Morphological bias of the MNI152 brain

Yishai Valter[1,2], Yu Huang[1], Niranjan Khadka[1], Abhishek Datta[1,2], Marom Bikson[2]*

[1] Research and Development, Soterix Medical Inc., Woodbridge, NJ, USA
[2] Biomedical Engineering Department, City College of New York, New York, NY, USA
*Corresponding author: Yishai Valter yvalter@soterixmedical.com

Abstract

The MNI152 template is widely regarded as a representative average brain in neuroimaging, computational modeling, and neuromodulation research; however, its fidelity as a true population mean has not been systematically evaluated. In this study, we compared the MNI152 6th generation nonlinear template to 436 individual MRI from a publicly available dataset, including Asian, Black, and White participants. We quantified the gross brain dimensions and extracted the mean scaling, shear, and voxel-wise Jacobian determinants from the linear and nonlinear registrations between the template and each subject. Across all racial groups, the MNI152 brain exhibited substantially larger radii than the population means, with z-scores frequently exceeding 1.0. The linear scaling factors indicated consistent contraction of the template, and voxel-wise Jacobian fields revealed spatially heterogeneous deformations, demonstrating that the template differs from real brains in both size and shape. These findings suggest that the MNI152 template does not reflect the average morphology of contemporary population samples and that linear registration alone cannot resolve these discrepancies. Therefore, more robust and unbiased template-generation pipelines may be necessary for applications requiring anatomically accurate head models.

Introduction

Standardized brain templates are the foundation of neuroimaging and modeling. They provide a common spatial reference that enables comparisons across individuals, studies, and modalities and underpin practically all group-level analyses in MRI, source localization, and computational modeling. Among these templates, the Montreal Neurological Institute (MNI) series has become the *de facto* coordinate system for human-brain research.

The earliest MNI template (MNI305) was created in the 1990s by manually identifying anatomical landmarks in 305 young adult MRI scans and linearly registering them to approximate Talairach space. Because only linear transformations were used, interindividual differences in cortical folding, cranial proportions, and local curvature accumulated in the group average, producing a template that was systematically larger than any individual brain with a “virtual convolution” (Evans et al., 1993, 2021) due to the inherent limitations of early registration methods

The subsequent MNI152 (ICBM152) template improved the image quality by averaging 152 higher-resolution scans; however, it inherited the global morphology of the original MNI305 space (Mazziotta et al., 1995, 2001a, 2001b). The Colin27 template (Holmes et al., 1998) and nonlinear

MNI152 templates (Evans et al., 2012) provided sharper and more accurate anatomical detail but did not fundamentally alter the large-scale cranial geometry established by the earlier linear templates. Extended-coverage models, such as the New York Head (Huang et al., 2016) and the Big Field of View template (Kreilkamp et al., 2023), also preserved the overall MNI152 size and shape.

Prior work has shown that MNI152 is larger than the Talairach atlas and several race-specific templates, including Chinese, Indian, and Korean population averages (Lancaster et al., 2007; Tang et al., 2010; Lee et al., 2005). However, no study has directly compared MNI152 to a large set of individual subject scans across multiple racial groups, including Caucasians. Such a comparison is essential for determining whether the MNI152 brain truly represents an “average” human brain or whether its morphology reflects the artifacts of template construction.

In this study, we evaluated the morphology of the MNI152 nonlinear template relative to 436 individual MRI scans from a publicly available dataset. Using gross morphological measurements, linear affine transforms, and voxel-based morphometry (Ashburner & Friston, 2000), we assessed whether the MNI152 brain accurately reflects population averages or whether systematic deviations persist. Our findings revealed substantial and spatially nonuniform discrepancies, suggesting that reliance on MNI152 may introduce unintended biases in applications requiring anatomically realistic head models.

