Neuroradiology (2022) 64:473–481
https://doi.org/10.1007/s00234-021-02782-z

DIAGNOSTIC NEURORADIOLOGY

Quantitative imaging features predict spinal tap response in normal pressure hydrocephalus

Eyal Lotan[1] · Brianna E. Damadian[1] · Henry Rusinek[2] · Megan Griffin[1] · Benjamin Ades-Aron[2] · Ning Lu[1] · James Golomb[3] · Ajax E. George[1]

Received: 19 May 2021 / Accepted: 30 July 2021 / Published online: 20 August 2021

Abstract

Purpose Gait improvement following high-volume lumbar puncture (HVLP) and continuous lumbar drain (cLD) is widely used to predict shunt response in patients with suspected normal pressure hydrocephalus (NPH). Here, we investigate differences in MRI volumetric and traditional measures between HVLP/cLD responders and non-responders to identify imaging features that may help predict HVLP/cLD response.

Methods Eighty-two patients with suspected NPH were studied retrospectively. Gait testing was performed before and immediately/24 h/72 h after HVLP/cLD. A positive response was defined as improvement in gait post-procedure. Thirty-six responders (26 men; mean age 79.3 ± 6.3) and 46 non-responders (25 men; mean age 77.2 ± 6.1) underwent pre-procedure brain MRI including a 3D T1-weighted sequence. Subcortical regional volumes were segmented using FreeSurfer. After normalizing for total intracranial volume, two-way type III ANCOVA test and chi-square test were used to characterize statistical group differences. Evans' index, callosal angle (CA), and disproportionately enlarged subarachnoid space hydrocephalus were assessed. Multivariable logistic regression models were tested using Akaike information criterion to determine which combination of metrics most accurately predicts HVLP/cLD response.

Results Responders and non-responders demonstrated no differences in total ventricular and white/gray matter volumes. CA (men only) and third and fourth ventricular volumes were smaller; and hippocampal volume was larger in responders ($p < 0.05$). Temporal horns volume correlated with degree of improvement in gait velocity in responders ($p = 0.0006$). The regression model was 76.8% accurate for HVLP/cLD response.

Conclusion CA and third and fourth ventricular volumes and hippocampal volume may serve as potentially useful imaging features that may help predict spinal tap response and hence potentially shunt response.

Keywords Normal pressure hydrocephalus · Spinal tap · Volumetric analysis · MRI

Introduction

In the early 1960s, Colombian neurosurgeon Salomón Hakim first noted that two patients with severe hydrocephalus and motor impairment did not have increased intracranial pressure and unexpectedly improved after lumbar puncture. After documenting these cases [1], Hakim convinced Raymond Adams, a Harvard neurologist and neuropathologist, of his findings. Together, Hakim and Adams first described normal pressure hydrocephalus (NPH) and defined the now well-known clinical triad of gait impairment, urinary incontinence, and dementia as its hallmarks [2].

The prevalence of idiopathic NPH in the USA is about 0.2% among patients aged ≥ 60 years and about 6% among those 80 years and older [3]. The complete triad is neither

Presentation on a meeting: Findings were presented during the Alzheimer's Association International Conference (AAIC), 26–30 July, 2021.

Eyal Lotan
elotan@gmail.com

1 Department of Radiology, New York University Langone Medical Center, 660 1st Ave, 2nd Floor, New York, NY 10016, USA

2 Center for Biomedical Imaging, Department of Radiology, New York University Grossman School of Medicine, 660 First Ave, New York, NY 10016, USA

3 Department of Neurology, New York University Langone Medical Center, New York, NY, USA

necessary nor sufficient to make the diagnosis and today, in the presence of ventriculomegaly on MRI, gait impairment, which can be severe and incapacitating, is used to identify patients with suspected NPH. For further diagnosis, various clinical and imaging studies are performed, with a more definitive diagnosis (and treatment strategy) involving a positive response to cerebrospinal fluid (CSF) diversion via ventricular shunt surgery. The response to shunting is usually assessed through subjective assessment as well as quantitative gait studies, as gait impairment is the primary symptom and the one most likely to improve with shunting [4]. A reliable quantitative gait measure often used is the functional ambulation performance (FAP) score, which combines different spatiotemporal aspects of patients' gait at a self-selected walking speed [5].

In order to select patients who are likely to benefit from shunt surgery, most clinicians rely on a favorable response to high-volume CSF lumbar puncture (HVLP) or continuous lumbar drain (cLD), a CSF removal procedure over a 3-day period [6]. Owing to the invasive nature of shunting surgery, a large number of studies since the 1960s have attempted to identify predictors of shunt response. More recent studies have aimed to characterize imaging biomarkers of patients with shunt-responsive NPH [7–13]. Here, based on detailed gait studies following high-volume CSF removal, we develop a predictive model for selecting NPH patients with likely beneficial outcome. Potential predictors included demographic data, volumetric brain segmentation measures, and traditional NPH imaging biomarkers. The latter consist of the callosal angle (CA), Evans' index (EI) [10], and assessment of disproportionately enlarged subarachnoid space hydrocephalus (DESH) [12], all demonstrated positive predictive value in differentiating patients with NPH from patients without NPH. We hypothesized that structural volumetric and linear brain imaging features can help predict responders vs non-responders to CSF removal.

