Prognostic Utility of Disproportionately Enlarged Subarachnoid Space Hydrocephalus in Idiopathic Normal Pressure Hydrocephalus Treated with Ventriculoperitoneal Shunt Surgery: A Systematic Review and Meta-analysis


ABSTRACT

BACKGROUND: Disproportionately enlarged subarachnoid space hydrocephalus is a specific radiologic marker for idiopathic normal pressure hydrocephalus. However, controversy exists regarding the prognostic utility of disproportionately enlarged subarachnoid space hydrocephalus.

PURPOSE: Our aim was to evaluate the prevalence of disproportionately enlarged subarachnoid space hydrocephalus in idiopathic normal pressure hydrocephalus and its predictive utility regarding prognosis in patients treated with ventriculoperitoneal shunt surgery.

DATA SOURCES: We used MEDLINE and EMBASE databases.

STUDY SELECTION: We searched for studies that reported the prevalence or the diagnostic performance of disproportionately enlarged subarachnoid space hydrocephalus in predicting treatment response.

DATA ANALYSIS: The pooled prevalence of disproportionately enlarged subarachnoid space hydrocephalus was obtained. Pooled sensitivity, specificity, and area under the curve of disproportionately enlarged subarachnoid space hydrocephalus to predict treatment response were obtained. Subgroup and sensitivity analyses were performed to explain heterogeneity among the studies.

DATA SYNTHESIS: Ten articles with 812 patients were included. The pooled prevalence of disproportionately enlarged subarachnoid space hydrocephalus in idiopathic normal pressure hydrocephalus was 44% (95% CI, 34%–54%). The pooled prevalence of disproportionately enlarged subarachnoid space hydrocephalus was higher in the studies using the second edition of the Japanese Guidelines for Management of Idiopathic Normal Pressure Hydrocephalus compared with the studies using the international guidelines without statistical significance (52% versus 43%, \( P = .38 \)). The pooled sensitivity and specificity of disproportionately enlarged subarachnoid space hydrocephalus for prediction of treatment response were 59% (95% CI, 38%–77%) and 66% (95% CI, 57%–74%), respectively, with an area under the curve of 0.67 (95% CI, 0.63–0.71).

LIMITATIONS: The lack of an established method for assessing disproportionately enlarged subarachnoid space hydrocephalus using brain MR imaging served as an important cause of the heterogeneity.

CONCLUSIONS: Our meta-analysis demonstrated a relatively low prevalence of disproportionately enlarged subarachnoid space hydrocephalus in idiopathic normal pressure hydrocephalus and a poor diagnostic performance for treatment response.

ABBREVIATIONS: DESH = disproportionately enlarged subarachnoid space hydrocephalus; HSROC = hierarchical summary receiver operating characteristic; iNPH = idiopathic normal pressure hydrocephalus; NPV = negative predictive value; PPV = positive predictive value; VP = ventriculoperitoneal

Received January 11, 2021; accepted after revision March 17.
From the Department of Radiology and Research Institute of Radiology (H.Y.P., C.H.S., M.J.K., W.H.S., S.J.K.), Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea; and Department of Medical Science (C.R.P.), Asan Medical Institute of Convergence Science and Technology, University of Ulsan College of Medicine, Seoul, Korea.

H.Y. Park and C.R. Park contributed equally to the manuscript.

http://dx.doi.org/10.3174/ajnr.A7168


Diopathic normal pressure hydrocephalus (iNPH) is a potentially reversible disease characterized by the triad of gait disturbance, dementia, and urinary incontinence.1,2 Currently, the only effective treatment for iNPH is CSF shunt surgery, with ventriculoperitoneal (VP) shunt surgery being the most commonly performed.3 However, the procedure is invasive and is associated
with complications such as shunt obstruction, which may require multiple revision surgeries.\textsuperscript{5,5} Hence, careful selection of patients for this surgery is critical. The CSF tap test or drainage test and CSF infusion test have been proposed to predict treatment response in iNPH.\textsuperscript{1,2,6} However, several studies have questioned the predictive values of these tests.\textsuperscript{7,9} In addition, these tests are invasive and pose the potential risk of infection.\textsuperscript{8,10,11}