Methods

Dataset

Structural MR scans were obtained from the Dallas Lifespan Brain Study, a large publicly available dataset hosted on OpenNeuro (Park et al., 2025). The full dataset included 464 unique participant scans spanning a wide adult age range (21-89 years; mean 58.31 ±18.14). For the present analysis, we downloaded T1w volumes from participants who self-identified as Asian (n=14), Black/African-American (n=27), or White/Caucasian (n=395). Participants labeled as mixed race, unknown, or other races were excluded because of the limited sample size. For the analysis of MNI152, we used the MNI152 T1 scan with a 2 mm resolution provided as part of the FSL software package.

Brain Extraction

All MRI volumes were processed using FSL (version 6.0.7.17). Brain extraction was performed using FSL’s *bet* tool on each subject’s scan and on the MNI152 template. The segmented brains were visually inspected for their accuracy. Six subject brains were discarded because of failed brain segmentation.

Rigid Alignment to MNI Space

Each subject’s brain-extracted image was aligned with a 6-degree-of-freedom (rigid body) transformation (FSL, *flirt -6 dof*) using the extracted MNI152 brain as the reference image. This step ensured that subsequent morphological comparisons reflected true size and shape

differences, rather than head orientation or position. The aligned brains were visually inspected for accuracy of transformation. Twenty-nine brains were discarded due to a failed alignment (e.g., inverted brain).

Gross Morphological Measurements

To quantify large-scale anatomical differences between the MNI152 brain and individual brains, we computed three linear radii for each subject:

- X-radius: half the x-distance between the leftmost and rightmost brain voxels.
- Y-radius: half the y-distance between the most anterior and most posterior voxels.
- Z-radius: z-distance between the x-y plane (z=0) and the most superior voxel.

A custom Python script identified the relevant boundary voxels in each brain (subjects and MNI152), computed the corresponding distances, and converted these measurements from voxel coordinates to world space (mm). For each racial group, we calculated the mean and standard deviation of each radius and obtained the MNI152 brain's z-score relative to the group distribution.

Affine transform analysis

To quantify the global linear differences between the MNI152 template and individual brains, the MNI152 brain was linearly registered to each subject's brain-extracted image (FSL, *flirt*). The resulting 4×4 affine matrix was saved and decomposed (FSL, *avscale*) to extract its scaling and shear components. Scaling factors reflect axis-specific expansion or contraction, whereas shear terms capture non-orthogonal distortions. Shear factors were converted to degrees using the inverse tangent of the shear parameter, representing the angular deviation from the orthogonal axes ($\theta = arctan\,(\text{shear})$).

Nonlinear Deformation Analysis

Voxel-based morphometry was used to quantify voxel-wise shape differences between the MNI152 brain and individual brain anatomies. For each subject, the MNI152 brain was nonlinearly warped to the subject's extracted brain volume (FSL, *fnirt*) and the resulting 3D map of Jacobian determinants was saved. The Jacobian determinant ($|J|$) provides a local measure of volumetric change, indicating expansion ($|J| > 1$) or contraction ($|J| < 1$) relative to the reference image (Chung et al., 2001). For example, a $|J|$ of 0.9 indicates a local contraction of 10%. Therefore, the full 3D field of Jacobian determinants represents a voxel-wise map of differences with respect to the template (Freeborough & Fox, 1998; Gee & Bajcsy, 1999). When aggregated across subjects, these maps offer a quantitative assessment of how well a template captures the anatomical characteristics of a given population (Yang et al., 2020).

To compute the mean Jacobian determinant at each voxel, we stacked the 3D Jacobian maps for each racial group into 4D arrays and averaged across the 4th dimension. The resulting volume, $|J|_{mean}$, is the average deformation at each MNI152 voxel.