Methods

Patient cohort and clinical measures

This retrospective, anonymized, single-center study was approved by the institutional review board with a waiver of consent and was Health Insurance Portability and Accountability Act-compliant. Eighty-two consecutive patients with suspected NPH were studied. Subjects were chosen from our institution's adult hydrocephalus clinic, where they were referred for suspected NPH due to gait impairment in the presence of ventriculomegaly (with or without urinary incontinence or cognitive decline). Our initial cohort included 321 consecutive patients who underwent HVLP (drainage of at least 30 mL CSF) or cLD. It is generally believed that cLD is a more robust predictor of response to shunt surgery than HVLP [14–16]. The decision to offer cLD or HVLP was largely driven by clinical factors. Patients with more advanced dementia were preferentially sent for HVLP as were patients with a very prominent and consistent gait disturbance. Very slowly declining patients with mild or minimal cognitive deficits and a relatively minor gait disturbance tended to be sent for cLD.

From our cohort, we selected patients based upon the following inclusion criteria: (a) an available preprocedural MRI of the brain that included high-resolution magnetization-prepared rapid acquisition with gradient echo (MPRAGE) sequence and (b) documented clinical follow-up evaluation with gait testing before, immediately after, 24 h after, and 72 h after CSF removal and a recommendation of whether or not to proceed to shunt surgery. The majority (185/239) of patients were excluded because they failed to meet criterion "a." Thirty-nine patients were excluded because they failed to meet criterion "b." Seven patients with secondary NPH and confounding diagnoses (e.g., vascular dementia, territorial infarct, tumor resection) and 8 patients with motion artifact on MRI were also excluded, leading to a final sample size of $n = 82$.

Gait was evaluated objectively by using FAP, a validated method which provides a numerical score from 30 to 100 to quantitatively represent gait performance [5]. The FAP score focuses on functional aspects of gait and represents a quantification of patients' gait based on a selection of mean spatiotemporal gait parameters obtained at a self-selected speed. The following parameters are calculated: step time (in seconds), step length (SL) to leg length (LL) ratio, normalized velocity (V/LL) for each leg, and the degree of asymmetry for SL/LL ratio between both limbs. The values calculated for each parameter are then used to deduct points from the maximum score of 100 based on the differences from the normal range. FAP was determined using the GaitRite System (CIR Systems, Havertown, Pa), a computerized device with an electronic walkway that can detect foot contact and motion in order to analyze spatiotemporal aspects of gait [17].

Response to CSF removal was assessed by a multidisciplinary team led by a neurologist with over 25 years of experience. A positive response was considered positive if agreement was noted across three factors: (1) increase in FAP score, (2) subjective gait improvement noted by neurologist, and (3) subjective gait improvement noted by patient and/or caregiver.

MR image acquisition

MRI was performed on one of several local 1.5 T and 3 T MR imaging systems (Siemens AG, Erlangen, Germany) and included high-resolution T1-weighted MR images acquired by using an MPRAGE sequence. We used local three-dimensional MPRAGE sequences optimized for 1.5 or for 3 T imaging. Specific protocols were as follows: 3 T: repetition time 2100–2200 ms; echo time 2.3–4.0 ms; inversion time 1100 ms; flip angle 9–12°; matrix size 256 × 256 × 192; slice thickness 0.9–1.2 mm; and bandwidth 200/260 Hz/pixel; 1.5 T: repetition time 2100–2200 ms; echo time 3.8–4.0 ms; inversion time 1100 ms; flip angle 12°; matrix size 256 × 256 × 160; slice thickness 0.9–1.2 mm; and bandwidth 160/200 Hz/pixel.

Volumetric analysis

All T1-weighted MPRAGE sequences were visually inspected for satisfactory image quality prior to analysis. Brain segmentation was conducted using the automatic FreeSurfer image analysis suite, version 5.3 (https://surfer.nmr.mgh.harvard.edu) using default processing parameters and workflow [18]. The segmentation process was successfully completed in each case, and it took approximately 8 h per case. We recorded left and right hemisphere volumes of all brain substructures segmented and labelled by FreeSurfer. Also recorded were ventricular compartments and the total intracranial volume (TIV) estimates. Both absolute (mm^3) and normalized (i.e., expressed as the fraction of TIV) measures were analyzed.

Traditional measures

Two fellowship-trained neuroradiologists, blinded to clinical data, independently measured EI and CA and evaluated the cerebral sulci for DESH. EI was determined by measuring the maximum transverse width of the frontal horns on transaxial view and dividing by the maximum transverse internal skull diameter [19]. The T1-weighted MPRAGE sequence at the midsagittal plane was used to generate a reformatted coronal section at the level of the posterior commissure, with plane perpendicular to the anterior–posterior commissure line. CA, defined as the angle between the medial superior borders of the left and right lateral ventricles, was measured using the same coronal view as above [20]. The presence of DESH was identified by the characteristic pattern of crowding of the sulci superiorly near the vertex accompanied by enlargement of CSF spaces more inferiorly, particularly in the Sylvian fissures [21]. The presence of widely disproportionate sulci and fissures was also deemed to reflect a variant of DESH. All three-dimensional reformatting was performed using a Picture Archiving and Communications in Medicine multiplanar reconstruction tool (Intellispace PACS Enterprise v4.4.516; Philips Healthcare, Amsterdam, the Netherlands).

