J Neurosurg 94:202–210, 2001

Predicting shunt failure on the basis of clinical symptoms and signs in children

Hugh J. L. Garton, M.D., M.H.Sc., John R. W. Kestle, M.D., M.Sc., and James M. Drake, F.R.C.S.(C)

Section of Neurosurgery, Department of Surgery, University of Michigan School of Medicine, Ann Arbor, Michigan; Division of Pediatric Neurosurgery, Primary Children's Medical Center, University of Utah, Salt Lake City, Utah; and Division of Pediatric Neurosurgery, Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada

Object. In evaluating pediatric patients for shunt malfunction, predictive values for symptoms and signs are important in deciding which patients should undergo an imaging study, whereas determining clinical findings that correlate with a low probability of shunt failure could simplify management.

Methods. Data obtained during the recently completed Pediatric Shunt Design Trial (PSDT) were analyzed. Predictive values were calculated for symptoms and signs of shunt failure. To refine predictive capability, a shunt score based on a cluster of signs and symptoms was derived and validated using multivariate methods.

Four hundred thirty-one patient encounters after recent shunt insertions were analyzed. For encounters that took place within 5 months after shunt insertion (early encounters), predictive values for symptoms and signs included the following: nausea and vomiting (positive predictive value [PPV] 79%, likelihood ratio [LR] 10.4), irritability (PPV 78%, LR 9.8), decreased level of consciousness (LOC) (PPV 100%), erythema (PPV 100%), and bulging fontanelle (PPV 92%, LR 33.1). Between 9 months and 2 years after shunt insertion (late encounters), only loss of developmental milestones (PPV 83%, LR 36.7) and decreased LOC (PPV 100%) were strongly associated with shunt failure. However, the absence of a symptom or sign still left a 15 to 29% (early encounter group) or 9 to 13% (late encounter group) chance of shunt failure. Using the shunt score developed for early encounters, which sums from 1 to 3 points according to the specific symptoms or signs present, patients with scores of 0, 1, 2, and 3 or greater had shunt failure rates of 4%, 50%, 75%, and 100%, respectively. Using the shunt score derived from late encounters, patients with scores of 0, 1, and 2 or greater had shunt failure rates of 8%, 38%, and 100%, respectively.

Conclusions. In children, certain symptoms and signs that occur during the first several months following shunt insertion are strongly associated with shunt failure; however, the individual absence of these symptoms and signs offers the clinician only a limited ability to rule out a shunt malfunction. Combining them in a weighted scoring system improves the ability to predict shunt failure based on clinical findings.

Key Words • sensitivity • specificity • predictive value • likelihood ratio • clinical prediction • shunt malfunction

Evaluation of patients with possible CSF shunt failure is a common practice in pediatric neurosurgery, given the 39% 1-year and 53% 2-year failure rates that have been observed following initial shunt placement.[5] Predictive values for the signs and symptoms of shunt failure are important in deciding which patient should undergo an imaging study and/or should be transferred to a facility with neurosurgical expertise. Likewise, recognition of clinical factors that correlate with a low probability of shunt failure could simplify management.

A number of prior publications have addressed the use of CT scanning, magnetic resonance imaging, radionucleotide, or iodinated contrast studies; shunt taps; Doppler ultrasonography; and intracranial pressure monitoring to aid in the diagnosis of patients with shunt malfunction.[3,7,9,13,14,16,21,26–28] Fewer studies have concentrated on the symptoms and signs of shunt malfunction that are available from the history compilations and physical examinations in such patients, and most of these are descriptive.[1,6,10,12,24] Predictive values that can be used to estimate the likelihood of malfunction in the presence or absence of a clinical finding[22] are very informative, but are considered more rarely. Watkins and colleagues[29] and Piatt[19,20] evaluated the predictive value of clinical findings, concentrating on one or two symptoms, signs, or results of diagnostic studies. The clinician, however, is simultaneously presented with a number of these factors from which to shape impressions as to the likelihood of shunt failure. To assess the predictive ability of the clinical presentation of patients with possible shunt malfunction, we evaluated data collected during the PSDT. Our goal was to determine predictive values for individual symptoms and signs

Abbreviations used in this paper: CI = confidence interval; CSF = cerebrospinal fluid; CT = computerized tomography; LOC = level of consciousness; LR = likelihood ratio; NPV = negative predictive value; PPV = positive predictive value; PSDT = Pediatric Shunt Design Trial; ROC = receiver-operating characteristic.

and, if possible, to identify combinations of these that might improve predictive ability.