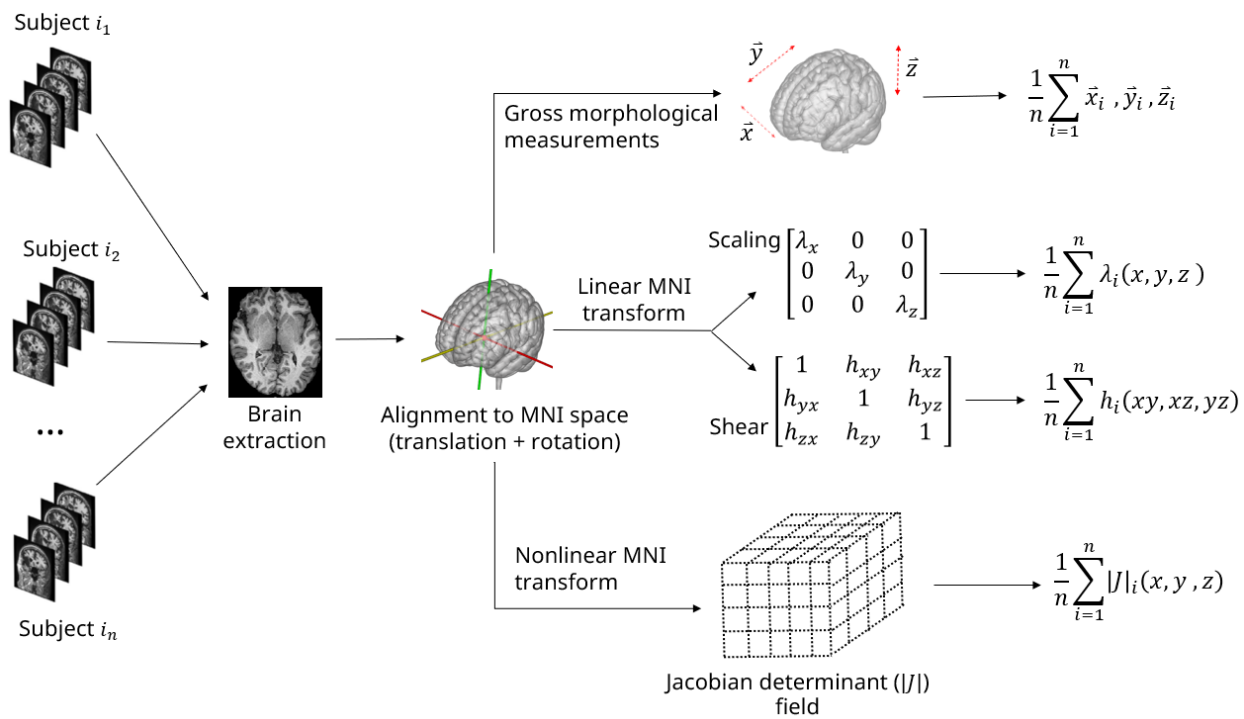

Figure 1: The brain was extracted from each subject's T1w volume (FSL, *bet*) and rigidly aligned to the MNI152 template brain with a six degrees of freedom transform (FSL, *flirt -6 dof*). Three analyses were then performed in parallel. The linear radii (X, Y, and Z) were computed from the aligned brain mask. Second, the MNI152 brain was linearly transformed to each brain and the resulting scaling and shear matrices were extracted (FSL, *avscale*). The MNI152 brain was also nonlinearly warped to each subject's brain (FSL, *fnirt*) and the resulting Jacobian determinant field was saved. The mean radii, scaling factors, shear factors, and mean voxel-wise Jacobian determinants were calculated for each racial group.

Results

Gross morphological measurements

Across all three racial groups, the MNI152 brain exhibited substantially larger linear dimensions than the corresponding population averages (Table 1). MNI152 z-scores were 0.24–1.65 for the X-radius, 1.42–2.32 for the Y-radius, and 0.49–1.65 for the Z-radius. The largest discrepancy occurred along the anterior–posterior (Y) axis, where the MNI152 brain had a length more than 5 mm greater than the mean of any racial group.