Table. 1 Demographic data, gait measures, and traditional imaging measures (mean ± standard deviation) for responders vs non-responders to CSF removal

Characteristic	Responders $N=36$	Non-responders $N=46$	p-value[#]
Age (y)	79.3 ± 6.3	77.2 ± 6.1	0.12
Gender (male/tot) (%)	26/36 (72%)	25/46 (54%)	0.10
Change FAP immediate (%)	9.6 ± 16.4	− 0.7 ± 6.9	0.0007*
Change FAP 24 h (%)	16.8 ± 16.4	− 2.0 ± 5.5	< 0.0001*
Change FAP 72 h (%)	9.8 ± 12.3	− 0.4 ± 8.8	0.005*
Callosal angle (degree)	77.9 ± 23.9	89.0 ± 19.0	0.02*
Evans' index	0.36 ± 0.04	0.37 ± 0.05	0.40
DESH (positive)	29/36 (81%)	37/46 (80%)	0.98

[#]Group differences are tested after correcting for age and gender effects using type III ANCOVA test

[*]Asterisk indicates statistically significant p-value (<0.05)

Statistical analysis

A dichotomous variable HVLP/cLD outcome (responders vs. non-responders) was created and used as a dependent variable in a multivariate logistic regression model. Metrics evaluated included age, gender, CA, EI, DESH, and the volumetric metrics computed with FreeSurfer. Both normalized and raw volumes were considered. For each metric, we tested its interaction with gender, since gender is associated with brain volumes [22]. For the final model, significant predictors were selected in a stepwise approach based on improvement (lower) in Akaike information criterion. Logistic regression results consisted of the area under the receiver operating characteristic curve (AUC), accuracy, sensitivity, and specificity. To characterize differences in individual quantitative variables between responder and non-responder groups, we used two-way type III ANCOVA test, correcting each measure for age and sex. Chi-square test was used for categorical variables. Intraclass correlation coefficients (ICCs) and Cohen's K were used to test for reliability between observers. For continuous variables (CA, EI), ICCs were calculated; for dichotomous variable (DESH), Cohen's K was calculated. In cases of discrepancy in the evaluation, the images were re-evaluated, and a consensus was reached. All statistical analyses were performed using statistical software (R version 4.0.2, The R Foundation for Statistical Computing, and SPSS version 23; IBM, Armonk, NY). Statistical significance was set at $p < 0.05$.

Results

Clinical and traditional imaging characteristics of the responder and the non-responder patient groups are listed in Table 1. The average CSF removed was 36.4 mL for the non-responders and 40.5 mL for the responders ($p=0.13$). There was no difference between the two subgroups as per age ($p=0.12$) and gender ($p=0.10$). As expected, the percentage change in FAP at all three time points post CSF removal was larger in responders than non-responders ($p<0.05$). Linear measurement of CA in men ($p=0.004$), but not in women ($p=0.61$), was a strong discriminator between the subgroups (Fig. 1). Neither EI nor DESH was different between the subgroups, and neither was affected by gender.

There was excellent interobserver agreement for CA (ICC = 0.96; 95% confidence interval: 0.94, 0.98) and good to excellent agreement for EI (ICC = 0.87; 95% confidence interval: 0.81, 0.95) and for DESH (Cohen's kappa coefficient = 0.78; 95% confidence interval: 0.72, 0.89).

Figure 2 and Table 2 show group distribution of brain structural volumes that demonstrated statistically significant difference for either raw or normalized volumes, respectively. Responders and non-responders demonstrated no differences in total ventricular volume or the total white and gray matter volumes. Of the regional volumes contained in the FreeSurfer ASEG subcortical atlas, only third and fourth ventricle volumes showed discrimination between the subgroups for both raw and normalized volumes, being smaller in responders prior to CSF removal ($p<0.05$; only a trend toward significance ($p=0.09$) for the raw volume of the third ventricle). Hippocampal (left and right sides combined) and brainstem volumes demonstrated differences between the subgroups for the raw volumes ($p<0.05$), which was lost after normalization to TIV. Correlation between temporal horns volume and percentage change in gait velocity was different between the subgroups ($p=0.0001$); larger temporal horn volume correlated strongly with percentage change in gait velocity ($r=0.71$) only in responders ($p<0.0006$).

Table 3 and Fig. 3 show the results of the multivariate logistic regression model constructed to distinguish responders from non-responders to CSF removal. Measures of age, gender, third and fourth ventricles volumes, combined hippocampal volume, and CA were found to be significant predictors. An online tool (HVLP/cLD Response Calculator; *https://www.firevoxel.org/im/tap_response_estimator.xlsx*) derived from this multivariate regression model was created to quantify probability of response to CSF removal on the basis of input of the above values. With a computed AUC of 0.81 (Fig. 3), the diagnostic accuracy of the model is 76.8%; sensitivity 66.7%; and specificity 84.7%.

Out of the total of 82 MRI exams, 28 were acquired at 1.5 T and 54 at 3 T field strength. There were 13/28 or 46% responders among those examined at 1.5 T vs 23/54 (43%) among those examined at 3 T, not a significant difference (chi-square test $p=0.74$). MRI field strength was not associated with age, gender, or any MRI volumetric measure. Moreover, the multivariate logistic regression model for the subset of 54 patients who underwent 3 T MRI shows a highly comparable performance to the model for all 82 patients, with the diagnostic accuracy of 79.6%; sensitivity 78.3%; specificity 80.7%, and AUC of 0.80.