Clinical Material and Methods

Pediatric Shunt Design Trial

Data were derived from the PSDT. In this trial, 344 children with newly diagnosed hydrocephalus were randomized into groups to receive one of three types of shunt valves (a standard differential pressure valve, a Delta Valve [PS Medical–Medtronic, Goletta, CA], or a Sigma Valve [Cordis, Miami, FL]) and were followed for a minimum of 1 year. Follow-up examinations were performed at 3, 12, 24, and 36 months after shunt insertion, as well as whenever clinically indicated. The details of the study protocol of this prospective multicenter randomized controlled trial have been published elsewhere.[4] At each follow-up visit, symptoms, signs, and test results were recorded in detail. Data forms, clinical notes, and results of imaging studies were forwarded to the trial methodology center. Shunt function or failure was determined by the evaluating surgeon based on standardized criteria and was reviewed by an adjudication committee blinded to the randomization status of the patient.

Definitions and Exclusions

In the primary analysis of the PSDT, the presence or absence of shunt failure was assessed using all available information including symptoms, signs, and results of imaging studies. In some cases the symptoms and/or signs were the determining factors in the assessment of whether there was shunt failure and, rarely, they were the sole basis for the decision (if imaging study results were unavailable). For example, if a child presented with ventricles that had failed to decrease in size compared with findings of the preoperative study, the determination of the status of the shunt at that particular visit depended solely on the presence or absence of symptoms and signs. In the present study we excluded such cases and based our analysis on those patients for whom objective measures of shunt function were available. For this analysis a diagnosis of shunt failure was given if at least one of the following conditions was present: 1) obstruction—ventricular enlargement compared with results of a baseline study, obvious migration or disruption of the catheter system demonstrated on an imaging study, or an obvious persistent CSF leak from the wound; 2) overdrainage—the presence of subdural fluid collections revealed on an imaging study; 3) loculation—the presence of new, loculated compartments within the ventricular system; and 4) infection—a positive CSF culture, purulent discharge from the wound, or erosion of the shunt system.

Shunt success was defined during the early postoperative period (< 5 months after shunt insertion) as a decrease in ventricle size compared with results of a previous imaging study along with the lack of any features meeting the definition of shunt failure. During later follow up (> 9 months after shunt insertion), stable ventricle size (unchanged from a study in which researchers had previously documented decreased ventricle size) was accepted as proof of adequate shunt function. If no imaging study had been performed at the particular follow-up visit, but a subsequent study demonstrated smaller ventricles, or if the patient had no evidence of shunt failure for more than 1 year after the time of the particular visit, this was accepted as sufficient evidence of shunt function at the time of that visit.

Data Collection

We reviewed each patient encounter to assess whether the patient met the criteria of shunt failure or success at each visit. Ventricular size comparisons were made using a measured frontal/occipital horn ratio.[17] In some cases in which these measurements were not available, the study clinician's judgment concerning results demonstrated on imaging (same, larger, or smaller ventricle size) was accepted. If no determination could be made, the encounter was excluded from consideration.

Because it is recognized that the clinical presentation of shunt failure is different for patients in whom shunt insertion was performed recently than for those in whom it was performed longer ago, we divided the follow-up visits into two epochs.[11,18] Early encounters included those visits occurring within the first 5 months following shunt insertion, and were clustered mostly around the planned 3-month follow-up visit. Late encounters included visits occurring 9 months or more after the time of shunt insertion, and were clustered mostly around the planned 1-year follow-up visit. In cases in which duplicate visits occurred within the same period, only the first visit was included.