X-Radius		**Y-Radius**		**Z-Radius**	
MNI152	71 mm	MNI152	91 mm	MNI152	82 mm
Asian-American / Pacific Islander	70.27 ± 2.93 mm MNI z-score: 0.24	Asian-American / Pacific Islander	84.18 ± 2.92 mm MNI z-score: 2.32	Asian-American / Pacific Islander	80.55 ± 2.97 mm MNI z-score: 0.49
African-	66.43 ± 2.76 mm	African-	85 ± 3.35 mm	African-	77.04 ± 3.01 mm

American / Black	MNI z-score: 1.65	American / Black	MNI z-score: 1.78	American / Black	MNI z-score: 1.65
Caucasian / White	67.84 ± 2.61 mm MNI z-score: 1.21	Caucasian / White	85.72 ± 3.69 mm MNI z-score: 1.43	Caucasian / White	77.67 ± 3.42 mm MNI z-score: 1.26

Linear transform

Across all racial groups, mean scaling factors were consistently below 1, indicating that MNI152 must be contracted to match individual anatomy for most subjects. In contrast, mean shear components were very small (<0.04) corresponding to angular distortions of <2 degrees, demonstrating that shear contributes minimally to the linear mismatch between the template and real brains.

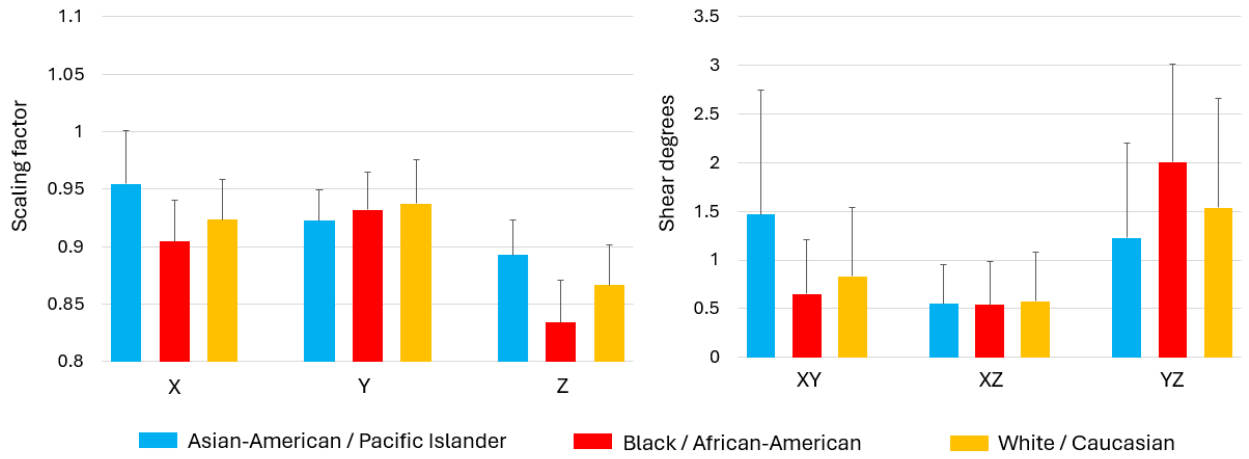

Figure 2: (A) Mean scaling factors for each racial group. A scaling factor <1 indicates compression and >1 indicates stretching. (B) Absolute shear factors for each subject were converted into degrees, and the mean degree tilt was then calculated for each racial group.

Nonlinear Deformation Analysis

Nonlinear deformation fields revealed that the MNI152 brain differs from individual brains not only in overall size but also in local shape. Importantly, contraction was not spatially uniform. Regions such as the temporal cortices showed strong contraction, whereas other regions, including deep white matter, showed systematic expansion. These patterns were consistent across racial groups (Figure 3).

Histograms of voxel-wise percent volume change (Figure 4) illustrate a broad distribution of deformation magnitudes, with long tails indicating localized expansions exceeding 20–30% and contractions exceeding 30–40%. Most voxels experienced systematic contraction confirming that MNI152 is generally larger than the population average. Mean volume reductions ranged from approximately 10% to 14% depending on racial group.

