ON-LINE APPENDIX: MATERIALS AND METHODS

Brain MR Imaging Analysis

Image Preprocessing. Analyses were performed using the open-access FSL.[1-3] The T1-weighted images were denoised and bias-field–corrected (FSL-SUSAN; https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/SUSAN). Wrap-around artifacts were verified not to overlay the skull and then were manually corrected. The FSL Brain Extraction Tool (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/BET) was then used to identify and remove voxels representing the skull and to produce a separate image representing the outer skull surface. The performance of the skull-stripping algorithm was improved by the use of nearest-neighbor interpolation and weighting (×1000) of the superior calvaria (parietal and frontal bones).[4] The images of the outer skull were used to coregister the pre- and postflight images, as described below.

Volumetric Analysis by Brain Tissue Segment. To perform ventricular volume analysis, we created hand-traced masks for the right lateral, left lateral, third, and fourth ventricles in Montreal Neurological Institute space (Montreal Neurological Institute-152 template).[5] Next the ventricular masks in Montreal Neurological Institute space were registered to the brain images in subject space using nonlinear registration (FMRIB Nonlinear Registration Tool, FNIRT; http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FNIRT), which allowed the estimation of ventricular volume for each mask.

Percentage change in volume of individual ventricles was then calculated as

$$\text{Volume \% Change} = \frac{\mathrm{V}_{\text{Aligned Post}} - \mathrm{V}_{\text{Pre}}}{\mathrm{V}_{\text{Pre}}} \times 100,$$

where V_{Pre} represents the preflight volume and $\mathrm{V}_{\text{Aligned Post}}$ represents the postflight volume. The percentage change in total ventricular volume was also calculated directly using the longitudinal SIENA algorithm of the FSL package (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/SIENA).[6]

Global and Local Structural Change Estimation. Pre- to postflight changes in brain morphology were characterized on the basis of the spatial displacement of voxels representing brain parenchyma. Brain parenchyma was defined as the sum of partial volume estimation (PVE) images for gray and white matter based on the FMRIB Automated Segmentation Tool (FAST; https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FAST) segmentation algorithm.[7] As in previous deformation-based morphometry studies,[8,9] volumetric displacement was estimated using linear scaling of large-scale structures and nonlinear warping of locally specific voxels. A focus on linear volumetric changes was motivated by previous observations of linear shifts in brain position relative to the skull for long-duration space flights (superior shift)[10] and a microgravity proxy method involving head-down bed rest (posterior shift).[11] The following methods were chosen to estimate these changes and avoid longitudinal measurement bias that can occur with asymmetric normalization of one image to another.

Global Structural Change

Pre- to postflight global structural changes of the brain were estimated for each astronaut with a 12-parameter affine-transformation matrix (3 parameters each for translation, rotation, skewing, and global scaling in x, y, and z).[1] The skull was used as a fixed anatomic reference between the pre- and postflight scans to account for differences in astronaut head positioning. To avoid any rescaling due to changes in soft tissue between the pre- and postflight scans (for example, the thickness of the subcutaneous fat layer in the scalp), which would result in an underestimation of the brain changes, we constrained the scaling factor to the outer skull surface.[12] A neuroradiologist (D.R.R.) reviewed the overlay of the skull on the T1-weighted scan for the pre- and postflight images as well as for the aligned postflight image. If any misalignment was observed, the skull-extraction algorithm was repeated for that subject, with heavier weighting toward the vertex of the skull.

The affine transforms for the postflight skull-specific alignment were applied to the postflight-summed PVE parenchyma images.[1] The PVE images from the preflight T1-weighted image (B_{PVE}^{pre}) were then affine-transformed to the skull-aligned postflight PVE images ($B_{PVE}^{post-aligned}$), which provided an estimate for the shift in brain position and global morphology during spaceflight based on the 12-parameter affine-transformation matrix ($B_{PVE}^{pretranformed}$).

Local Compression/Displacement of the Brain-CSF Interface

Local compression or displacement of brain parenchyma was assessed by first subtracting the affine-transformed preflight brain parenchyma image ($B_{PVE}^{pretranformed}$) from the skull-aligned postflight brain parenchyma image ($B_{PVE}^{postaligned}$) as follows:

$$\begin{aligned} \mathit{Local\ Change\ PVE\ Image} &= \Delta B_{local} \\ &= B_{PVE}^{postaligned} - B_{PVE}^{pretranformed}. \end{aligned}$$

Each difference image was then mapped to the Montreal Neurological Institute template by normalizing the preflight T1 image to the Montreal Neurological Institute template using FNIRT[13] and then applying the normalization parameters to the pre- and postflight difference image, which was subsequently Jacobian-modulated to estimate local volume change.