Several studies have attempted identifying radiologic markers to predict treatment response in iNPH, including the callosal angle and disproportionately enlarged subarachnoid space hydrocephalus (DESH).\textsuperscript{1,2,15} DESH refers to a morphologic pattern of communicating hydrocephalus that features uneven distribution of CSF between the superior and inferior subarachnoid spaces.\textsuperscript{15} DESH has gained recognition for its prognostic and diagnostic utility in iNPH.\textsuperscript{1,16-18} Consequently, the third edition of the Japanese Guidelines for Management of Idiopathic Normal Pressure Hydrocephalus adopted DESH as an imaging marker in the diagnostic criteria for iNPH.\textsuperscript{19} The seminal study supporting this decision was published in 2010 by Hashimoto et al.,\textsuperscript{17} which reported a high prevalence of DESH in iNPH and an excellent positive predictive value (PPV) of DESH for treatment response. However, this study had not included a control group (patients negative for DESH) and hence did not report the negative predictive value (NPV). Several studies published since have presented contradictory results.\textsuperscript{12,16,20-25} Of note, a few studies have questioned the clinical value of DESH for its poor NPV.\textsuperscript{20,21,26} In addition, the reported prevalence of DESH varies widely across studies, ranging from 22% to 96%.\textsuperscript{12,16,17,27}

To our knowledge, the diagnostic performance of DESH in the prediction of treatment response and the prevalence of DESH in iNPH have not been systematically evaluated. Therefore, we conducted a systematic review and meta-analysis to investigate the prevalence of DESH in iNPH and evaluate its clinical value in predicting treatment response.

**MATERIALS AND METHODS**

**Evidence Acquisition**

This study was performed and reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.\textsuperscript{28} Neither institutional review board approval nor written informed consent was required owing to the nature of the study.

**Literature Search**

A systematic literature search of the MEDLINE and EMBASE databases was conducted to identify studies that reported the prevalence of DESH in iNPH or the diagnostic performance of DESH in predicting treatment response in patients who underwent VP shunt surgery. The search term was (normal pressure hydrocephalus OR (NPH)) AND ((disproportionately enlarged subarachnoid space hydrocephalus) OR (DESH)), using the combination of a MESH term and free-text terms. The search was performed on October 10, 2020.

**Inclusion Criteria**

DESH was diagnosed when all 3 components, enlarged ventricles, tight high convexity, and dilated Sylvian fissure, were present (Fig 1).\textsuperscript{1} Some studies have used the term “incomplete” DESH to describe a status of features of DESH being partially present.\textsuperscript{2,19} However, to maintain consistency among the results of the included studies, we focused on DESH with all 3 features. To investigate the prevalence of DESH, we selected studies if they performed performing brain imaging to evaluate DESH in patients with iNPH. To evaluate the diagnostic performance of DESH for prediction of treatment response, we selected studies if they met all of the following criteria: 1) VP shunt surgery performed for iNPH; 2) brain MR imaging performed for the evaluation of DESH; 3) inclusion of the analysis of treatment response after VP shunt; and 4) availability of adequate information for the reconstruction of 2 × 2 tables to calculate the diagnostic performance of DESH.

**Exclusion Criteria**

Studies were excluded if they met any of the following criteria: 1) case reports and case series with <5 patients, 2) conference abstracts, 3) reviews, 4) letters to the editor or editorials, and 5) incomplete data for reconstruction of 2 × 2 tables. Two reviewers (H.Y.P, with 4 years of experience in diagnostic radiology, and C.H.S., with 9 years of experience in diagnostic radiology) independently evaluated the eligibility of each article. Interviewer disagreements were resolved through discussions to form a consensus.

**Data Extraction and Quality Assessment**

A standardized extraction form was used to extract data from the studies. The extracted data included the following: 1) study characteristics: author, year of publication, institution, country of origin, study period, study design, consecutive-versus-nonconsecutive enrollment, reference standards, outcome measures, blinding to
outcome measures, and follow-up periods; 2) patient characteristics: number of patients with iNPH, number of patients responsive to VP shunts, mean age, age range, and male-to-female ratio; and 3) characteristics of brain MR imaging: magnet strength, vendor, scanner, MR image, and image plane for the evaluation of DESH, method of evaluation, and the number and experience of the readers. On the basis of the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool, we evaluated the quality of the studies.30 Data extraction and quality assessment were independently conducted by 2 reviewers (H.Y.P. and C.H.S.).

Data Synthesis and Analysis

The primary outcome of our study was the prevalence of DESH among patients with iNPH. The pooled proportion and its 95% confidence interval were calculated using the DerSimonian and Laird random effects model, and a forest plot was constructed. The Cochran Q test and Higgins inconsistency index (I^2) test were used to evaluate heterogeneity, and a funnel plot was constructed to assess publication bias.31,32 To explain heterogeneity among the studies, we performed subgroup analysis based on which iNPH guidelines were being used for the patient inclusion.

