AJNR

Normative Data for Brainstem Structures, the Midbrain-to-Pons Ratio, and the Magnetic Resonance Parkinsonism Index

S.T. Ruiz, R.V. Bakklund, A.K. Håberg and E.M. Berntsen

AJNR Am J Neuroradiol 2022, 43 (5) 707-714
doi: https://doi.org/10.3174/ajnr.A7485
http://www.ajnr.org/content/43/5/707

Normative Data for Brainstem Structures, the Midbrain-to-Pons Ratio, and the Magnetic Resonance Parkinsonism Index

S.T. Ruiz, R.V. Bakklund, A.K. Håberg, and E.M. Berntsen

ABSTRACT

BACKGROUND AND PURPOSE: Imaging biomarkers derived from different brainstem structures are suggested to differentiate among parkinsonian disorders, but clinical implementation requires normative data. The main objective was to establish high-quality, sex-specific data for relevant brainstem structures derived from MR imaging in healthy subjects from the general population in their sixth and seventh decades of life.

MATERIALS AND METHODS: 3D T1WI acquired on the same 1.5T scanner of 996 individuals (527 women) between 50 and 66 years of age from a prospective population study was used. The area of the midbrain and pons and the widths of the middle cerebellar peduncles and superior cerebellar peduncles were measured, from which the midbrain-to-pons ratio and Magnetic Resonance Parkinsonism Index [MRPI = (Pons Area / Midbrain Area) × (Middle Cerebellar Peduncles / Superior Cerebellar Peduncles)] were calculated. Sex differences in brainstem measures and correlations to age, height, weight, and body mass index were investigated.

RESULTS: Inter- and intrareliability for measuring the different brainstem structures showed good-to-excellent reliability (intraclass correlation coefficient = 0.785–0.988). There were significant sex differences for the pons area, width of the middle cerebellar peduncles and superior cerebellar peduncles, midbrain-to-pons ratio, and MRPI (all, $P < .001$; Cohen D = 0.44–0.98), but not for the midbrain area ($P = .985$). There were significant very weak–to-weak correlations between several of the brainstem measures and age, height, weight, and body mass index in both sexes. However, no systematic difference in distribution caused by these variables was found, and because age had the highest and most consistent correlations, age-/sex-specific percentiles for the brainstem measures were created.

CONCLUSIONS: We present high-quality, sex-specific data and age-/sex-specific percentiles for the mentioned brainstem measures. These normative data can be implemented in the neuroradiologic work-up of patients with suspected brainstem atrophy to avoid the risk of misdiagnosis.

ABBREVIATIONS: BMI = body mass index; MCP = middle cerebellar peduncle; M/P = midbrain-to-pons; MRPI = Magnetic Resonance Parkinsonism Index; MSA = multiple system atrophy; PSP = progressive supranuclear palsy; PD = Parkinson disease; SCP = superior cerebellar peduncle

Distinguishing the different parkinsonian disorders, such as progressive supranuclear palsy (PSP), multiple system atrophy (MSA) of the parkinsonian type, and Parkinson disease (PD) can be difficult due to their overlapping clinical presentations, especially in the early stages when the clinical presentations are ambiguous.[1-7] Differentiating these disorders is, however, highly relevant because PSP progresses more rapidly than PD,[8] and neither PSP nor MSA responds well to levodopa therapy in contrast to PD.[8,9] Reliable neuroradiologic biomarkers derived from standard MR imaging scans analyzed with conventional radiologic tools can be important for an early and correct diagnosis of these parkinsonian disorders.[1-3,7,10-12] However, for such biomarkers to be reliable, they must be accurate and have high external validity, preferably based on normative data.

Neuroimaging used to identify specific patterns of atrophy is included in the diagnostic criteria for some parkinsonian disorders.[9,13] In PSP, atrophy of the superior cerebellar peduncle

Received July 29, 2021; accepted after revision February 11, 2022.

From the Departments of Circulation and Medical Imaging (S.T.R., R.V.B., E.M.B.), Faculty of Medicine and Health Sciences, and Neuromedicine and Movement Sciences (A.K.H.), Norwegian University of Science and Technology, Trondheim, Norway; and Department of Radiology and Nuclear Medicine (A.K.H., E.M.B.), St. Olavs University Hospital, Trondheim, Norway.

S.T. Ruiz and R.V. Bakklund contributed equally to the work and share the first authorship.

Please address correspondence to Erik M. Berntsen, MD, PhD, EDiNR, EDiPNR, Department of Circulation and Medical Imaging, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology, Mailbox 8905, 7491 Trondheim, Norway; e-mail: erik.berntsen@ntnu.no; @BerntsenErikM

Indicates open access to non-subscribers at www.ajnr.org

Indicates article with online supplemental data.

http://dx.doi.org/10.3174/ajnr.A7485

(SCP) and midbrain is characteristic, the latter showing the well-known Hummingbird sign.[1-3,5,12,14-17] In MSA, atrophy of the putamen is seen, and in the cerebellar variant of the disease, atrophy of the middle cerebellar peduncle (MCP) is typical.[3,9,18] Both qualitative and quantitative analysis of the degree of midbrain atrophy, the midbrain-to-pons ratio (M/P ratio), and the Magnetic Resonance Parkinsonism Index [MRPI = (Pons Area / Midbrain Area) × (MCP / SCP)] have been used to investigate and distinguish patients with PSP, MSA, and PD from one another and from healthy controls.[1-7,10,12,16-25] These studies demonstrate the potential of using easily accessible quantitative brainstem biomarkers to differentiate among parkinsonian disorders and between healthy controls and patients.