It is important to note that virtually all of the responders were referred to shunt surgery and 33 of them eventually underwent shunting, of whom 31 had available medical records post shunting. Out of these 31, one patient suffered severe post procedural complication and 3 did not improve. Twenty-seven out of the remaining 30 (90%) reported clinical improvement after the operation.

Discussion

Our results suggest that structural volumetric analysis and CA predict HVLP/cLD response and may potentially predict a stronger shunt response as ~90% of our responders clinically improved after the surgery. We found that smaller

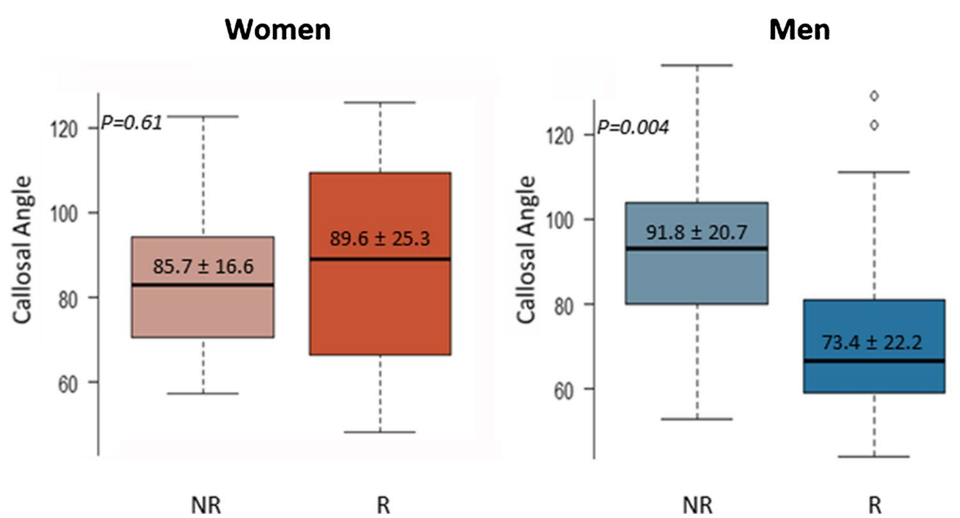

Fig. 1 Box-and-whisker plots and mean ± standard deviation of callosal angle gender distribution (in degrees) of responders and non-responders to CSF removal. R = responders; NR = non-responders