The secondary outcome of our study was the diagnostic performance of DESH in predicting treatment response in patients with iNPH who underwent VP shunt surgery. Pooled sensitivity, pooled specificity, and their corresponding 95% CIs were calculated using a bivariate random effects model, and coupled forest plots were constructed. On the basis of the calculated prevalence of patients responsive to treatment (total number of patients responsive to the shunt/total number of patients with iNPH) and pooled sensitivity and specificity, the PPV and NPV of DESH were obtained using the following formulæ:

\[
\text{PPV} = \frac{(\text{Sensitivity} \times \text{Prevalence}) + (1 - \text{Prevalence})}{\text{Sensitivity} \times \text{Prevalence} + (1 - \text{Prevalence}) \times (1 - \text{Specificity})}
\]

\[
\text{NPV} = \frac{[\text{Specificity} \times (1 - \text{Prevalence})] / [\text{Specificity} \times (1 - \text{Prevalence}) + \text{Prevalence} \times (1 - \text{Specificity})]}
\]

A hierarchical summary receiver operating characteristic (HSROC) curve with 95% confidence and prediction regions was plotted. Sensitivity analyses were performed to explain study heterogeneity. Heterogeneity among the studies was determined using the Cochran Q test and Higgins inconsistency index (I^2) test; a Cochran Q test with a P < .05 or I^2 > 50% was considered to indicate heterogeneity.31,32 In addition, the difference between the 95% confidence region and prediction region in the HSROC curve was visually analyzed, with a large difference indicating heterogeneity. A threshold effect (ie, positive correlation between sensitivity and the false-positive rate) was evaluated by visual assessment of the coupled forest plots and the Spearman correlation coefficient. A correlation coefficient of > 0.6 indicated a threshold effect.33 Publication bias was assessed using the Deeks funnel plot, and statistical significance was evaluated using the Deeks asymmetry test.34 For statistical analyses, the metandi and midas modules in STATA, Version 15.0 (StataCorp), and R, Version 3.6.3 (http://www.r-project.org), were used.

RESULTS

Literature Search

Figure 2 summarizes our study-selection process. In total, 130 articles were obtained from the systematic search. After we removed 4 duplicate articles, 126 articles were screened for eligibility on the basis of titles and abstracts, and 110 articles were excluded. One additional eligible study was identified from the bibliographies of articles.12 After full-text reviews of 17 articles, 7 articles were further excluded for the following reasons: Three articles were not in the field of interest,14,18,35 and 4 articles had insufficient information for reconstruction of 2 × 2 tables.13,15,17,36 Of the 4 articles, 3 articles reported an association between DESH and postsurgical outcome based on correlation and regression analyses.13,15,36 However, sensitivity and specificity could not be calculated in these studies, and they were excluded from the analysis. Finally, 10 original articles were included in the study.12,16,20–27

Characteristics of the Included Studies

The study and patient characteristics of the 10 articles are described in the Online Supplemental Data. All studies reported the prevalence of DESH. Of those, 8 studies reported the diagnostic performance of DESH with regard to treatment response.12,16,20–25 Three studies were prospective,16,22,25 and 7 studies were retrospective.12,20,21,23,24,26,27 Consecutive enrollment was performed in 7 studies,12,16,20,22,24,25 whereas 3 studies did not detail patient enrollment.23,26,27 Four studies used the international guidelines,20,22,25 while 3 studies used the second edition of the Japanese guidelines for the diagnosis of iNPH and the selection of patients receiving a VP shunt.23,24,27 The other 3 studies did not mention the reference standard.12,16,26 Of note, only patients with positive CSF tap test results underwent VP shunt surgery in 4 studies, possibly causing a selection bias.12,23,24 For the outcome measurements, 5 studies used either the INPH grading scale or the NPH Eide scale, which are systematic grading scales that focus on 3 domains: gait, cognitive, and urinary disturbance.12,16,22–24 The rest of the studies used various quantitative or qualitative methods including the Timed 10-Meter Walk Test or the revised Wechsler Adult Intelligence Scale neuropsychology report.20,21,25 In 5 studies, readers were blinded to outcome measures.16,20,21,26,27 The remaining 5 studies did not report on blinding.12,22–25 Most studies assessed treatment response 12 months after VP shunt surgery.16,20,22,25 In 1 study, treatment response was evaluated 10 days after VP shunt surgery,23 whereas another study included long follow-up periods (51.8 months; interquartile range, 22.2–64.2 months).12 One study did not report the time point for treatment-response evaluation.24