Statistical Analysis

Occurrences of particular symptoms and signs were tabulated against the fate of the shunt at each visit to calculate sensitivity, specificity, PPVs, NPVs, LRs, and associated CIs.[25,30] The PPV represents the percentage of those patients presenting with a particular symptom who, in fact, have a shunt failure. The NPV is the percentage of those patients without a particular symptom who do not, in fact, have a shunt failure. The quantity $1 - \text{NPV}$ gives the percentage of patients who, despite absence of the symptom, nevertheless have a shunt failure. Likelihood ratios express the odds that a symptom is present in a patient with shunt failure compared with its presence in a patient without shunt failure.[22]

To assess the ability of a combination of symptoms and signs to improve prediction of shunt success or failure, the encounters with patients were randomly divided into two sets, a derivation set for developing the model and a validation set for testing the model. Using only the derivation set of encounters, factors identified on the basis of a univariate analysis and clinical experience were entered into a multivariate logistic regression model to identify independent predictors of outcome and to develop a regression model. The β coefficients from the model were then assigned as weights to the predictive factors.[2] Whenever necessary, recursive partitioning methods were used to improve the efficacy of the developed model.[15] For each patient encounter, a total shunt score was calculated as the sum of the weighting factors for each symptom and sign present. Data obtained in patients with clinical factors that appeared only in cases of shunt failure were excluded from the regression analysis to avoid destabilizing the regression model, but these clinical factors were included in

TABLE 1
Analysis of eligible encounters and end point assessment

Component	Encounter Group (no. of encounters)	
	Early	Late
no. of patients followed up in PSDT	344*	189†
not seen during interval or no follow-up form	44	7
evaluated patients	300	182
excluded patients	24	27
survival by absence of symptoms & signs alone	7	8
no imaging studies & unable to confirm survival >1 yr	5	18
failure based on symptoms alone	12	1
successful shunts	203	137
ventricles smaller (measured)	128	62
ventricles smaller (clinician's report)	58	16
no image obtained, ventricle smaller on later image	14	7
clinically stable for >1 yr after visit	3	10
stable on scan after prior reduction	0	42
failed shunts	73	18
ventricles larger (measured)	11	5
ventricles larger (clinician's report)	15	8
migration/malposition/disruption	13	2
CSF leak	1	0
overdrainage/subdural fluid collection	7	3
loculation (independent of larger ventricles)	1	0
infection	25	0

* Original cohort in the trial.
† Members of original cohort remaining in trial after 9 months of follow up.

the developed shunt score with an assigned weight. Colinear variables, as observed for some symptoms and signs of infection, were simplified to a single representative factor. The logistic regression analysis was performed using a forward conditional method with an entry level of significance of 0.05. Goodness of fit of the logistic regression was tested using the Hosmer–Lemeshow statistic. Statistical analysis was performed using commercially available statistical computer software (SPSS, version 8.0 for Windows; SPSS, Inc., Chicago, IL).

The validity of the scoring system was assessed by comparing shunt failure rates for each level of the total shunt score in the validation set of encounters. The reliability of the model was determined by comparing the scoring system's predictive properties in the two datasets, whereas the discriminatory power of the model was assessed using ROC curve analysis to calculate the area under the curve and associated CIs. The analytical protocol was repeated for both early and late encounter groups.

Results

During the early period, by applying the objective definition of shunt failure we were able to assess data in 300 of the 344 patients who had initially been randomized. During this interval, in 44 patients either no encounter occurred or no follow-up form was obtained at the time of shunt failure. Of the 300 patients who were evaluated, 24 patients had to be excluded because the fates of their shunts had been determined exclusively or primarily on the basis of presenting symptoms and signs. Of the remaining 276 patients, shunts were functioning in 203 and had failed in 73. On average, therefore, children presenting for evaluation had a 26% chance of shunt failure.

During the late period, among the 189 patients in the trial whose shunts functioned longer than 9 months, there were no follow-up data forms for seven patients and shunt fate could not be confirmed by objective criteria in 27 patients. In the remaining 155 patients, there were 137 successful shunts and 18 failed shunts, for a 12% failure rate (Table 1).

Clinical Characteristics

The study population consisted principally of very young children, which is not surprising given the requirement that participants had to have undergone their first shunt placement. For the group of early encounters, the median patient age at the visit, correcting for prematurity, was 143 days (4.6 months) and the mean ± standard deviation was 610 ± 1180 days. The median patient age at the visit in the late encounter group was 502 days (16.5 months) and the mean ± standard deviation was 897 ± 1056 days. The clinical features of patients participating in this analysis are presented in Table 2. In this table, patient age at the visit has been dichotomized to older or younger than 15 months to assist in assessing the prevalence of age-specific symptoms such as headache, fontanelle tension, and school performance.