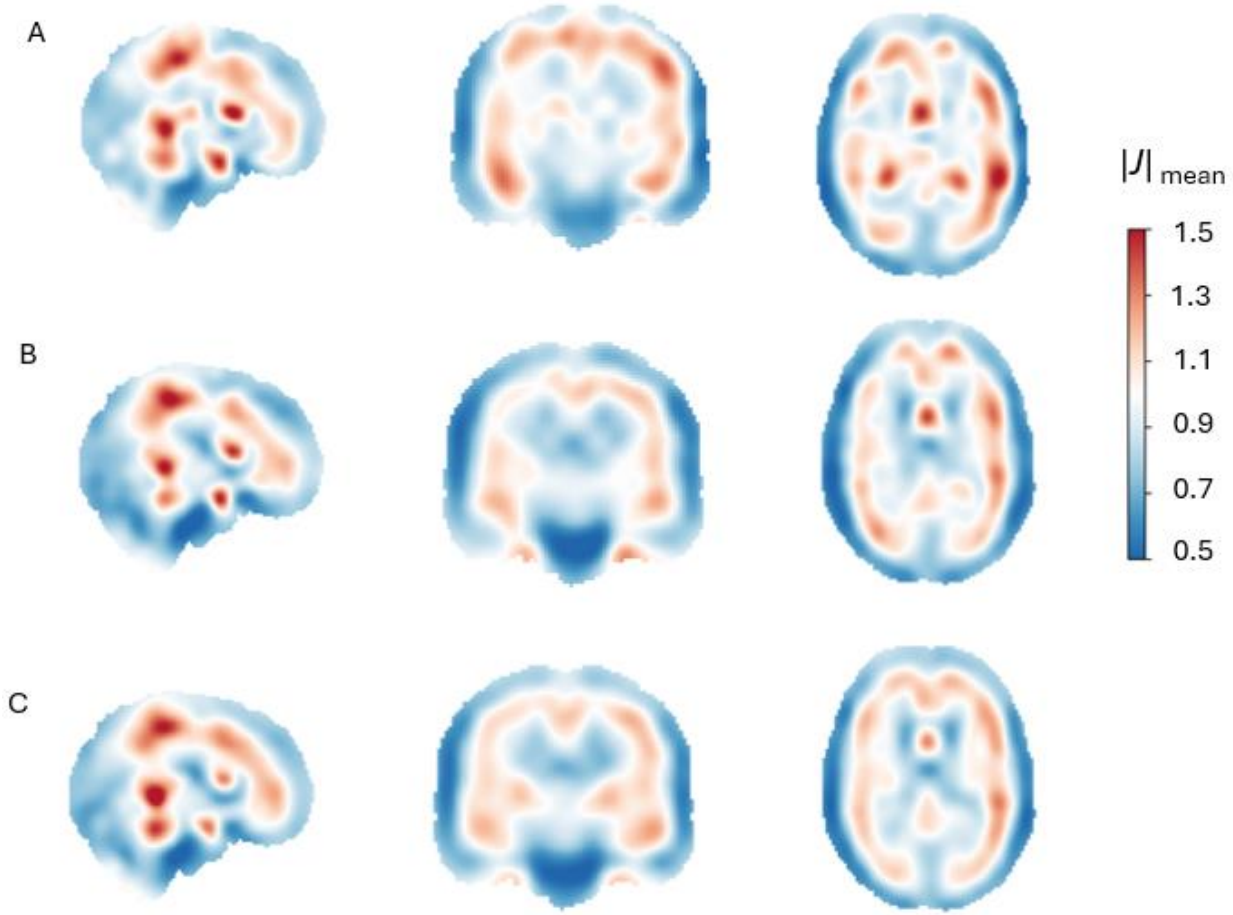

Figure 3: Mean Jacobian determinant fields over mid-sagittal, mid-coronal, and mid-axial slices showing systematic deformation between the MNI152 brain and individual brains for (A) Asian-American / Pacific Islander, (B) African-American / Black, and (C) Caucasian / White subjects. Values <1 indicate contraction of the MNI template, and values >1 indicate expansion. All groups show widespread contraction, demonstrating that the MNI152 brain is globally larger than real brains, with localized regions of expansion reflecting shape deformations.