Nevertheless, what is missing for effective clinical translation of these quantitative brainstem biomarkers to neuroradiologic practice is sex-specific normative data from a large number of healthy subjects from the general population in the appropriate decades of life when parkinsonian disorders first present. The largest study to date consists of 92 healthy controls, mainly in their 60s, not separated by sex and not otherwise specified.[25] The largest study to date providing data separated by sex consists of 85 healthy individuals (42 women) between 50 and 80 years of age who were recruited as part of 3 different studies and who all had normal neurologic examination findings and no history of neurologic diseases.[26] The latter study found significant sex differences only for the area of the pons and not for the M/P ratio or the MRPI, concluding that there is no need to consider age or sex when using these biomarkers to differentiate parkinsonian disorders. There is, however, rising awareness of sex differences in the brain and brain disorders, justifying further investigation into potential sex differences in neuroimaging biomarkers.[27,28]

The main goal of our study was to establish high-quality, sex-specific normative data for the midbrain and pons area, the width of the MCPs and SCPs, the M/P ratio, and the MRPI for use in clinical neuroradiology, facilitating a better diagnostic work-up of parkinsonian diseases. We also examined whether these brainstem measures were correlated to age, height, weight, and body mass index (BMI), which are reported to be correlated to brain size.[29,30]

MATERIALS AND METHODS

Population

Participants in our study are from the HUNT MRI study,[31] which is a part of the geographically defined prospective population study Nord-Trøndelag Health Study (HUNT),[32] in which MR imaging was performed between July 2007 and December 2009. Inclusion criteria were age between 50 and 65 years at the time of inclusion; participation in HUNT1 (1984–86), HUNT2 (1995–1997), and HUNT3 (2006–2008); and residency within 45 minutes from Levanger Hospital where scanning was performed. The only exclusion criteria were standard MR imaging contraindications. Seventy-three percent of the invited participants accepted and have been shown to be representative of the whole population.[31,33] Health information for each participant was obtained from questionnaires, blood samples, and a limited clinical examination. Participants with intracranial findings on MR imaging were contacted and underwent a clinical interview. [31] The present study excluded all subjects with known neurodegenerative disease based on their hospital records and MR imaging findings that could potentially confound our results. One subject with PSP, 3 with MS, 3 with pontine lacunar infarcts, 1 with a basilar dolichoectasia affecting the pons, 1 due to aberrant morphology, and 1 due to movement artifacts were excluded. In total, 996 subjects (99%) of the original 1006 from HUNT MRI were included in this study. The Regional Committee for Medical and Health Research Ethics in Central Norway has approved both the HUNT MRI study and this study (2011/456 and 2018/2231).

Measurements

Subjects' height and weight were measured in centimeters and kilograms to the first decimal using standardized methods in the HUNT3 study. Measurements of the different brainstem structures for this study were performed using the PACS software on aligned and saved images from a non-contrast-enhanced sagittal 3D T1WI IR-FSPGR volume (TR = 10.2 ms, TE = 4.1 ms, flip angle = 10°, section thickness = 1.2 mm, in-plane resolution = 0.975 × 0.975 mm^2) acquired on the same 1.5T scanner (Signa HDx; GE Healthcare) with an 8-channel head coil. The first authors performed measurements on 496 and 500 subjects each, after receiving extensive training by a board-certified neuroradiologist (E.M.B.), who also was consulted and assisted in difficult cases. The resulting measurements were used to calculate the M/P ratio and the MRPI.

Measurements of the midbrain and pons area were performed on midsagittal images according to the method of Oba et al.[1] The midsagittal plane was obtained and saved to the PACS using the MPR module of the software in which the center of the interpeduncular cistern was aligned with the center of the aqueduct in the transverse plane and along the falx in the coronal plane. After magnifying the saved midsagittal image 4 times, we drew a straight line between the superior pontine notch and the inferior point of the quadrigeminal plate (line A, Figure). Due to the ambiguous caudal outline of the quadrigeminal plate, parasagittal views were inspected before defining the exact location of this inferior point. A parallel line to line A was placed at the inferior pontine notch, thus defining the lower border of the pons (line B, Figure). Once these borders were defined, the raters manually traced the midbrain area ventral to the aqueduct above line A, while the area of the pons was manually traced ventral to the fourth ventricle and between lines A and B (Figure). The software automatically calculated the areas traced.

The widths of the MCPs were measured according to the method of Quattrone et al.[2] The previously defined midsagittal plane was used as a starting point to identify the parasagittal images in which the left and right MCPs were clearly surrounded by peripeduncular CSF and visible between the pons and cerebellum (Figure). The widths of the MCPs were then measured by drawing a straight line between the superior and inferior borders of the MCPs. The mean width of the 2 MCPs was calculated and used for further analysis.

The widths of the SCPs were also measured according to the method of Quattrone et al.[2] In the previously defined midsagittal plane, the axis defining the coronal plane was placed parallel to the rhomboid fossa, creating an oblique-coronal plane. Moving posteriorly in this oblique-coronal plane, we defined the first section where the inferior colliculi and SCPs were clearly separated