The characteristics of brain MR imaging performed in the studies are summarized in the Online Supplemental Data. In 6 studies, T1-weighted imaging was used,16,20,22,24,25,27 whereas the remaining 4 studies did not mention which MR image was used.12,21,23,26 Visual qualitative assessment was performed in all studies for the evaluation of DESH based on the axial and/or coronal image plane. The number of readers ranged from 1 to 3.
Quality Assessment

On the basis of the QUADAS-2 criteria, 6 of 10 articles satisfied at least 4 of the 7 items, indicating reasonable analysis quality (Online Supplemental Data). The detailed description of quality assessment is provided in the Online Supplemental Data.

Prevalence of DESH among Patients with iNPH

Ten studies reported the prevalence of DESH in patients with iNPH, which ranged from 22% to 64%.12,16,20-27 The pooled prevalence of DESH was 44% (95% CI, 34%–54%) (Fig 3). The Cochran Q test and the I² test demonstrated considerable heterogeneity among the studies (Q = 66.84; P < .001; I² = 87%). No significant publication bias was observed on the funnel plot (Online Supplemental Data).

Subgroup analysis demonstrated that the pooled prevalence of DESH was higher in the studies using the second edition of the Japanese guidelines compared with studies using the international guidelines, but without statistical significance (52% versus 43%, P = .38) (Online Supplemental Data). The heterogeneity among the studies was slightly decreased on subgroup analysis (international guidelines: I² = 77%; Japanese guidelines: I² = 72%).

Four studies used a CSF tap test as a prerequisite for surgical eligibility.12,23-25 No significant difference was observed in the prevalence of DESH among these 4 studies and the rest of the studies (P = .79).

Diagnosis Performance of DESH for Prediction of Treatment Response

Eight studies reported the diagnostic performance of DESH in predicting treatment response.12,16,20-27 The sensitivity and specificity of DESH in individual studies ranged from 20% to 93% and 48% to 100%, respectively. The pooled sensitivity and specificity were 59% (95% CI, 38%–77%) and 66% (95% CI, 57%–74%), respectively (Fig 4). The area under the HSROC curve was 0.67 (95% CI, 0.63–0.71) (Fig 5). Based on the pooled estimates, the calculated PPV and NPV were 80% and 41%, respectively.

The Cochran Q test and Higgins I² test demonstrated significant study heterogeneity regarding sensitivity (Q = 87.25; P = .00; I² = 92%) and specificity (Q = 15.45; P = .03; I² = 55%). In addition, a notable difference was observed between the 95% confidence region and the prediction region in HSROC, indicating considerable heterogeneity (Fig 5). Visual analysis of the coupled forest plot showed a high likelihood of the threshold effect (Fig 4), though the Spearman correlation coefficient between the sensitivity and the false-positive rate was not significant (correlation coefficient, 0.336; 95% CI, –0.483–0.842). The Deeks funnel plot showed a low likelihood of publication bias (P = .26) (Online Supplemental Data).

No significant difference was observed in the diagnostic performance of DESH among the 4 studies depending solely on CSF tap test for surgical eligibility and the rest of the studies (P = .74). All except 3 studies evaluated treatment response at 12 months after VP shunt surgery. When a sensitivity analysis was performed excluding those 3 studies,12,23,24 the diagnostic performance was slightly improved with a pooled sensitivity of 62% (95% CI, 37%–82%) and a specificity of 70% (95% CI, 60%–78%) (Online Supplemental Data). Additional sensitivity analysis was conducted regarding the 5 studies that used systematic grading
scales for the evaluation of treatment response. Pooled sensitivity was slightly improved to 70% (95% CI, 41%–88%), but no considerable change was observed in the pooled specificity of 64% (95% CI, 48%–76%) (Online Supplemental Data). The degree of heterogeneity remained similar in both sensitivity analyses.

DISCUSSION

DESH has been regarded as an imaging marker that aids in the diagnosis of iNPH and the prediction of treatment response after shunt surgery. However, our study demonstrated poor diagnostic performance of DESH with regard to the prediction of treatment response (pooled sensitivity of 59% and specificity of 66%) and a relatively low prevalence of DESH (44%) in iNPH. In addition, the calculated NPV of DESH based on the pooled estimates was poor (41%), which means that a substantial number of patients in the included studies had improvement of symptoms after the operation even without DESH. Therefore, our study suggests that patients negative for DESH should not be excluded from VP shunt surgery.