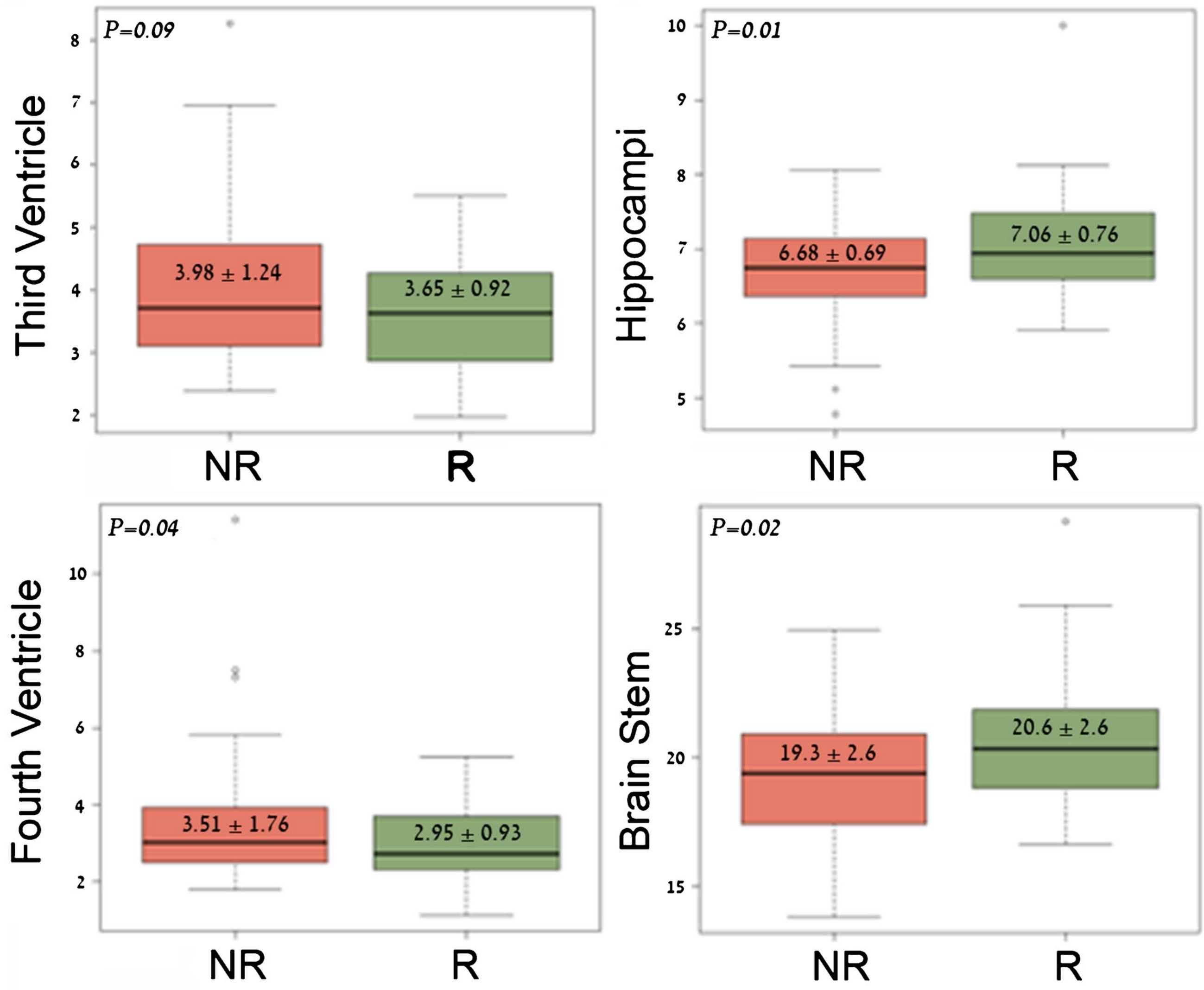

Fig. 2 Box-and-whisker plots and raw (non-normalized) mean ± standard deviation of brain structural volumes in responders and non-responders. R = responders; NR = non-responders

Table 2 Normalized volumes (mean ± standard deviation) of brain structures for responders to CSF removal vs non-responders

Brain structure	Normalized volume (% of TIV[†])		
	Responders	Non-responders	*p*-value[#]
Third ventricle	0.21 ± 0.05	0.23 ± 0.07	0.038*
Fourth ventricle	0.17 ± 0.05	0.21 ± 0.10	0.035*
Brainstem	1.17 ± 0.11	1.14 ± 0.13	0.27
Hippocampi (L + R)	0.40 ± 0.04	0.39 ± 0.05	0.57

[†]Normalization consists of division by TIV, the total intracranial volume as estimated by FreeSurfer software

[#]Group differences are tested after correcting for age and gender effects using type III ANCOVA test

[*]Asterisk indicates statistically significant *p*-value (< 0.05)

volumes of the 3rd and 4th ventricles, and smaller CA (for men), may be used to identify suspected NPH patients who are more likely to respond to CSF removal. In addition, we showed that in responders, larger volume of the temporal horns pre-treatment correlates with higher degree of gait improvement. These results, along with the predictive model we created, provide useful diagnostic and prognostic information by offering an additional supportive decision tool for clinicians considering a possible diagnosis of NPH.

For appropriately selected candidates, ventricular shunt surgery can lead to a dramatic improvement, resulting in ambulatory freedom and increased cognitive abilities [4]. As a result, multiple clinical trials are currently focusing on NPH for novel diagnostic, preventative, and therapeutic strategies [21]. The strategy for identifying NPH patients who are most likely to benefit from shunt surgery remains

Table. 3 Estimated coefficients of the optimal multivariate logistic regression model to distinguish responders from non-responders to CSF removal

Predictor	Coefficient	Odds ratio	*p*-value
Age	0.117	1.123	0.03*
Gender	4.85	128.1	0.03*
Third and fourth ventricles[a]	−0.346	0.707	0.04*
Hippocampi (L+R)	0.00111	1.001	0.01*
CA*Gender	−0.0544	0.947	0.04*

[a]The sum of the volumes of the third and fourth ventricles, normalized to the total ventricular volume to improve the precision of the measurements[20]

*Asterisk indicates statistically significant *p*-value (<0.05)

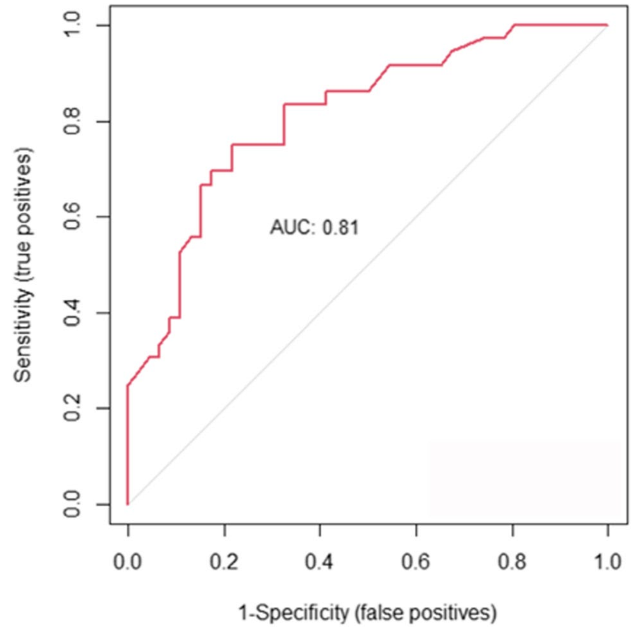

Fig. 3 Receiver-operating characteristic (ROC) curve and corresponding area under the curve (AUC) of the multivariate logistic regression prediction model for HVLP/cLD response

widely debated. DESH [22], high-convexity tightness [12], small CA [7, 23, 24], and temporal horn width [7, 25] have all been identified as possible imaging biomarkers to predict shunt response in NPH. Agerskov et al. argue that none of these biomarkers should be used to exclude patients from shunt surgery [13]. Agerskov's negative assessment is due to the fact that existing studies suffer from pre-selection bias. For example, most groups will tend to shunt only those patients who respond favorably to HVLP/cLD. Unlike the above studies, our analysis aimed to examine a presumably balanced sample of responders and non-responders.