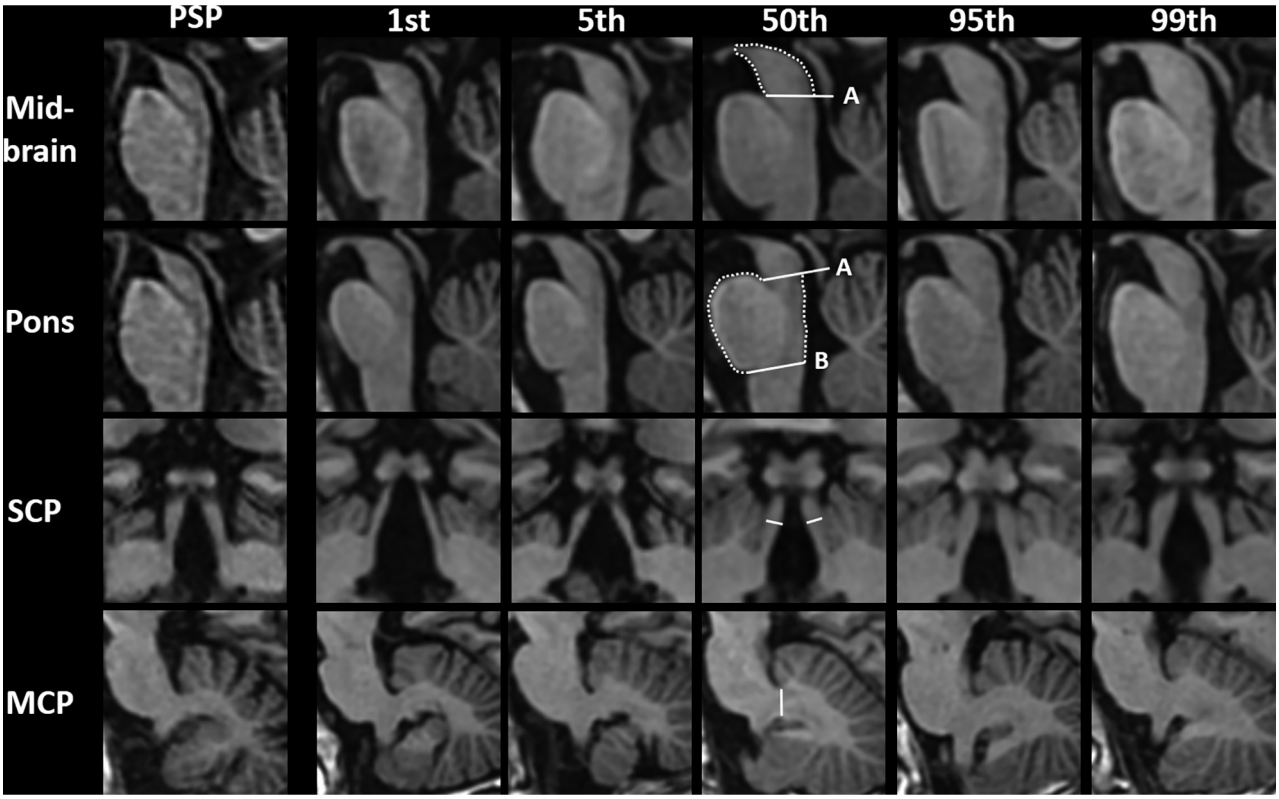

FIGURE. T1WI showing the 1st, 5th, 50th, 95th, and 99th percentiles for the midbrain and pons area, MCP and SCP widths for men as a whole group, as well as the same structures for the one male patient with verified PSP. Measurements of these structures are shown for the 50th percentile, which are also presented in high resolution in the Online Supplemental Data. The 50th percentile equals the mean. The line *A* is drawn between the superior pontine notch and the inferior point of the quadrigeminal plate, while line *B* is parallel to line *A* starting at the inferior pontine notch.

by CSF as the starting section and saved it. The SCPs were measured at their midpoint as the distance between the medial and lateral edges (Figure). Measurements were performed on both the left and right SCPs on the starting section and in the 2 following sections, resulting in a total of 3 saved images and 6 measurements. The mean value of these 6 measurements was calculated and used for further analysis.

Remeasurement

To investigate interrater reliability, each rater performed a second set of measurements on 50 random subjects from the other rater's pool of subjects. To investigate intrarater reliability, each rater performed a second set of measurements on 50 random subjects from their own pool of subjects. Remeasurement for inter-/intrarater reliability was performed twice. The first time, the 2 raters realigned and saved the new midsagittal, parasagittal, and oblique-coronal images on which remeasurement was performed. The second time, the remeasurement was performed on the original saved midsagittal, parasagittal, and oblique-coronal images where the original measurements had been performed. The remeasurement for inter-/intrarater reliability analysis was commenced and completed 4 weeks after the original measurements were finalized.

Statistical Analysis

SPSS statistical and computing software (IBM) was used for statistical analysis. The mean (SD) and range for the midbrain and pons areas, the width of the MCPs and SCPs, the M/P ratio, and the MRPI were calculated for all 996 subjects and for each sex separately. A statistically significant difference of the mean of the brainstem measurements between the sexes was defined as $P < .05$ with a 2-tailed independent-samples *t* test. The effect size of significant sex differences was investigated using the Cohen D, in which values <0.50 were interpreted as small; values between 0.50 and 0.80, as medium; and values >0.80, as large.[34] Correlations among age, sex, height, weight or BMI, and brainstem measurements were investigated using the Pearson correlation, and $P < .05$ was considered significant. A correlation coefficient of 0.19 was considered as very weak; 0.20–0.39, weak; 0.40–0.59, moderate; 0.60–0.79, strong; and 0.8–1.0, very strong.[35] Inter-/intrarater reliability was calculated using intraclass correlation coefficients using a 2-way mixed-effects model with absolute agreement and single measures. Values between 0.75 and 0.90 were interpreted as good, and values >0.90 were interpreted as excellent.[36]

RESULTS

In total, 527 women (mean age, 58.7 [SD, 4.3] years; mean height, 165.2 [SD, 5.8] cm; mean weight, 72.6 [SD, 11.8] kg; mean BMI, 26.6 (SD, 4.1] kg/m^2) and 469 men (mean age, 59.2 [SD, 4.2] years; mean height, 178.0 [SD, 6.1] cm; mean weight, 86.8 [SD, 11.1] kg; mean BMI, 27.4 [SD, 3.1] kg/m^2) were included in the present study. The descriptives for the brainstem measures are presented in Table 1. There were statistically significant differences between the sexes for the pons area with a large effect size,

Table 1: Descriptives for all subjects and for each sex[a]