The concept of DESH was first introduced in the study by Hashimoto, et al., in which they reported a high prevalence of DESH (96%). However, patients were included in this study only if they showed tight high convexity and ventriculomegaly on brain MR imaging. Therefore, the calculated prevalence of DESH in the study was inevitably high because the included patients already had at least 2 features of DESH. On the contrary, our study demonstrated a pooled prevalence of DESH of 44% in patients with iNPH. Our results support previous studies reporting that a considerable number of patients had no or partial features of DESH. Partial features of DESH or incomplete DESH was defined in a few studies as 1 or 2 definite features of the 3 components of DESH. Currently, the evidence is scarce regarding the prevalence of incomplete DESH in patients with iNPH and its role in the diagnosis. In addition, it is difficult to differentiate incomplete DESH from mere brain atrophy. In our meta-analysis, patients with incomplete DESH were excluded to maintain consistency among the studies, possibly further lowering the prevalence of DESH. Indeed, the prevalence of DESH increased from 64% to 87% in 1 of the included studies when incomplete DESH was taken into account.

We speculated that the prevalence of DESH would be higher in the studies that used the second edition of the Japanese guidelines for patient inclusion, because DESH was part of the diagnostic criteria in the guidelines. In subgroup analysis, a higher pooled prevalence of DESH (52%) was observed in the studies using the second Japanese guidelines but without statistical significance (P = .38). The study heterogeneity was slightly decreased on subgroup analysis.

Our study demonstrated poor performance of DESH in predicting treatment response with the pooled sensitivity and specificity.
The calculated PPV and NPV in our study were 80% and 80% of DESH based on its high PPV in treatment-response evaluation. However, NPV was not provided by this study. Since then, a few recent studies have presented promising results on the quantification of DESH based on the scoring system or the ratio between subarachnoid spaces of the Sylvian fissure and the vertex sulci. These quantification methods need to be validated.

Our study has several limitations. First, only 10 studies were included. Because DESH is a relatively recent imaging marker of iNPH, studies on the performance of DESH remain scarce. Nevertheless, >800 patients with iNPH were included in our study. Second, differences existed between the studies with regard to time points and methods of treatment-response evaluation. These may have affected the study heterogeneity. However, study heterogeneity remained unchanged in sensitivity analyses. A slight improvement in diagnostic performance was seen when performing the sensitivity analysis after excluding 3 studies that used different time intervals for the outcome measurement. One of the excluded studies that reported a substantially lower sensitivity of DESH used a very long time interval (mean, 51.9 months). In contrast, another study that used a very short time interval (0.3 months) demonstrated a substantially lower specificity of DESH.

We assumed that the follow-up period might have affected the diagnostic performance of DESH, given that that symptom improvements were most remarkable in the short term, with a gradual decline in the long term. Our results suggest the need for a large multicenter study including patients with and without DESH subjected to shunt surgery and with outcomes evaluated in a standardized manner.

Third, potential selection bias may be present in the study. Some of the studies used positive findings on the CSF tap tests for the selection of surgical candidates, which might have caused selection bias because patients with negative CSF tap test results could also benefit from VP shunt surgery. Moreover, 3 studies demonstrating an association between DESH and favorable postsurgical outcome were excluded because the sensitivity and specificity could not be calculated. This issue might have resulted in the underestimation of the diagnostic performance of DESH in our study. Finally, patients with incomplete DESH were excluded from the analysis. Despite an additional search performed to incorporate incomplete DESH, no eligible article was found. A previous study reported that patients satisfying DESH features except for a dilated Sylvian fissure showed a higher improvement rate after VP shunt surgery than patients satisfying DESH features except for the tight high convexity (87.5% versus 27.2%). Although the evidence is lacking, this finding may suggest that tight high convexity is a more important prognostic factor than the other features of DESH.
CONCLUSIONS

Our meta-analysis demonstrates a relatively low prevalence of DESH among patients with iNPH and a poor diagnostic performance of DESH with regard to treatment-response prediction.

REFERENCES

32. Suh CH, Park SH. Successful publication of systematic review and meta-analysis of studies evaluating diagnostic test accuracy. Korean J Radiol 2016;17:5–6 CrossRef Medline
34. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol* 2005;58:882–93 CrossRef Medline