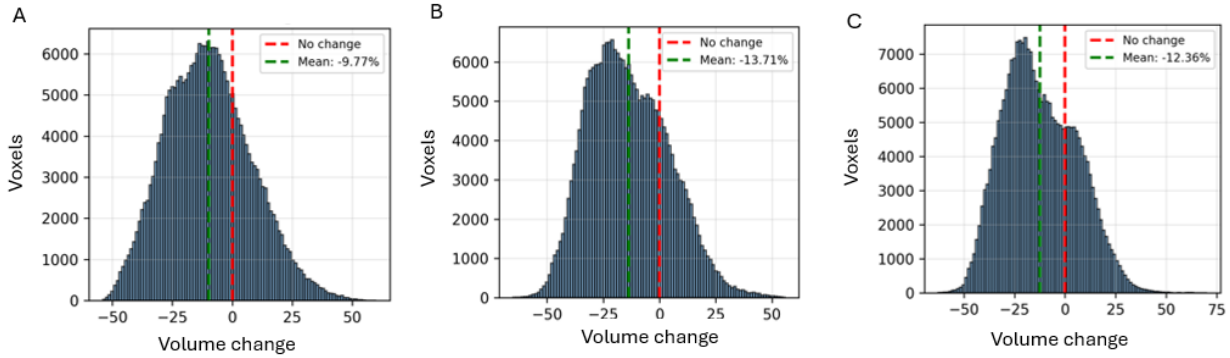

Figure 4: Distribution of voxel-wise percent volume change when deforming the MNI152 brain to subject-specific anatomy. Histograms show the spread of Jacobian-derived volume changes for (A) Asian-American / Pacific Islander, (B) African-American / Black, (C) Caucasian / White. The distributions are centered below zero, indicating systematic contraction of MNI152. Long tails reflect localized regions requiring large expansions or contractions, demonstrating that the mismatch between MNI152 and real brains is spatially heterogeneous.

Discussion

This study provides evidence that the widely used MNI152 template does not reflect the average size or shape of real human brains. Across more than 400 subjects spanning three racial groups, the MNI152 brain was consistently larger in all linear dimensions, with z-scores frequently exceeding 1.0. These findings extend prior work showing that the MNI152 template is larger than the Talairach atlas and larger than several race-specific templates and demonstrate that this discrepancy persists when compared directly to individual subject scans, including those of Caucasian subjects. Across all racial groups, the affine scaling factors were substantially below 1, indicating that the MNI152 template must usually be contracted along all three axes to approximate individual anatomy. However, the shear components were small, demonstrating that non-orthogonal distortions contribute minimally to template–subject alignment. The minimal shear also aligns with our previous work showing that certain MNI coordinates can be localized for TMS targeting with scaling alone, without requiring shear corrections (Valter et al., 2025).

The systematic volumetric bias we observed may contribute to previously reported inaccuracies when using MNI-based models, such as poor electrode-position optimization for interferential stimulation (Brahma et al., 2025).

More broadly, the results highlight the importance of using unbiased, population-representative template-generation pipelines, particularly for applications requiring anatomically realistic head models. Modern diffeomorphic averaging methods, race-specific templates, or individualized head models may provide more accurate alternatives.

References

Ashburner, J., & Friston, K. J. (2000). Voxel-based morphometry—the methods. Neuroimage, 11(6), 805-821.

Brahma, T., Guillen, A., Moreno, J., Datta, A., & Huang, Y. (2025). On the need of individually optimizing temporal interference stimulation of human brains due to inter-individual variability. *Brain stimulation*.

Evans, A. C., Collins, D. L., Mills, S. R., Brown, E. D., Kelly, R. L., & Peters, T. M. (1993, October). 3D statistical neuroanatomical models from 305 MRI volumes. In 1993 IEEE conference record nuclear science symposium and medical imaging conference (pp. 1813-1817). IEEE.

Evans, A. C., Janke, A. L., Collins, D. L., & Baillet, S. (2012). Brain templates and atlases. Neuroimage, 62(2), 911-922.

Freeborough, P. A., and Fox, N. C. 1998. Modelling brain deformations in Alzheimer disease by ¯uid registration of serial MR images. J. Comput. Assisted Tomogr. 22:838±843.