	All Subjects ($n = 996$)	Women ($n = 527$)	Men ($n = 469$)	P Value[b]	Cohen D
Age (yr)	59.0 (SD, 4.2) (50.5–66.8)	58.7 (SD, 4.3) (50.5–66.3)	59.2 (SD, 4.2) (51.0–66.8)	$P = .103$	
Height[c] (cm)	171.2 (SD, 8.7) (148.4–196.7)	165.2 (SD, 5.8) (148.4–186.7)	178.0 (SD, 6.1) (161.9–196.7)	$P < .001$	2.17
Weight[c] (kg)	79.3 (SD, 13.5) (47.2–127.0)	72.6 (SD, 11.8) (47.2–119.6)	86.8 (SD, 11.1) (59.8–127.0)	$P < .001$	1.24
BMI[c]	27.0 (SD, 3.7) (18.3–41.7)	26.6 (SD, 4.1) (18.3–41.7)	27.4 (SD, 3.1) (19.6–40.1)	$P < .001$	0.21
Midbrain (mm^2)	136.7 (SD, 20.5) (76.0–222.4)	136.7 (SD, 19.7) (81.9–211.9)	136.7 (SD, 21.4) (76.0–211.9)	$P = .985$	
Pons (mm^2)	542.2 (SD, 57.8) (386.5–812.5)	518.2 (SD, 47.8) (386.5–695.7)	569.1 (SD, 56.2) (416.5–812.5)	$P < .001$	0.98
MCP (mm)	9.4 (SD, 0.8) (5.3–13.0)	9.1 (SD, 0.8) (5.3–11.5)	9.7 (SD, 0.8) (6.8–13.0)	$P < .001$	0.76
SCP (mm)	3.8 (SD, 0.5) (2.5–5.5)	3.7 (SD, 0.5) (2.6–5.2)	3.9 (SD, 0.5) (2.5–5.5)	$P < .001$	0.44
M/P ratio	0.25 (SD, 0.04) (0.14–0.41)	0.26 (SD, 0.04) (0.15–0.41)	0.24 (SD, 0.03) (0.14–0.36)	$P < .001$	0.67
MRPI	10.0 (SD, 1.9) (5.5–18.1)	9.5 (SD, 1.7) (5.5–15.6)	10.6 (SD, 2.0) (6.1–18.1)	$P < .001$	0.59

[a] Data are mean (SD) and range.
[b] Women versus men.
[c] Missing data for 2 women and 1 man.

the width of the MCP with a medium effect size, the SCP with a small effect size, as well as the M/P ratio and MRPI both with moderate effects sizes (all, $P < .001$; Cohen D = 0.44–0.98), but none for the midbrain area. Scatterplots for each measured brainstem structure by age, height, weight, and BMI are presented for each sex in the Online Supplemental Data.

Correlations between the brainstem measurements and age, height, weight, and BMI were investigated in each sex separately. The significant correlations were as follows: For women, there were weak age-related correlations for the midbrain area and the M/P ratio [both $r(525) = -0.28$ ($P < .001$)] and very weak correlations for the width of the MCP [$r(525) = -0.18$ ($P < .001$)] and the MRPI [($r(525) = 0.15$ ($P < .001$)]. Height was weakly correlated to the pons area [$r(523) = 0.20$ ($P < .001$)] and very weakly to the width of the MCP [$r(523) = 0.12$ ($P = .009$)] and the midbrain area [$r(523) = 0.11$ ($P = .014$)]. Weight was very weakly correlated to the pons area [$r(523) = 0.09$ ($P = .049$)] and the M/P ratio [$r(523) = -0.11$ ($P = .009$)]. BMI was very weakly correlated to the midbrain area [$r(523) = -0.10$ ($P = .027$)] and the M/P ratio [$r(523) = -0.11$ ($P = .013$)]. For men, there were weak age-related correlations for the midbrain area [$r(467) = -0.23$ ($P < .001$)] and the M/P ratio [$r(467) = -0.25$ ($P < .001$)] and a very weak correlation for MRPI [$r(467) = 0.16$ ($P < .001$)]. Height was weakly correlated to the midbrain area [$r(466) = 0.22$ ($P < .001$)] and the pons area [$r(466) = 0.21$ ($P < .001$)] and very weakly correlated to the width of the SCP [$r(466) = 0.11$ ($P = .018$)], the M/P ratio [$r(466) = 0.09$ ($P = .043$)], and the MRPI [$r(466) = -0.12$ ($P = .008$)]. Weight was very weakly correlated to the pons area [$r(466) = 0.10$ ($P = .041$)]. BMI was very weakly correlated to the midbrain area and the M/P ratio [both $r(466) = -0.14$ ($P = .002$)] as well as the MRPI [$r(466) = 0.13$ ($P = .006$)].

The significant correlations were in the same direction and magnitude for both sexes, and no systematic difference between the sexes caused by these variables could be seen in the scatterplots. Age was the variable with the highest and most consistent correlation with the brainstem measures. The normative percentile data are reported for 5-year age groups for men and women separately (Table 2).

Intraclass correlation coefficient values for both inter-/intrarater reliability showed good-to-excellent reliability for all measurements based on the limits defined by Koo and Li.[36] When remeasurements were based on realigned and saved images, the intraclass correlation coefficient values ranged from 0.785 to 0.978. When remeasurements were based on the original saved images, intraclass correlation coefficient values ranged from 0.891 to 0.988 (Online Supplemental Data).

DISCUSSION

We present sex-specific normative data for the midbrain and pons areas, the widths of the MCPs and SCPs, the M/P ratio, and the MRPI based on representative general population data from individuals between 50 and 66 years of age.[31,33] These are the decades of life when the parkinsonian disorders have their onset.[8,13,37] Our measurements have mostly excellent reliability when measured both on the original and realigned, saved images, ensuring their usefulness in a clinical setting. There were highly significant sex differences for all brainstem measures except for the midbrain area. There were significant very weak–to-weak correlations of midbrain area, the M/P ratio, and the MRPI with age for both sexes. There were also some significant very weak–to-weak correlations for some of the brainstem measures and height, weight, and BMI, which are in line with findings in previous studies.[29,30] However, because no systematic differences between the sexes caused by these variables were found, age-/sex-specific percentiles were created. These percentiles are easy to use in the clinic; nevertheless, height did have some very weak–to-weak correlations to some of the brainstem measures, which may be considered on an individual basis.