While total ventricular volume was not different between responders and non-responders, our analysis indicates that the more caudal portions of the ventricular system (third and fourth ventricles) are smaller in responders. Similarly, Yamada et al. showed that tap responders and non-responders have no difference in total ventricular volumes. They also showed that tap responders have primarily *z*-axis (superior-inferior) expansion of the lateral ventricles rather than more global expansion or *x* or *y*-axial expansion of the ventricular system [26]. Our findings in combination with the results of Yamada et al. suggest that a key difference between tap responders and non-responders lies in the distribution of CSF within the ventricular system.

The correlation between larger temporal horns volume pre-treatment and a stronger response to CSF removal is supported by non-volumetric measurements showing correlation between wider temporal horns and shunt response [7], i.e., gait improvement. As opposed to our cohort of NPH patients, Annweiler et al. reported association between larger temporal horns to slower gait speed and greater stride-to-stride variability among 115 community-dwelling adults (mean age = 70 years) with mild cognitive impairment [27]. In these patients and more so in patients with Alzheimer's disease the common finding of larger temporal horns is probably due to hippocampal atrophy [28], whereas in NPH the hippocampi are more likely to be compressed by the increased CSF volume in the temporal horns, possibly contributing to the characteristic reduced cognitive function in these patients [29]. Taken together, we speculate that our finding of stronger response to CSF removal in patients with larger temporal horns may result from more substantial "decompression" of the mass effect on the hippocampi.

We also found that responders have larger hippocampal and brainstem raw volume (but not normalized). This may reflect decreased rate of other comorbidities that contribute to hippocampal and brainstem atrophy, such as possible concomitant pathology of early Alzheimer's disease, in responders as compared to non-responders [30–32]. Similar to our finding, Savolainen et al. showed that the hippocampi were larger in those for whom shunting improved gait disturbance or incontinence [33]. Inverse association between multiple gait measures and hippocampal volume was also reported by Ezzati et al. [34] in 112 community-residing adults, age 70 years and over, although adding memory performance to the models attenuated the association. Considering the known involvement of hippocampi in memory function, it is unclear if this association is related to normal brain aging processes or the influence of pre-clinical dementia. An interesting and somewhat surprising finding as part of our sub-analysis is that smaller CA in male gender, but not female, predisposes for response to CSF removal ($p = 0.004$). CA has been previously shown to be useful both in selection of shunt candidates (prognostic value) [35, 36] and in distinguishing NPH patients from patients with Alzheimer's disease (diagnostic value) [10], but no relationship with gender was reported. A gender difference in NPH pathophysiology and ventricular shape that is captured by CA metric are possible, but to date still largely unknown. Age-related gender differences were found to be localized to medial prefrontal regions [37]. Women have lower lateral ventricular volumes than men [38], implying that the shape and size of the ventricular system volume changes differently in males and females with aging. However, other studies found no gender differences in ventricular system volume and intracranial areas after normalization [39]. The significance of gender dependence of CA and the effect on its utility as an imaging biomarker for NPH needs to be evaluated in future studies to confirm these results and to elucidate the underlying mechanism.

DESH and EI were not found to predict response to CSF removal in our study, as supported by other studies that evaluated the prediction value of EI [7], and DESH [40] in suspected NPH patients after shunting.

Some limitations of our study should be considered. First, this work is based on a retrospective cohort from a single institution with somewhat small subgroubs. In particular, imaging data available were limited and 239 patients had to be excluded. Future prospective studies are required to avoid potential bias. Second, HVLP achieves varying levels of sensitivity and specificity in predicting positive response to shunt, which may limit the extrapolation of our findings to predict shunt responsiveness [41]. Ideally, it was preferable to predict response to shunt surgery rather than to HVLP/cLD. However, in our cohort, ~ 90% of HVLP/cLD responders clinically improved following shunt surgery. Consequently, the subgroup of non-responders to shunt surgery was too small (~ 3 patients) for further statistical analysis (poor statistical power). We therefore chose HVLP/cLD response to dichotomize the broad sample of patients meeting clinical and radiographic criteria for NPH into those with a very high probability of improving with shunt surgery and those with a lower probability of improving. While not a perfect surrogate for actual shunt responsiveness, we anticipated that such dichotomization would permit a valid investigation of the proposed radiographic markers. Similar to our results, the predictive value of spinal tap test is progressively improving over the recent years of with evidence of high positive predictive value (> 90%) [42]. Third, volumetric analysis of brain structures yields multiple measures that may be correlated and raise the well-known problem of multiple comparisons. While some of our results (Tables 1 and 2) are not corrected for multiple comparisons, we dealt with this issue by limiting the measures to those previously reported in the literature focusing on scientifically sensible comparisons; combining measures into composites, like left and right hippocampus, and third and fourth ventricles; and focusing on the analysis on the multivariate model (Table 3).

Conclusion

Our results showed that imaging biomarkers, such as CA, third and fourth ventricular volumes, and hippocampal volume, combined in a predictive model, may provide a useful noninvasive diagnostic supportive tool potentially helpful in assessing the need for HVLP/cLD by predicting its success chances and hence the success of shunt surgery.

Author contribution All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by all authors. The first draft of the manuscript was written by Eyal Lotan and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Data availability Full anonymized data will be shared at the request from any qualified investigator.

Code availability Not applicable.

Declarations

Ethical approval The study as performed in line with the principles of the Declaration of Helsinki. It was approved by the IRB committee of NYU Langone Medical Center.

Consent to participate Not applicable.

Consent for publication Not applicable.