Regional Deformation of Brain Parenchyma

A tensor-based morphometry image analysis pipeline developed in-house was used to longitudinally investigate the local structural changes throughout the brain parenchyma in response to spaceflight. The pipeline included 2 separate nonlinear warpings. First, the preflight brain image was normalized to the postflight brain. The regional volumetric change, represented by the Jacobian determinant, denotes a difference between the 2 time points in the postflight native space. For statistically comparing across the astronauts, the mentioned Jacobian determinant is then normalized to a stereotaxic space (e.g. Montreal Neurological Institute-152 brain template). Thus, the second nonlinear warping is pooling the volumetric change across the astronauts.

Functional Task Test

The Activity Board Test measured crewmembers' ability to perform manual assembly and repair tasks, with the primary metric being time to complete all subtasks. Dynamic Postural Stability was assessed using 1 of the Sensory Organization Test conditions provided by EquiTest system platform (NeuroCom, Clackamas, Oregon). During testing, subjects were instructed to maintain stable upright posture for 20-second trials with feet positioned shoulder-width apart on a force plate, eyes closed, and arms folded across the chest. Trials were conducted with the head still or moving with a sway-referenced support surface in the anterior-posterior direction intended to disrupt somatosensory feedback.[14] The primary metric was a computed continuous equilibrium score based on the sway angle of the astronauts' estimated center of mass. The Recovery from Fall/Stand Test measured the ability to maintain postural control after standing up from a prone position. The primary metric was mean sway speed, centimeter $\times$ second^{-1}, defined as the average rate of the center of pressure displacement during standing on a force plate.

The Supine Egress and Walk Test measured the ability to rise from a supine position and walk while avoiding obstacles, with the primary metric being course completion times. The Pegboard Test was used to assess fine-motor control, and the primary performance metric was time to complete the entire board. The Seated Egress and Walk Test measured the ability to remove a harness, rise from a seated position, and walk while avoiding obstacles, with the primary metric being course-completion times. The Object Translation Test measured the ability to pick up and move objects such as tools from one location to another, with the primary metric being completion times. The Ladder Climb Test measured the amount of time for astronauts to climb 40 rungs of a ladder. The Hatch Opening Test measured the astronaut's ability to turn a hatch on a simulated spacecraft with the primary performance metric being peak force produced. The Jump Down Test evaluated the ability of crewmembers to jump down from a height of 30 cm onto a force plate and remain still in a standing position for 10 seconds. The primary performance metric was postural settling time defined as the time from touchdown until the shear ground-reaction force remained within 3 SDs of the corresponding ground-reaction force during quiet stance. Changes in dynamic balance control were assessed using the Tandem Walk Test. Astronauts walked in a heel-to-toe fashion at a self-selected speed for 10 steps per trial with their arms crossed on their chests and their eyes closed. The primary metric was the average percentage of correct steps.