Gee, J. C., and Bajcsy, R. K. 1999. Brain Warping, Chap. 11, pp. 183±198. Academic Press, San Diego.

Holmes, C. J., Hoge, R., Collins, L., Woods, R., Toga, A. W., & Evans, A. C. (1998). Enhancement of MR images using registration for signal averaging. Journal of computer assisted tomography, 22(2), 324-333.

Huang, Y., Parra, L. C., & Haufe, S. (2016). The New York Head—A precise standardized volume conductor model for EEG source localization and tES targeting. NeuroImage, 140, 150-162.

Kreilkamp, B. A., Martin, P., Bender, B., la Fougère, C., van de Velden, D., Stier, C., ... & Focke, N. K. (2023). Big Field of View MRI T1w and FLAIR Template-NMRI225. Scientific Data, 10(1), 211.

Lancaster, J. L., Tordesillas-Gutiérrez, D., Martinez, M., Salinas, F., Evans, A., Zilles, K., ... & Fox, P. T. (2007). Bias between MNI and Talairach coordinates analyzed using the ICBM-152 brain template. Human brain mapping, 28(11), 1194-1205.

Lee JS, Lee DS, Kim J, Kim YK, Kang E, Kang H, Kang KW, Lee JM, Kim JJ, Park HJ, Kwon JS, Kim SI, Yoo TW, Chang KH, Lee MC. Development of Korean standard brain templates. J Korean Med Sci. 2005 Jun;20(3):483-8. doi: 10.3346/jkms.2005.20.3.483.

Mazziotta, J. C., Toga, A. W., Evans, A. C., Fox, P., & Lancaster, J. (1995).
A probabilistic atlas of the human brain: Theory and rationale for its development. NeuroImage, 2(2), 89–101.

Mazziotta, J. C., Toga, A. W., Evans, A. C., Fox, P. T., Lancaster, J., Zilles, K., Woods, R., Paus, T., Simpson, G., Pike, B., Holmes, C., Collins, D. L., Thompson, P., MacDonald, D., Iacoboni, M., Schormann, T., Amunts, K., Palomero-Gallagher, N., Geyer, S., Parsons, L., Narr, K., Kabani, N., LeGoualher, G., Boomsma, D., Cannon, T., Kawashima, R., & Mazoyer, B. (2001).
A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 356(1412), 1293–1322.

Mazziotta, J. C., Toga, A. W., Evans, A. C., Fox, P. T., Lancaster, J., Zilles, K., Woods, R., Paus, T., Simpson, G., Pike, B., Holmes, C. J., Collins, D. L., Thompson, P., MacDonald, D., Iacoboni, M., Schormann, T., Amunts, K., Palomero-Gallagher, N., Geyer, S., Parson, L., Narr, K., Kabani, N., LeGoualher, G., Boomsma, D., Cannon, T., Kawashima, R., & Mazoyer, B.; International Consortium for Brain Mapping. (2001). Four-dimensional probabilistic atlas of the human brain. Journal of the American Medical Informatics Association, 8(5), 401–430

Park, D.C. et al. The Dallas Lifespan Brain Study: A Comprehensive Adult Lifespan Data Set of Brain and Cognitive Aging. Sci Data 12, 846 (2025). https://doi.org/10.1038/s41597-025-04847-7

Tang, Y., Hojatkashani, C., Dinov, I. D., Sun, B., Fan, L., Lin, X., ... & Toga, A. W. (2010). The construction of a Chinese MRI brain atlas: a morphometric comparison study between Chinese and Caucasian cohorts. Neuroimage, 51(1), 33-41.

Valter, Y., Truong, D. Q., Kawasumi, Y., Bikson, M., & Datta, A. (2025). TMS Targeting Using MRI-free Head Modeling. Transcranial Magnetic Stimulation, 100188.

Yang G, Zhou S, Bozek J, et al. Sample sizes and population differences in brain template construction. NeuroImage. 2020;206:116318. doi:10.1016/j.neuroimage.2019.116318