Conflicts of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

1. Hakim S (1964) Some observations on C.S.F. Pressure. Hydrocephalic syndrome in adults with "normal" C.S.F. pressure (recognition of a new syndrome). Dissertation [Thesis No. 957]. Javeriana University School of Medicine, Bogota
2. Adams RD, Fisher CM, Hakim S et al (1965) Symptomatic occult hydrocephalus with "normal" cerebrospinal-fluid pressure. A treatable syndrome. N Engl J Med 273:117–126. https://doi.org/10.1056/NEJM196507152730301
3. Alvi MA, Brown D, Yolcu Y et al (2020) Prevalence and trends in management of idiopathic normal pressure hydrocephalus in the United States: insights from the national inpatient sample. World Neurosurg S1878–8750(20):32020–32029. https://doi.org/10.1016/j.wneu.2020.09.012
4. Shaw R, Everingham E, Mahant N et al (2016) Clinical outcomes in the surgical treatment of idiopathic normal pressure hydrocephalus. J Clin Neurosci 29:81–86. https://doi.org/10.1016/j.jocn.2015.10.044
5. Gouelle A (2014) Use of Functional Ambulation Performance Score as measurement of gait ability: review. J Rehabil Res Dev 51:665–674. https://doi.org/10.1682/JRRD.2013.09.0198

6. Marmarou A, Young HF, Aygok GA et al (2005) Diagnosis and management of idiopathic normal-pressure hydrocephalus: a prospective study in 151 patients. J Neurosurg 102:987–997. https://doi.org/10.3171/jns.2005.102.6.0987
7. Virhammar J, Laurell K, Cesarini KG, Larsson EM (2014) Preoperative prognostic value of MRI findings in 108 patients with idiopathic normal pressure hydrocephalus. AJNR Am J Neuroradiol 35:2311–2318. https://doi.org/10.3174/ajnr.A4046
8. Serulle Y, Rusinek H, Kirov II et al (2014) Differentiating shunt-responsive normal pressure hydrocephalus from Alzheimer disease and normal aging: pilot study using automated MRI brain tissue segmentation. J Neurol 261:1994–2002. https://doi.org/10.1007/s00415-014-7454-0
9. Kojoukhova M, Koivisto AM, Korhonen R et al (2015) Feasibility of radiological markers in idiopathic normal pressure hydrocephalus. Acta Neurochir (Wien) 157:1709–1719. https://doi.org/10.1007/s00701-015-2503-8
10. Miskin N, Patel H, Franceschi AM et al (2017) Diagnosis of normal-pressure hydrocephalus: use of traditional measures in the era of volumetric MR Imaging. Radiology 285:197–205. https://doi.org/10.1148/radiol.2017161216
11. Kockum K, Lilja-Lund O, Larsson EM et al (2018) The idiopathic normal-pressure hydrocephalus Radscale: a radiological scale for structured evaluation. Eur J Neurol 25:569–576. https://doi.org/10.1111/ene.13555
12. Narita W, Nishio Y, Baba T et al (2016) High-convexity tightness predicts the shunt response in idiopathic normal pressure hydrocephalus. AJNR Am J Neuroradiol 37:1831–1837. https://doi.org/10.3174/ajnr.A4838
13. Agerskov S, Wallin M, Hellström P et al (2019) Absence of disproportionately enlarged subarachnoid space hydrocephalus, a sharp callosal angle, or other morphologic MRI markers should not be used to exclude patients with idiopathic normal pressure hydrocephalus from shunt surgery. AJNR Am J Neuroradiol 40:74–79. https://doi.org/10.3174/ajnr.A5910
14. Walchenbach R, Geiger E, Thomeer RT, Vanneste J (2002) The value of temporary external lumbar CSF drainage in predicting the outcome of shunting on normal pressure hydrocephalus. J Neurol Neurosurg Psychiatry 72:503–506. https://doi.org/10.1136/jnnp.72.4.503
15. Williams MA, Razumovsky AY, Hanley DF (1998) Comparison of Pcsf monitoring and controlled CSF drainage diagnose normal pressure hydrocephalus. Acta Neurochir Suppl (Wien) 71:328–330. https://doi.org/10.1007/978-3-7091-6475-4_95
16. Haan J, Thomeer RT (1988) Predictive value of temporary external lumbar drainage in normal pressure hydrocephalus. Neurosurgery 22:388–391. https://doi.org/10.1227/00006123-198802000-00020
17. Webster KE, Wittwer JE, Feller JA (2005) Validity of the GAITRite walkway system for the measurement of averaged and individual step parameters of gait. Gait Posture 22:317–321. https://doi.org/10.1016/j.gaitpost.2004.10.005
18. Fischl B, Salat DH, Busa E et al (2002) Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron 33:341–355. https://doi.org/10.1016/s0896-6273(02)00569-x
19. Orrison WW (2000) Neuroimaging. Saunders, Philadelpha
20. Ishii K, Kanda T, Harada A et al (2008) Clinical impact of the callosal angle in the diagnosis of idiopathic normal pressure hydrocephalus. Eur Radiol 18:2678–2683. https://doi.org/10.1007/s00330-008-1044-4
21. Sasaki M, Honda S, Yuasa T et al (2008) Narrow CSF space at high convexity and high midline areas in idiopathic normal pressure hydrocephalus detected by axial and coronal MRI. Neuroradiology 50:117–122. https://doi.org/10.1007/s00234-007-0318-x