REFERENCES

1. Jenkinson M, Bannister P, Brady M, et al. **Improved optimization for the robust and accurate linear registration and motion correction of brain images.** *Neuroimage* 2002;17:825–41 CrossRef Medline
2. Jenkinson M, Smith S. **A global optimisation method for robust affine registration of brain images.** *Med Image Anal* 2001;5:143–56 CrossRef Medline
3. Greve DN, Fischl B. **Accurate and robust brain image alignment using boundary-based registration.** *Neuromage* 2009;48:63–72 CrossRef Medline
4. Jenkinson M, Pechaud M, Smith S. **MR-based estimation of brain, skull and scalp surfaces.** *In: Proceedings of the Annual Meeting of the Organization for Human Brain Mapping*, Toronto, Ontario, Canada**.** June 12–16, 2005
5. Mazziotta J, Toga A, Evans A, et al. **A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM).** *Phil Trans R Soc Lond B* 2001;356:1293–1322. CrossRef Medline
6. Smith SM, Zhang Y, Jenkinson M, et al. **Accurate, robust, and automated longitudinal and cross-sectional brain change analysis.** *Neuroimage* 2002;17:479–89 CrossRef Medline
7. Zhang Y, Brady M, Smith S. **Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm.** *IEEE Trans Med Imaging* 2001;20:45–57 CrossRef Medline
8. Ashburner J, Hutton C, Frackowiak R, et al. **Identifying global anatomical differences: deformation-based morphometry.** *Human Brain Mapp* 1998;6:348–57
9. Eckert MA, Tenforde A, Galaburda AM, et al. **To modulate or not to modulate: differing results in uniquely shaped Williams syndrome brains.** *Neuroimage* 2006;32:1001–07 CrossRef Medline
10. Roberts DR, Albrecht MH, Collins HR, et al. **Effects of spaceflight on astronaut brain structure as indicated on MRI.** *N Engl J Med* 2017;377:1746–53 CrossRef Medline
11. Roberts DR, Zhu X, Tabesh A, et al. **Structural brain changes following long-term 6° head-down tilt bedrest as an analog for spaceflight.** *AJNR Am J Neuroradiol* 2015;36:2048–54 CrossRef Medline
12. Smith SM, De Stefano N, Jenkinson M, et al. **Normalized accurate measurement of longitudinal brain change.** *J Comput Assist Tomogr* 2001;25:466–75 CrossRef Medline
13. Jenkinson M, Beckmann CF, Behrens TE, et al. **FSL.** *Neuroimage* 2012;62:782–90 CrossRef Medline
14. Mulavara AP, Peters BT, Miller CA, et al. **Physiological and functional alterations after spaceflight and bed rest.** *Med Sci Sports Exerc* 2018;50:1961–80 CrossRef Medline
15. Benjamini Y, Krieger AM, Yekutieli D. **Adaptive linear step-up procedures that control the false discovery rate.** *Biometrika* 2006;93:491–507 CrossRef

On-line Table 1: Descriptive statistics on ventricular volume measurements, stratified by ISS versus Shuttle

	Shuttle ($n = 7$)				ISS ($n = 12$)				
	Preflight (Mean)	Postflight (Mean)	Percentage Change (Mean)	Within-Group *P* Value[a]	Preflight (Mean)	Postflight (Mean)	Percentage Change (Mean)	Within-Group *P* Value[a]	Between-Group *P* Value[b]
Gray matter	577.3 ± 45.6	573.0 ± 42.5	−0.67% ± 2.42%	.47	593.0 ± 42.2	590.9 ± 43.5	−0.34% ± 2.17%	.52	.64
White matter	579.1 ± 43.9	582.5 ± 39.5	0.65% ± 1.92%	.22	602.0 ± 63.2	601.0 ± 67.0	−0.23% ± 1.86%	.91	.47
Total ventricular volume percentage change			0.0% ± 1.70%	.58			10.70% ± 5.88%	<.001	<.001

[a] *P* values testing within-group change were obtained using the Wilcoxon signed rank test.

[b] *P* values testing between-group differences in percentage change and shift were obtained using the Wilcoxon rank sum test.

On-line Table 2: Descriptive statistics on functional task testing measurements, ISS (*n* = 7), and Shuttle (*n* = 1)

Subtest	Preflight (Mean)	Postflight (Mean)	Percentage Change (Mean)	Within-Group Effect Size	*P* Value[a]
Activity Board	51.4 ± 10.4	53.1 ± 8.4	4.6% ± 12.7%	0.34	.74
Dynamic Postural Stability	72.8 ± 7.3	32.7 ± 26.1	−56.7% ± 32.8%	−1.81	.008[b]
Recovery from Fall/Stand	7.4% ± 1.3	11.9 ± 3.0	60.0% ± 19.9%	2.32	.008[b]
Supine Egress and Walk	17.7 ± 2.1	23.1 ± 3.8	31.5% ± 23.2%	1.34	.02[b]
Pegboard Test	56.3 ± 6.6	62.5 ± 7.9	11.5% ± 12.1%	0.94	.04
Seated Egress and Walk	16.9 ± 1.9	21.9 ± 3.5	30.6% ± 22.8%	1.35	.02[b]
Object Translation	14.7 ± 1.7	19.4 ± 3.4	32.0% ± 17.7%	1.66	.008[b]
Ladder Climb	18.7 ± 4.5	20.5 ± 5.8	9.5% ± 14.4%	0.65	.25
Hatch Opening	698.2 ± 188.2	625.4 ± 178.0	−9.7% ± 15.8%	−0.68	.11
Jump Down	198.1 ± 38.2	321.0 ± 62.9	0.7% ± 0.6%	1.41	.008[b]
Tandem Walk	73.8 ± 15.1	43.0 ± 17.4	39.9% ± 24.6%	−1.54	.008[b]

[a] *P* values testing within-group change were obtained using the Wilcoxon signed rank test.