Table 2: Sex-specific percentiles in the different age groups

	Women (n = 527)					Men (n = 469)				
	1st	5th	50th	95th	99th	1st	5th	50th	95th	99th
Midbrain area										
50–54 yr (n = 121/91)	93	111	142	178	207	104	113	141	183	222
55–59 yr (n = 187/159)	101	110	137	174	196	95	107	140	176	195
60–66 yr (n = 219/219)	88	102	130	166	178	88	101	130	166	184
Pons area										
50–54 yr (n = 121/91)	405	431	517	591	616	417	478	565	670	703
55–59 yr (n = 187/159)	414	456	517	608	642	458	480	567	678	785
60-66 yr (n = 219/219)	413	436	516	610	664	443	476	565	659	689
MCP width										
50–54 yr (n = 121/91)	7.9	8.2	9.4	10.6	11.3	8.0	8.2	9.7	11.3	11.9
55–59 yr (n = 187/159)	7.7	8.1	9.2	10.4	10.8	7.5	8.6	9.8	11.2	12.2
60–66 yr (n = 219/219)	7.2	7.6	9.0	10.5	10.9	7.7	8.4	9.7	11.0	11.7
SCP width										
50–54 yr (n = 121/91)	2.6	3.1	3.7	4.5	5.1	2.5	3.2	3.9	4.7	5.5
55–59 yr (n = 187/159)	2.8	3.1	3.8	4.5	4.9	2.8	3.1	4.0	4.7	4.8
60–66 yr (n = 219/219)	2.7	2.9	3.7	4.5	4.9	2.9	3.2	4.0	4.8	5.2
M/P ratio										
50–54 yr (n = 121/91)	0.21	0.22	0.28	0.34	0.38	0.18	0.21	0.25	0.32	0.36
55–59 yr (n = 187/159)	0.20	0.21	0.27	0.32	0.37	0.16	0.18	0.25	0.30	0.32
60–66 yr (n = 219/219)	0.17	0.19	0.26	0.32	0.33	0.17	0.18	0.23	0.29	0.30
MRPI										
50–54 yr (n = 121/91)	6.2	6.7	9.0	12.5	13.6	6.1	7.6	9.4	13.6	16.5
55–59 yr (n = 187/159)	5.6	7.2	9.2	12.2	14.0	7.3	7.8	10.1	14.8	17.1
60–66 yr (n = 219/219)	6.7	7.1	9.5	13.2	15.3	6.8	8.0	10.6	14.5	17.1

Note:—n = X/Y: X women and Y men.

This is the first study to present sex-specific percentiles for these imaging biomarkers in different age groups. Previous studies have had smaller cohorts not suitable for estimating percentiles. Our study is 11 times larger than the second largest,[25] enabling us to precisely estimate and illustrate the normal variance in both men and women. In the following discussion, we compare our results with those in studies with cohorts of >20 healthy individuals from convenience or not-specified populations in which the midbrain and pons areas and the widths of the MCPs and SCPs have been measured according to Oba et al[1] and Quattrone et al.[2] The data from these studies are also presented in the Online Supplemental Data.[1,2,4-7,15,16,19,20,25,26]

Brainstem Structures

We found a mean midbrain area of 137mm^2 for both sexes with no significant sex difference but with a significant weak age-correlated atrophy for both sexes. Previous studies have shown that the midbrain area is smaller in patients with PSP compared with healthy controls.[1,2,5-7,15,19,20] Previous estimates of the midbrain area in healthy subjects ranged from 118 to 142 mm^2, with most of these studies reporting a smaller midbrain area than we do.[1,2,5-7,15,16,19,20,25,26] The aforementioned studies have subjects with a mean age that is 7–12 years older than our cohort's mean age. This difference could explain their reported smaller midbrain area, considering that our study and others show significant weak age-related atrophy of midbrain area.[13,31]

We found a mean pons area for all subjects of 542 mm^2, with a significantly different pons area of 518 mm^2 for women and 569 mm^2 for men with a large effect size (Cohen D = 0.98), but no significant age correlation for either sex. Our results are slightly higher than previous reported values for the pons area for both sexes combined, which ranged from 469 to 541 mm^2.[1,2,5-7,15,19,20,25,26] The 2 articles separating the sexes also found a significantly smaller pons area in women,[1,26] while the 2 articles investigating age did not find any significant correlation.[15,26]

We found a mean SCP width for all subjects of 3.8 mm, with significantly different SCP widths of 3.7 mm for women and 3.9 mm for men, with a low effect size (Cohen D = 0.44) and no significant age correlation for either sex. Earlier estimates of SCP width ranged from 3.5 to 3.9 mm for both sexes combined, which is very similar to our results.[2,5-7,15,19,25,26] The only previous study separating the sexes found nonsignificant differences (P = .121) of 3.7 mm for women and 3.8 mm for men,[26] probably due to their lower sample size and lower statistical power, while we show a small but significant difference between the sexes with a small effect size. Neither of the 2 studies investigating age effects found any significant correlation between age and SCP width, corresponding to our findings.[15,26]

We found a mean MCP width for all subjects of 9.4 mm, with significantly different MCP widths of 9.1 mm for women and 9.7 mm for men, with a medium effect size (Cohen D = 0.76) and a significant age correlation for women but not men. Previous studies have found MCP widths from 8.6 to 10.0 mm,[2,5-7,15,19,25,26] with the only study separating the sexes finding a nonsignificant difference (P = .345) with 9.9 mm for women and 10.1 mm for men.[26] Again, the sex difference uncovered by our study is probably due to the larger sample size of our study. None of the 2 studies investigating age correlation found any significant correlation to age for MCP width,[15,26] while we found a significant correlation for women but not men. As far as we know, this sex-specific, age-related MCP width atrophy in women has not been reported before, but sex-specific and age-related atrophy is described in different regions of the brain.[38]

M/P Ratio

We found a mean M/P ratio for all subjects of 0.25, with significantly different M/P ratios of 0.26 for women and 0.24 for men, with a medium effect size (Cohen D = 0.67) and a significant age correlation for both sexes. These results are in line with previous estimates of the M/P ratio in healthy subjects, which ranges from 0.23 to 0.27 for both sexes combined (studies reporting the P/M ratio were converted as follows: M/P ratio = [P/M ratio]$^{-1}$).[1,2,4-7,19,20,26] The only other study separating the sexes found an almost significant difference ($P = .052$) of 0.23 for women and 0.22 for men,[26] which resembles our results. This discrepancy is probably due to their lower sample size. The same study did not find a significant age correlation for the M/P ratio ($P = .109$); however, Morelli et al[15] did report a significant correlation between age and the M/P ratio for the sexes combined ($P < .001$). Given our large cohort and previous literature indicating a sex difference in the pons area, we believe the presence of a significant sex difference in the pons area with a large effect size also indicates both age-/sex-specific differences in the M/P ratio, which need to be considered when used for clinical purposes.