22. Ritchie SJ, Cox SR, Shen X et al (2018) Sex differences in the adult human brain: evidence from 5216 UK biobank participants. Cereb Cortex 28:2959–2975. https://doi.org/10.1093/cercor/bhy109
23. Virhammar J, Laurell K, Cesarini KG, Larsson EM (2014) The callosal angle measured on MRI as a predictor of outcome in idiopathic normal-pressure hydrocephalus. J Neurosurg 120:178–184. https://doi.org/10.3171/2013.8.JNS13575
24. Grahnke K, Jusue-Torres I, Szujewski C et al (2018) The quest for predicting sustained shunt response in normal-pressure hydrocephalus: an analysis of the callosal angle's utility. World Neurosurg 115:e717–e722. https://doi.org/10.1016/j.wneu.2018.04.150
25. ilja-Lund O, Kockum K, Hellström P et al (2020) Wide temporal horns are associated with cognitive dysfunction, as well as impaired gait and incontinence. Sci Rep 10:18203. https://doi.org/10.1038/s41598-020-75381-2
26. Yamada S, Ishikawa M, Yamamoto K (2015) Optimal diagnostic indices for idiopathic normal pressure hydrocephalus based on the 3D quantitative volumetric analysis for the cerebral ventricle and subarachnoid space. AJNR Am J Neuroradiol 36:2262–2269. https://doi.org/10.3174/ajnr.A4440
27. Annweiler C, Montero-Odasso M, Bartha R et al (2014) Association between gait variability and brain ventricle attributes: a brain mapping study. Exp Gerontol 57:256–263. https://doi.org/10.1016/j.exger.2014.06.015
28. Frisoni GB, Geroldi C, Beltramello A et al (2002) Radial width of the temporal horn: a sensitive measure in Alzheimer disease. AJNR Am J Neuroradiol 23:35–47
29. Muñoz-López M (2015) Past, present, and future in hippocampal formation and memory research. Hippocampus 25:726–730. https://doi.org/10.1002/hipo.22452
30. Cabral D, Beach TG, Vedders L et al (2011) Frequency of Alzheimer's disease pathology at autopsy in patients with clinical normal pressure hydrocephalus. Alzheimers Dement 7:509–513. https://doi.org/10.1016/j.jalz.2010.12.008
31. Ji X, Wang H, Zhu M et al (2021) Brainstem atrophy in the early stage of Alzheimer's disease: a voxel-based morphometry study. Brain Imaging Behav 15(1):49–59. https://doi.org/10.1007/s11682-019-00231-3
32. Golomb J, de Leon MJ, George AE et al (1994) Hippocampal atrophy correlates with severe cognitive impairment in elderly patients with suspected normal pressure hydrocephalus. J Neurol Neurosurg Psychiatry 57:590–593. https://doi.org/10.1136/jnnp.57.5.590
33. Savolainen S, Laakso MP, Paljärvi L et al (2000) MR imaging of the hippocampus in normal pressure hydrocephalus: correlations with cortical Alzheimer's disease confirmed by pathologic analysis. AJNR Am J Neuroradiol 21:409–414
34. Ezzati A, Katz MJ, Lipton ML et al (2015) The association of brain structure with gait velocity in older adults: a quantitative volumetric analysis of brain MRI. Neuroradiology 57:851–861. https://doi.org/10.1007/s00234-015-1536-2
35. Klasson N, Olsson E, Eckerström C, Malmgren H, Wallin A (2018) Estimated intracranial volume from FreeSurfer is biased by total brain volume. Eur Radiol Exp 2:24. https://doi.org/10.1186/s41747-018-0055-4
36. Abraham ME, Povolotskiy R, Gold J et al (2020) The current state of clinical trials studying hydrocephalus: an analysis of ClinicalTrials.gov. Cureus 12:e10029. https://doi.org/10.7759/cureus.10029
37. Cowell PE, Sluming VA, Wilkinson ID et al (2007) Effects of sex and age on regional prefrontal brain volume in two human cohorts. Eur J Neurosci 25:307–318. https://doi.org/10.1111/j.1460-9568.2006.05281.x
38. Brix MK, Westman E, Simmons A et al (2017) The Evans' Index revisited: new cut-off levels for use in radiological assessment of

ventricular enlargement in the elderly. Eur J Radiol 95:28–32. https://doi.org/10.1016/j.ejrad.2017.07.013
39. Cosgrove KP, Mazure CM, Staley JK (2007) Evolving knowledge of sex differences in brain structure, function, and chemistry. Biol Psychiatry 62:847–855. https://doi.org/10.1016/j.biopsych.2007.03.001
40. Ahmed AK, Luciano M, Moghekar A et al (2018) Does the presence or absence of desh predict outcomes in adult hydrocephalus? AJNR Am J Neuroradiol 39:2022–2026. https://doi.org/10.3174/ajnr.A5820
41. Mihalj M, Dolic´ K, Kolic´ K, Ledenko V (2016) CSF tap test - obsolete or appropriate test for predicting shunt responsiveness? A systemic review. J Neurol Sci 362:78–84. https://doi.org/10.1016/j.jns.2016.01.028
42. Wikkelso C, Hellstrom P, Klinge PM, Tans JT (2013) The European iNPH Multicentre Study on the predictive values of resistance to CSF outflow and the CSF Tap Test in patients with idiopathic normal pressure hydrocephalus. J Neurol Neurosurg Psychiatry 84:562–568. https://doi.org/10.1136/jnnp-2012-303314