[b] *P* values that remained significant after correction for multiple comparisons with a false discovery rate of 5% using the method of Benjamini et al.[15]

On-line Table 3: Spearman correlation matrix representing correlations between percentage change in total ventricular volume and percentage changes in functional task testing metrics among 8 astronauts (7 ISS astronauts and 1 Shuttle astronaut)

Functional Task Subtests	Total Ventricular Volume (% Change)	Unadjusted *P* Value
Activity Board	−0.05	.91
Dynamic Postural Stability	−0.52[a]	.18
Recovery from Fall/Stand	0.74[a]	.04
Supine Egress and Walk	0.48	.23
Pegboard Test	−0.07	.87
Seated Egress and Walk	0.36	.39
Object Translation	0.05	.91
Ladder Climb	−0.14	.74
Hatch Opening	−0.64[a]	.00
Jump Down	0.14	.74
Tandem Walk	0.10	.82

[a] Correlations >.50. None of the correlations remained significant after correction for multiple comparisons.

On-line Table 4: Descriptive statistics on WinSCAT domain measurements, ISS only (*n* = 12)

Subtest	Preflight (Mean)	Postflight (Mean)	Percentage Change (Mean)	Within-Group Effect Size	*P* Value[a]
CDS					
Accuracy (% correct)	98.9 ± 0.7	98.1 ± 1.2	−0.8% ± 1.1%	−0.73	.02[b]
Reaction time (ms)	1127.2 ± 229.2	1019.6 ± 221.5	−9.1% ± 10.8%	−0.78	.021[b]
CPT					
Accuracy (% correct)	95.7 ± 3.2	96.8 ± 3.2	1.2% ± 2.3%	0.51	.10
Reaction time (ms)	544.1 ± 85.1	493.3 ± 88.6	−9.4% ± 8.2%	−1.12	.001[b]
MTH					
Accuracy (% correct)	88.3 ± 7.8	89.2 ± 8.5	1.3% ± 10.1%	0.09	.77
Reaction time (ms)	2545.4 ± 458.0	2299.6 ± 229.3	−7.0% ± 19.2%	−0.63	.06
MSP					
Accuracy (% correct)	97.6 ± 1.5	95.8 ± 4.7	−1.8% ± 5.3%	−0.34	.50
Reaction time (ms)	1374.7 ± 286.9	1232.5 ± 324.8	−9.7% ± 18.7%	−0.55	.13
CDD					
Accuracy (% correct)	97.3 ± 3.4	97.0 ± 3.6	−0.2% ± 4.4%	−0.08	.66
Reaction time (ms)	995.3 ± 162.8	956.4 ± 195.8	−3.3% ± 16.9%	−0.21	.57

Note:—MTH indicates Mathematical Processing; MSP, Delayed Matching to Sample; CDD, Code Substitution Delayed Recognition.

[a] *P* value testing within-group change was obtained using Wilcoxon signed rank test.

[b] *P* values that remained significant after correction for multiple comparisons with a false discovery rate of 5% using the method of Benjamini et al.[15]

On-line Table 5: Spearman correlation matrix representing correlations between percentage changes in total ventricular volume and percentage changes in cognitive function metrics among ISS astronauts ($n = 12$)

WinSCAT Subtest	Total Ventricular Volume (% Change)	Unadjusted *P* Value
CDS % change		
Accuracy	−0.60[a]	.05
Reaction time	−0.10	.75
CPT % change		
Accuracy	−0.01	.92
Reaction time	−0.62[a]	.03
MTH % change		
Accuracy	0.25	.44
Reaction time	−0.58[a]	.05
MSP % change		
Accuracy	−0.01	.97
Reaction time	0.12	.71
CDD % change		
Accuracy	−0.06	.85
Reaction time	0.04	.90

Note:—MTH indicates Mathematical Processing; MSP, Delayed Matching to Sample; CDD, Code Substitution Delayed Recognition.

[a] Correlations >.50. None of the correlations remained significant after correction for multiple comparisons.