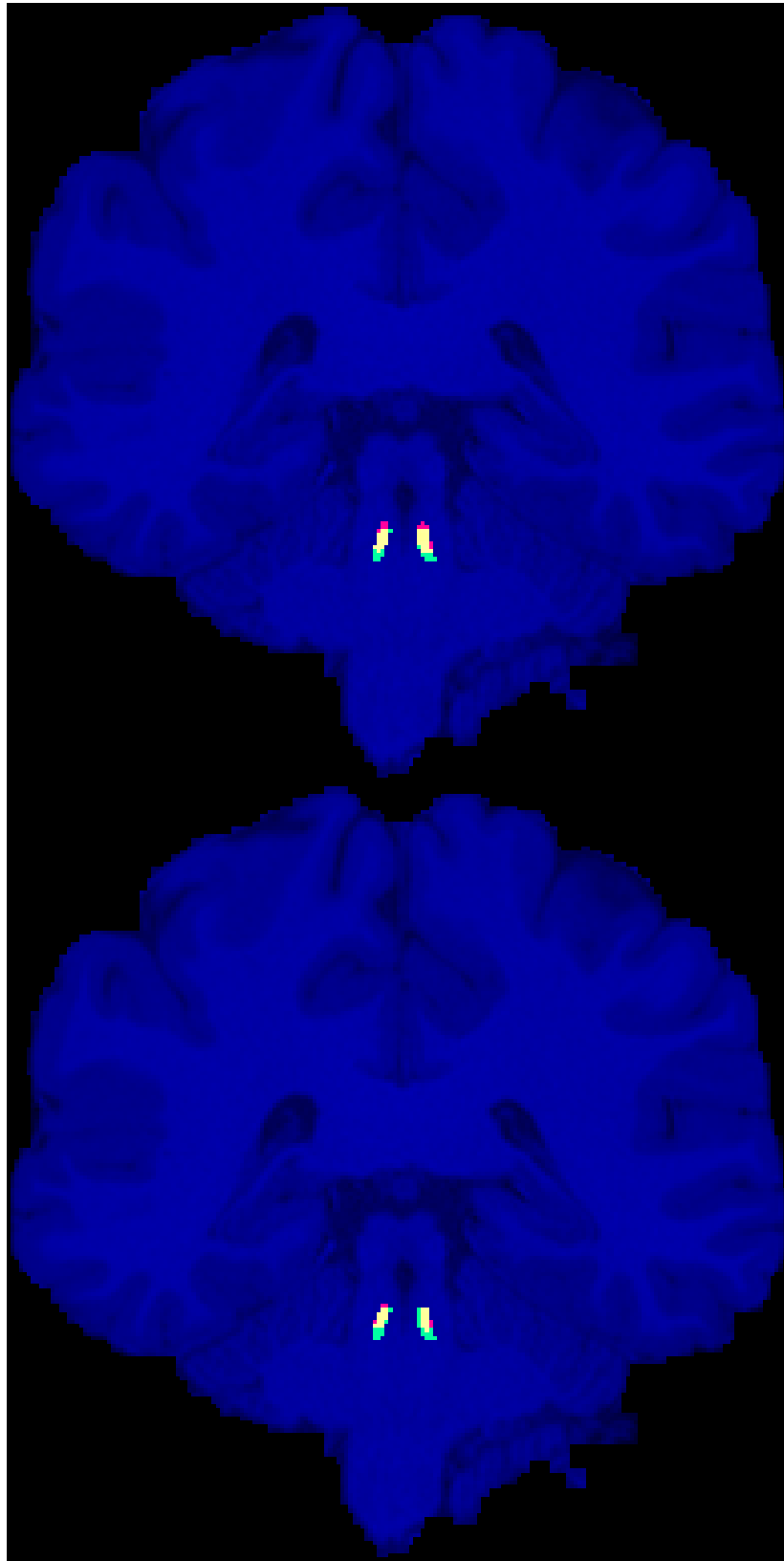


Figure 1. Manually annotated presumptive LC areas (green), automatic localization of LC using the ELV (top) and U-Net (bottom) methods (red), and their overlap (yellow), are shown on the coronal slice with the largest intersection with the manual LC areas for the representative HCP subject (with the median ELV Dice score).

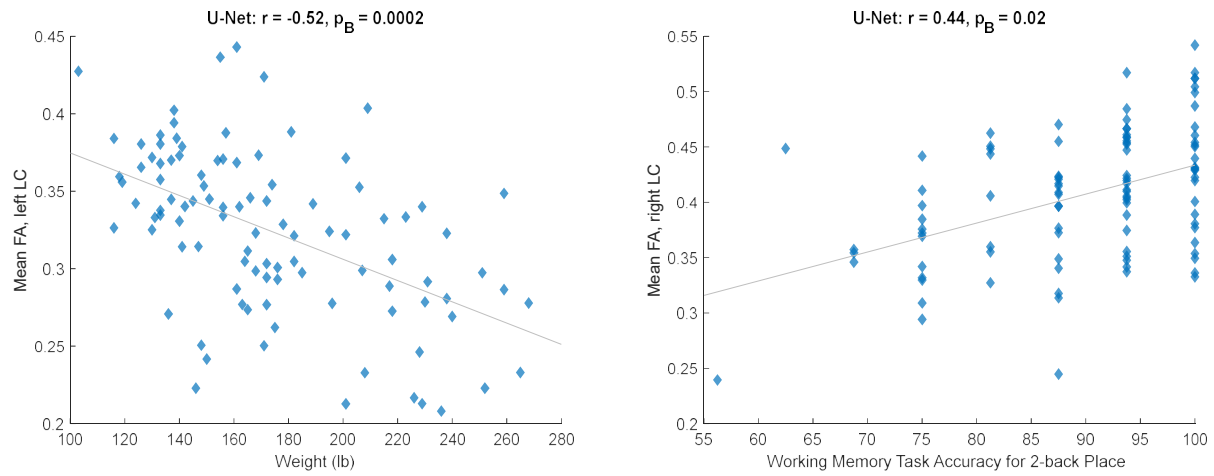


Figure 2. Significant relationships between the mean fractional anisotropy (FA) inside the automatically estimated LC region and non-MRI variables (corresponding to Table 2, unadjusted for ICV).