MRPI

We found a mean MRPI of 10.0 for all subjects, with a significantly different MRPI of 9.5 for women and 10.6 for men, giving a medium effect size (Cohen D = 0.59). There was also a significant positive age correlation for both sexes. Most previous estimates of the MRPI ranged from 9.1 to 10.5, concurring with our findings.[2,4-7,19,25] Mangesius et al[26] did, however, find an MRPI for all subjects of 12.2, with a nonsignificantly different MRPI of 12.0 for women and 12.3 for men ($P = .362$). They also reported lower values of the midbrain area and larger values of the MCP width than in our study and the literature in general, both resulting in an increase in the MRPI. The reason for this discrepancy in the midbrain area and MCP width measurements is unclear because they also used a 3D T1 MPRAGE sequence and performed their measurements the same way that we did. Considering that we have shown age-related atrophy of the midbrain, one explanation for their lower midbrain area could be that their population was on average 7 years older than ours. We also found a significant age correlation for MRPI in both sexes in our study, which was not shown in the 2 previous studies addressing this subject.[15,26] Nevertheless, we do believe there are true age-/sex-specific differences, which should be addressed in the clinical and neuroradiologic settings through differentiated normal values for both age and sex, as we have presented in Table 2.

Which Biomarker to Use?

The Hummingbird sign has been shown to have high specificity (99.5%) but a rather low sensitivity (51.6%) when it comes to identifying patients with PSP.[14] The general consensus in the literature is that the MRPI has a higher diagnostic accuracy when distinguishing patients with PSP from healthy controls than the midbrain area,[1-3,12,17,22,23] SCP width,[2] and the M/P ratio.[1,6,16] One reason for the superiority of the MRPI is that the values of the midbrain and pons area and the MCP and SCP widths overlap between patients with PSP and healthy subjects, giving a low sensitivity and specificity.[1-3,16,19,20,23] The usefulness of the M/P ratio is more uncertain. Some studies have found it useful to differentiate between patients with PSP and healthy controls,[1,20] while others have not.[2,6,15,16,19] Different studies have estimated cutoff values for MRPI that separate patients with PSP from healthy subjects with a sensitivity and specificity of 100%, ranging from 13.2 to 13.6 combined for both sexes.[2,4,6,7,25] This cutoff corresponds well to our 95th percentile for MRPI for both sexes combined but becomes a bit more problematic when separated by sex, as discussed below.

Are Sex-Specific Norms Necessary?

Modern clinical medicine is slowly moving toward precision medicine, which includes sex-specific health care as there are important biologic sex differences requiring awareness in all aspects of medicine, including diagnostics.[39] Such sex differences have also been shown in patients with PSP.[40] We found significant sex-specific differences for all measurements except for midbrain area, and most notably in the MRPI, which is the most acknowledged biomarker for evaluating PSP. Currently, cutoff values are determined for entire cohorts and not for women and men separately. The given MRPI cutoff value of 13.2 to 13.6 in the literature corresponds well to the 95th percentile for our whole group, but when separated by sex, it corresponds to somewhere between the 95th and 99th percentile for women for all age groups and below the 95th percentile for men in all age groups. Thus, if the previous suggested cutoff values from the literature are applied, some women with PSP will risk not being diagnosed using the MRPI, while some men well within the normal variation will be misdiagnosed with PSP using the MRPI. This issue clearly shows the need for high-quality, sex-specific normative data when evaluating parkinsonian neurodegenerative diseases, which we provide in Table 2. Furthermore, considering that the sex difference could potentially increase the sensitivity and specificity for some of the other structures or ratios, it could perhaps make them as useful as the MRPI. Although the MRPI is considered the superior biomarker, it is still quite time-consuming in the clinical workflow, and measuring only the midbrain area or M/P ratio is far less time-consuming.

Strengths and Limitations

The predominant strength of our study is the large sample size, consisting of 996 subjects from the general population, imaged on the same scanner, with the same software, with a standardized MR imaging protocol and the same T1WI volume used in the Alzheimer Disease Neuroimaging Initiative, making our findings generalizable for everyone using this sequence.[41] This feature is an advantage of our study because previous studies have large variations in acquisition parameters and section thickness, potentially causing variations in the measurements. To the best of our knowledge, this is the largest study on brainstem structures in healthy subjects, giving it a considerable statistical power and making it less prone to random variation. This size enables us to find significant differences and calculate percentiles for several age groups. Moreover, the 996 individuals included in our study were between 50 and 66 years of age, encompassing the age of

onset for several of the parkinsonian disorders better than previous studies.[8,13,37]

An additional strength of our study is that we have manually measured the structures instead of using an automated approach. Currently, the manual approach is the preferred method and is considered the criterion standard. Given that our normative data are obtained manually and our inter-/intrarater intraclass correlation coefficients showed good-to-excellent reliability, we believe our data and percentiles are of high quality and easy to use for clinicians and neuroradiologists. Furthermore, these measurements are easy to perform and not dependent on additional implementation or steps in the workflow such as automated atlas-based approaches, which are not accessible to all. The future is, however, computer-aided, with either atlas-based volumetric evaluation or artificial intelligence approaches, as shown in several studies.[42,43] One multicenter study combining patients with PD ($n = 204$) and PSP ($n = 106$), MSA ($n = 81$), and healthy controls ($n = 73$) showed that fully automated atlas-based volumetry with subsequent support vector machine classification could differentiate the different syndromes on an individual level with sensitivities from 79% to 87% and specificities from 87% to 96%.[42] In their study, the midbrain showed the most atrophy in PSP and the MCP in MSA of the cerebellar type.