During the early patient encounters, among the routinely assessed symptoms, nausea and vomiting, irritability, bulging fontanelle, and increased head circumference all occurred at rates exceeding 10%. School performance was assessed less frequently in this group, which was expected given the age distribution of the patients. Surgeons in the trial assessed shunt valve performance (based on the questions "Was the shunt reservoir depressable?" and "Did it refill?") in only approximately 50% of the visits because of variations in examination procedures or shunt valve design or both. During late patient encounters, symptoms and signs were generally less prevalent. Among routinely assessed symptoms only nausea and vomiting, headache, irritability, and increased head circumference occurred at rates of approximately 5% or more, whereas assessment of school performance and valve function were reported less commonly.

Predictive Values

In the early encounter group (Table 3), irritability and nausea and vomiting were associated with PPVs greater than 70%. Among clinical signs, several occurred only in patients with failing shunts, and there were no false-positive occurrences. These included decreased LOC in seven patients, papilledema in two, erythema around the shunt in seven, and peritonitis in four, as well as several other features typical of infection. For these factors, the PPV was 100% and the LR could not be calculated. Similarly, bulging fontanelle and fluid tracking around the shunt were highly predictive of shunt failure. Fever (PPV 89%) was also highly predictive of failure and in each instance occurred in the infection category. No factors were individually successful in predicting the absence of shunt failure. The probability of a shunt malfunction, even when a particular symptom was absent (1 − NPV), ranged from 17

TABLE 2
*Characteristics of postop encounters**

Clinical Parameter	No. of Encounters (%)	
	Early Encounter Group	Late Encounter Group
no. of patients	276	155
patient age at visit		
<15 mos	216 (78.0)	62 (40.0)
>15 mos	60 (28.0)	93 (60.0)
cause of hydrocephalus		
IVH	67 (24.3)	48 (31.0)
MMC	56 (20.3)	21 (13.5)
tumor	25 (9.1)	15 (9.7)
AS	19 (6.9)	10 (6.5)
unknown & other	109 (39.5)	61 (39.4)
presenting symptom or sign		
nausea & vomiting	33 of 274 (12.0)	7 of 154 (4.5)
headache	14 of 232 (6.0)	9 of 149 (6.0)
irritability	36 of 274 (13.1)	8 of 154 (5.2)
LDM	10 of 257 (3.9)	6 of 150 (4.0)
worsening of SP	2 of 190 (1.1)	1 of 111 (0.9)
new or changed seizures	7 of 271 (2.6)	4 of 152 (2.6)
bulging fontanelle	39 of 267 (14.6)	0 of 123 (0.0)
decreased LOC	7 of 272 (2.6)	2 of 153 (1.3)
papilledema	2 of 221 (0.9)	1 of 132 (0.8)
fluid tracking around shunt	20 of 275 (7.3)	3 of 154 (1.9)
sixth cranial nerve palsy	5 of 270 (1.9)	2 of 154 (1.3)
loss of upgaze	6 of 267 (2.2)	0 of 152 (0.0)
unable to depress reservoir	5 of 152 (3.3)	1 of 56 (1.8)
reservoir does not refill	7 of 151 (4.6)	4 of 56 (7.1)
increased head circumference	45 of 250 (18.0)	15 of 128 (11.7)
abdominal pain	5 of 274 (1.8)	0 of 155 (0.0)
nuchal rigidity	2 of 269 (0.7)	0 of 155 (0.0)
erythema	7 of 276 (2.5)	0 of 155 (0.0)
abdominal mass	1 of 274 (0.4)	0 of 155 (0.0)
meningismus	2 of 275 (0.7)	0 of 155 (0.0)
peritonitis	4 of 274 (1.5)	0 of 155 (0.0)
fever	19 of 276 (6.9)	1 of 155 (0.6)

* The early encounter group encompasses patients whose follow-up visits occurred less than 5 months after shunt placement, and the late encounter group includes patients whose follow-up visits occurred more than 9 months after shunt placement. Numbers of patients for whom yes/no answer was available are given; data for remaining patients were either not reported or reported as unknown. Abbreviations: AS = aqueductal stenosis; IVH = intraventricular hemorrhage; LDM = loss of developmental milestones; MMC = myelomeningocele; SP = school performance.

to 20%, an insufficient reduction from the baseline 26% failure rate.

In the late encounter group (Table 3), fewer symptoms and signs were predictive of outcome. Among factors occurring more than once, only loss of developmental milestones and decreased LOC had predictive values for failure that were higher than 70%. Again, even in the absence of a particular symptom or sign, there remained roughly a 10 to 14% chance of shunt failure, offering little discrimination from the 12% baseline rate.