A limitation of our study is that the participants included did not undergo a thorough neurologic examination at the time of the MR imaging study, which potentially could have identified subjects with prodromal disease. However, given our large sample size, we believe the possible inclusion of these subjects is less likely to substantially skew our results.

A second limitation is that our normative data do not reach beyond 66 years of age, which is in the lower range of when some of the parkinsonian diseases typically have their onset. However, on an individual level, many patients have earlier onset or prodromal phases in which the need for high-quality normative data is also important. Whether there is a need for even more specified or tailored percentiles based on additional demographic and/or health variables remains to be determined and may become standard as big data and artificial intelligence become more integrated in the clinic. For this data set, however, we limited the percentile to sex and age because height, weight, and BMI were only weakly and inconsistently related to the brainstem measures.

Another limitation is that the 2 raters who performed the measurements were not experienced neuroradiologists. In comparable studies, neuroradiologists or experienced raters have independently performed measurements. Nevertheless, the 2 raters in our study were given extensive training by an experienced neuroradiologist, and we consider this to be the best approach when performing manual measurements in such large studies. Despite these limitations, the 2 raters in our study achieved inter-/intrarater reliability ranging from good to excellent. Comparable studies have not elaborated on the implementation of their reliability analysis. For example, they have not presented 95% CIs for their intraclass correlation coefficient values or mentioned whether the intraclass correlation coefficients were calculated on the basis of single or average measures or whether they had been measured on originally saved images or on realigned, saved images. Thus, comparing our intraclass correlation coefficient values with theirs is difficult. As expected, the intraclass correlation coefficient values in our study decreased when performing measurements on realigned, saved images, as several additional steps of the procedure were added.

CONCLUSIONS

We present high-quality sex-/age-specific normative data as means and percentiles for the midbrain and pons area, the widths of the MCPs and SCPs, the M/P ratio, and the MRPI based on manual measurements in 996 healthy subjects between 50 and 66 years of age. Furthermore, we show that using sex-specific data is important to avoid misdiagnosing patients with suspected brainstem atrophy.

Disclosure forms provided by the authors are available with the full text and PDF of this article at www.ajnr.org.

REFERENCES

1. Oba H, Yagishita A, Terada H, et al. **New and reliable MRI diagnosis for progressive supranuclear palsy.** *Neurology* 2005;64:2050–55 CrossRef Medline
2. Quattrone A, Nicoletti G, Messina D, et al. **MR imaging index for differentiation of progressive supranuclear palsy from Parkinson disease and the Parkinson variant of multiple system atrophy.** *Radiology* 2008;246:214–21 CrossRef Medline
3. Chougar L, Pyatigorskaya N, Degos B, et al. **The role of magnetic resonance imaging for the diagnosis of atypical parkinsonism.** *Front Neurol* 2020;11:665 CrossRef Medline
4. Quattrone A, Morelli M, Nigro S, et al. **A new MR imaging index for differentiation of progressive supranuclear palsy-parkinsonism from Parkinson's disease.** *Parkinsonism Relat Disord* 2018;54:3–8 CrossRef Medline
5. Longoni G, Agosta F, Kostić VS, et al. **MRI measurements of brainstem structures in patients with Richardson's syndrome, progressive supranuclear palsy-parkinsonism, and Parkinson's disease.** *Mov Disord* 2011;26:247–55 CrossRef Medline
6. Morelli M, Arabia G, Salsone M, et al. **Accuracy of magnetic resonance parkinsonism index for differentiation of progressive supranuclear palsy from probable or possible Parkinson disease.** *Mov Disord* 2011;26:527–33 CrossRef Medline
7. Nigro S, Morelli M, Arabia G, et al. **Magnetic resonance parkinsonism index and midbrain to pons ratio: which index better distinguishes progressive supranuclear palsy patients with a low degree of diagnostic certainty from patients with Parkinson disease?** *Parkinsonism Relat Disord* 2017;41:31–36 CrossRef Medline
8. Testa D, Monza D, Ferrarini M, et al. **Comparison of natural histories of progressive supranuclear palsy and multiple system atrophy.** *Neurol Sci* 2001;22:247–51 CrossRef Medline
9. Gilman S, Wenning GK, Low PA, et al. **Second consensus statement on the diagnosis of multiple system atrophy.** *Neurology* 2008;71:670–76 CrossRef Medline
10. Whitwell JL, Höglinger GU, Antonini A, et al; Movement Disorder Society-endorsed PSP Study Group. **Radiological biomarkers for diagnosis in PSP: where are we and where do we need to be?** *Mov Disord* 2017;32:955–71 CrossRef Medline
11. Meissner WG, Fernagut PO, Dehay B, et al. **Multiple system atrophy: recent developments and future perspectives.** *Mov Disord* 2019;34:1629–42 CrossRef Medline
12. Cosottini M, Ceravolo R, Faggioni L, et al. **Assessment of midbrain atrophy in patients with progressive supranuclear palsy with routine magnetic resonance imaging.** *Acta Neurol Scand* 2007;116:37–42 CrossRef Medline
13. Höglinger GU, Respondek G, Stamelou M, et al; Movement Disorder Society-endorsed PSP Study Group. **Clinical diagnosis of progressive**