Regression Analysis

Early Encounter Group. In the derivation subset of 159 patients (Table 4), symptoms and signs significantly associated with shunt failure ($p < 0.05$) in the univariate assessment included decreased LOC, irritability, bulging fontanelle, headache, fever, nausea and vomiting, increased head circumference, and fluid tracking around the shunt. Erythema, meningismus, abdominal pain, and peritonitis were also strongly statistically associated with shunt failure, although meningismus, abdominal pain, and peritonitis occurred only when erythema was also present, rendering these variables colinear with erythema and, thus, excluded from further analysis. Patients with decreased LOC and/or erythema always experienced shunt failures, and, thus, these clinical factors were included in the shunt score with an arbitrary weighting factor of 3. The remaining factors significant in the univariate analysis were entered into the logistic regression analysis: irritability, bulging fontanelle, headache, fever, nausea and vomiting, increased head circumference, and fluid tracking around the shunt. Also included were the clinically chosen variables: patient age at visit (as a continuous variable), cause of hydrocephalus, and valve type. Irritability, bulging fontanelle, fever, headache, and fluid tracking around the shunt were all independent predictors, whereas patient age, cause of hydrocephalus, and valve type did not significantly influence the associations. Whenever it could be calculated, the logistic regression model had an excellent goodness of fit (Hosmer–Lemeshow test: $\chi^2 = 0.0038$, $p = 0.9981$ [four of 10 deciles calculable]). These

TABLE 3

*Predictive values in early and late encounter groups**

Clinical Parameter	PPV (%)		1-NPV (%)	LR		LR Neg	Sensitivity (%)	Specificity (%)	No. of Encounters†			
	Value	95% CI		Ratio	95% CI				TP	FP	FN	TN
early encounter group												
nausea & vomiting	79	61–91	19	10.4	4.7–23.0	0.7	36	97	26	7	46	195
headache	57	29–82	21	4.4	1.6–12.1	0.9	15	97	8	6	46	172
irritability	78	61–90	18	9.8	4.7–20.5	0.6	39	96	28	8	44	194
LDM	20	3–56	25	0.8	0.2–3.5	1.0	3	96	2	8	62	185
worsening of SP	50	1–99	29	2.4	0.2–37.6	1.0	2	99	1	1	55	133
new or changed seizures	0	0–41	27	0.0	—	1.0	0	96	0	7	72	192
bulging fontanelle	92	79–98	15	33.1	10.5–104.2	0.5	51	98	36	3	35	193
decreased LOC	100	59–100	25	NR	—	0.9	10	100	7	0	66	199
papilledema	100	16–100	27	NR	—	1.0	3	100	2	0	59	160
fluid tracking around shunt	75	51–91	23	8.3	3.1–22.0	0.8	21	98	15	5	58	197
sixth cranial nerve palsy	40	5–85	25	1.9	0.3–11.4	1.0	3	99	2	3	67	198
loss of upgaze	50	12–88	24	3.0	0.6–14.7	1.0	5	99	3	3	63	198
unable to depress reservoir	80	28–99	21	13.4	1.5–115.8	0.9	11	99	4	1	31	116
reservoir does not refill	86	42–100	20	19.9	2.5–159.7	0.8	17	99	6	1	29	115
increased head circumference	67	51–80	17	5.7	3.3–9.9	0.6	46	92	30	15	35	170
abdominal pain	100	48–100	25	NR	—	0.9	7	100	5	0	68	201
nuchal rigidity	50	1–99	26	2.8	0.2–44.8	1.0	1	99	1	1	69	198
erythema	100	59–100	25	NR	—	0.9	10	100	7	0	66	203
abdominal mass (pseudocyst)	100	2–100	26	NR	—	1.0	1	100	1	0	72	201
meningismus	100	16–100	26	NR	—	1.0	3	100	2	0	70	203
peritonitis	100	40–100	26	NR	—	0.9	5	100	4	0	69	201
fever	89	67–99	22	23.6	5.6–99.8	0.8	23	99	17	2	56	201
late encounter group												
nausea & vomiting	43	10–82	10	5.7	1.4–23.3	0.9	17	97	3	4	15	132
headache	22	3–60	11	2.2	0.5–9.8	0.9	12	95	2	7	15	125
irritability	25	3–65	11	2.5	0.5–11.5	0.9	11	96	2	6	16	130
LDM	83	36–100	9	36.7	4.5–296.4	0.7	28	99	5	1	13	131
worsening of SP	100	2–100	13	NR	—	0.9	7	100	1	0	14	96
new or changed seizures	0	0–60	12	NR	—	1.0	0	97	0	4	18	130
bulging fontanelle	0	—	11	NR	—	1.0	0	100	0	0	13	110
decreased LOC	100	16–100	11	NR	—	0.9	11	100	2	0	16	135
papilledema	100	2–100	12	NR	—	0.9	6	100	1	0	16	115
fluid tracking around shunt	33	1–91	11	3.8	0.4–39.6	1.0	6	99	1	2	17	134
sixth cranial nerve palsy	50	1–99	11	7.6	0.5–115.6	1.0	6	99	1	1	17	135
loss of upgaze	0	—	12	NR	—	1.0	0	100	0	0	18	134
unable to depress reservoir	100	2–100	11	NR	—	0.9	14	100	1	0	6	49
reservoir does not refill	50	7–93	10	7.0	1.2–42.0	0.7	29	96	2	2	5	47
increased head circumference	27	8–55	10	2.7	1.0–7.5	0.8	27	90	4	11	11	102
abdominal pain	0	—	12	NR	—	1.0	0	100	0	0	18	137
nuchal rigidity	0	—	12	NR	—	1.0	0	100	0	0	18	137
erythema	0	—	12	NR	—	1.0	0	100	0	0	18	137
abdominal mass (pseudocyst)	0	—	12	NR	—	1.0	0	100	0	0	18	137
meningismus	0	—	12	NR	—	1.0	0	100	0	0	18	137
peritonitis	0	—	12	NR	—	1.0	0	100	0	0	18	137
fever	0	0–98	12	NR	—	1.0	0	99	0	1	18	136

* LR = likelihood ratio for shunt failure in presence of symptom; LR Neg = likelihood ratio for shunt failure in absence of symptom; NR = no result, division by zero; — = not applicable.

† Number of encounters in which there were false-negative (FN), false-positive (FP), true-negative (TN), and true-positive (TP) results.

five symptoms and signs were assigned weights based on their regression coefficients (Table 5). Summing these weights as well as those for decreased LOC and erythema for each symptom or sign present at a patient encounter, the clinical scores ranged from 0 to 8 in the derivation set. For example, a child with a decreased LOC and a bulging fontanelle scored a 3 + 2 = 5 total score (Table 5).

In the validation set of 118 patients, the overall risk of shunt failure was 28%. The risks of shunt failure associated with scores of 0, 1, 2, or 3 or more were 4% (three of 79), 50% (five of 10), 75% (nine of 12), and 100% (15 of 15), respectively. The ROC curve analysis demonstrated a high degree of homology between the derivation and validation sets, with the area under the curve equal to 0.95 (95% CI 0.92–0.99) for the derivation set and 0.93 (95% CI 0.88–0.98) for the validation set (Fig. 1).

Late Encounter Group. In the derivation subset of 100 patients, loss of developmental milestones and nausea and vomiting were the only variables significantly associated with shunt failure (Table 4). In the logistic regression analysis, which included these factors and patient age at visit,

TABLE 4

Results of univariate analysis, derivation set only, in the early and late encounter groups