supranuclear palsy: the Movement Disorder Society Criteria. *Mov Disord* 2017;32:853–64 CrossRef Medline
14. Mueller C, Hussl A, Krismer F, et al. **The diagnostic accuracy of the hummingbird and morning glory sign in patients with neurodegenerative parkinsonism.** *Parkinsonism Relat Disord* 2018;54:90–94 CrossRef Medline
15. Morelli M, Arabia G, Messina D, et al. **Effect of aging on magnetic resonance measures differentiating progressive supranuclear palsy from Parkinson's disease.** *Mov Disord* 2014;29:488–95 CrossRef Medline
16. Gröschel K, Kastrup A, Litvan I, et al. **Penguins and hummingbirds: midbrain atrophy in progressive supranuclear palsy.** *Neurology* 2006;66:949–50 CrossRef Medline
17. Bacchi S, Chim I, Patel S. **Specificity and sensitivity of magnetic resonance imaging findings in the diagnosis of progressive supranuclear palsy.** *J Med Imaging Radiat Oncol* 2018;62:21–31 CrossRef Medline
18. Brooks DJ, Seppi K; Neuroimaging Working Group on MSA. **Proposed neuroimaging criteria for the diagnosis of multiple system atrophy.** *Mov Disord* 2009;24:949–64 CrossRef Medline
19. Sankhla CS, Patil KB, Sawant N, et al. **Diagnostic accuracy of Magnetic Resonance Parkinsonism Index in differentiating progressive supranuclear palsy from Parkinson's disease and controls in Indian patients.** *Neurol India* 2016;64:239–45 CrossRef Medline
20. Ahn JH, Kim M, Kim JS, et al. **Midbrain atrophy in patients with presymptomatic progressive supranuclear palsy-Richardson's syndrome.** *Parkinsonism Relat Disord* 2019;66:80–86 CrossRef Medline
21. Möller L, Kassubek J, Südmeyer M, et al. **Manual MRI morphometry in Parkinsonian syndromes.** *Mov Disord* 2017;32:778–82 CrossRef Medline
22. Kato N, Arai K, Hattori T. **Study of the rostral midbrain atrophy in progressive supranuclear palsy.** *J Neurol Sci* 2003;210:57–60 CrossRef Medline
23. Warmuth-Metz M, Naumann M, Csoti I, et al. **Measurement of the midbrain diameter on routine magnetic resonance imaging: a simple and accurate method of differentiating between Parkinson disease and progressive supranuclear palsy.** *Arch Neurol* 2001;58:1076–79 CrossRef Medline
24. Armstrong MJ. **Progressive supranuclear palsy: an update.** *Curr Neurol Neurosci Rep* 2018;18:12 CrossRef Medline
25. Nigro S, Arabia G, Antonini A, et al. **Magnetic Resonance Parkinsonism Index: diagnostic accuracy of a fully automated algorithm in comparison with the manual measurement in a large Italian multicentre study in patients with progressive supranuclear palsy.** *Eur Radiol* 2017;27:2665–75 CrossRef Medline
26. Mangesius S, Hussl A, Tagwercher S, et al. **No effect of age, gender and total intracranial volume on brainstem MR planimetric measurements.** *Eur Radiol* 2020;30:2802–08 CrossRef Medline
27. Cahill L. **Why sex matters for neuroscience.** *Nat Rev Neurosci* 2006;7:477–84 CrossRef Medline
28. Tannenbaum C, Ellis RP, Eyssel F, et al. **Sex and gender analysis improves science and engineering.** *Nature* 2019;575:137–46 CrossRef Medline
29. Jancke L, Liem F, Merillat S. **Weak correlations between body height and several brain metrics in healthy elderly subjects.** *Eur J Neurosci* 2019;50:3578–89 CrossRef Medline
30. Willette AA, Kapogiannis D. **Does the brain shrink as the waist expands?** *Ageing Res Rev* 2015;20:86–97 CrossRef Medline
31. Håberg AK, Hammer TA, Kvistad KA, et al. **Incidental intracranial findings and their clinical impact: the HUNT MRI study in a general population of 1006 participants between 50-66 years.** *PLoS One* 2016;11:e0151080 CrossRef Medline
32. Krokstad S, Langhammer A, Hveem K, et al. **Cohort profile: the HUNT Study, Norway.** *Int J Epidemiol* 2013;42:968–77 CrossRef Medline
33. Honningsvåg LM, Linde M, Håberg A, et al. **Does health differ between participants and non-participants in the MRI-HUNT study, a population-based neuroimaging study? The Nord-Trøndelag health studies 1984-2009.** *BMC Med Imaging* 2012;12:23 CrossRef Medline
34. Cohen J. *Statistical Power Analysis for the Behavioral Sciences* (2nd ed). Hillsdale, NJ: Lawrence Earlbaum Associate; 1988
35. Swinscow TDV. *Statistics at Square One*. 9th ed. Revised by MJ Campbell. BMJ Publishing Group; 1996
36. Koo TK, Li MY. **A guideline of selecting and reporting intraclass correlation coefficients for reliability research.** *J Chiropr Med* 2016;15:155–63 CrossRef Medline
37. Miki Y, Foti SC, Asi YT, et al. **Improving diagnostic accuracy of multiple system atrophy: a clinicopathological study.** *Brain* 2019;142:2813–27 CrossRef Medline
38. Xu J, Kobayashi S, Yamaguchi S, et al. **Gender effects on age-related changes in brain structure.** *AJNR Am J Neuroradiol* 2000;21:112–18 Medline
39. Regitz-Zagrosek V. **Sex and gender differences in health: Science & Society Series on Sex and Science.** *EMBO Rep* 2012;13:596–603 CrossRef Medline
40. Mahale RR, Krishnan S, Divya KP, et al. **Gender differences in progressive supranuclear palsy.** *Acta Neurol Belg* 2021 Feb 17. [Epub ahead of print] CrossRef Medline
41. Mueller SG, Weiner MW, Thal LJ, et al. **The Alzheimer's Disease Neuroimaging Initiative.** *Neuroimaging Clin N Am* 2005;15:869–77. xi-xii CrossRef Medline
42. Huppertz HJ, Moller L, Sudmeyer M, et al. **Differentiation of neurodegenerative parkinsonian syndromes by volumetric magnetic resonance imaging analysis and support vector machine classification.** *Mov Disord* 2016;31:1506–17 CrossRef Medline
43. Scherfler C, Gobel G, Muller C, et al. **Diagnostic potential of automated subcortical volume segmentation in atypical parkinsonism.** *Neurology* 2016;86:1242–49 CrossRef Medline