Clinical Parameter	No. of Shunt Failures		No. of Shunt Survivals		Odds Ratio		p Value	MA Status*
	Present	Assessed	Present	Assessed	Ratio	95% CI		
early encounter group								
nausea & vomiting	17	41	5	118	16.0	5.3–47.6	0.000	I
headache	7	30	3	108	10.6	2.5–44.3	0.001	I
irritability	17	40	6	118	13.8	4.9–38.7	0.000	I
LDM	1	33	5	114	0.7	0.1–6.0	1.000	—
worsening of SP	1	29	1	82	2.9	0.2–47.8	0.456	—
new or changed seizures	0	40	4	117	0.0	—	0.573	—
bulging fontanelle	20	40	1	107	106.0	13.5–835.4	0.000	I
decreased LOC	4	41	0	116	—	—	0.004	E
papilledema	1	34	0	98	—	—	0.258	—
fluid tracking around shunt	7	41	4	118	5.9	1.6–21.3	0.009	I
sixth cranial nerve palsy	1	38	2	118	1.6	0.2–17.8	0.570	—
loss of upgaze	1	35	1	118	3.4	0.2–56.5	0.406	—
unable to depress reservoir	1	16	1	69	4.5	0.3–76.6	0.343	—
reservoir does not refill	2	17	1	68	8.9	0.7–105.1	0.101	—
increased head circumference	17	37	9	110	9.5	3.7–24.4	0.000	I
abdominal pain	4	41	0	117	—	—	0.004	E
nuchal rigidity	1	38	1	115	3.1	0.2–50.5	0.436	—
erythema	4	41	0	118	—	—	0.004	E
abdominal mass (pseudocyst)	1	41	0	117	—	—	0.259	—
meningismus	2	40	0	118	—	—	0.063	—
peritonitis	3	41	0	117	—	—	0.017	E
fever	9	41	1	118	32.9	4.0–269.4	0.000	I
late encounter group								
nausea & vomiting	2	11	1	89	19.6	1.6–237.4	0.031	I
headache	1	10	5	87	1.8	0.2–17.4	0.489	—
irritability	1	11	3	88	2.8	0.3–29.9	0.377	—
LDM	3	11	1	86	31.9	3.0–343.2	0.004	I
worsening of SP	1	8	0	64	—	—	0.111	—
new or changed seizures	0	11	4	87	0.0	—	1.000	—
bulging fontanelle	0	9	0	74	—	—	—	—
decreased LOC	1	11	0	88	—	—	0.111	—
papilledema	0	10	0	77	—	—	—	—
fluid tracking around shunt	1	11	1	89	8.8	0.5–151.8	0.209	—
sixth cranial nerve palsy	0	11	0	89	—	—	—	—
loss of upgaze	0	11	0	87	—	—	—	—
unable to depress reservoir	1	5	0	33	—	—	0.132	—
reservoir does not refill	1	5	2	33	3.9	0.3–53.1	0.353	—
increased head circumference	3	10	10	73	2.7	0.6–12.2	0.187	—
abdominal pain	0	11	0	89	—	—	—	—
nuchal rigidity	0	11	0	89	—	—	—	—
erythema	0	11	0	89	—	—	—	—
abdominal mass (pseudocyst)	0	11	0	89	—	—	—	—
meningismus	0	11	0	89	—	—	—	—
peritonitis	0	11	0	89	—	—	—	—
fever	0	11	1	89	0.0	—	0.724	—

* Status of variable in multivariate analysis (MA); E = excluded from the multivariate analysis because 100% predictive or always encountered with finding also excluded from multivariate regression; I = included in multivariate regression. See *Regression Analysis* for full explanation.

cause of hydrocephalus, and shunt valve type, both symptoms remained independently associated with shunt failure, whereas patient age, cause of hydrocephalus, and shunt valve type did not influence the association. However, the scoring system, in which a weight of 1 was assigned to each symptom and a cut point of greater than 0 indicated shunt failure, detected only 54% of the shunt failures (that is, with a score > 0). Based on recursive partitioning methods and clinical judgment the following factors were added: decreased LOC (assigned weight of 3), fluid tracking around the shunt (weight of 1), and increased head circumference (weight of 1) (Table 5). Clinical scores ranged from 0 to 4 in the derivation set, and 73% of the shunt failures were associated with scores higher than 0.

In the validation set of 55 patients, the risks of shunt failure associated with scores of 0, 1, or 2 or more were 8% (three of 46), 38% (three of eight), and 100% (one of one), respectively. The ROC curve analysis again demonstrated a high degree of homology between derivation and validation sets, with the area under the curve equal to 0.8 (95% CI 0.68–0.92) for the derivation set and 0.74 (95% CI 0.58–0.9) for the validation set (Fig. 1).

Sensitivity Assessment

Because of the risk of a missed diagnosis and to identi-

TABLE 5
*Multivariate analysis and symptom score**

Variable	Coefficient†	SE	OR (95% CI)	p Value	Weight‡
early encounter group					
fluid tracking around shunt	2.61	1.29	13.6 (1.1–173.2)	0.0443	1
headache	3.57	1.145	35.8 (3.8–339.3)	0.0018	1
irritability	3.80	0.929	44.6 (7.2–275.5)	<0.001	1
fever	4.70	1.392	110.3 (7.2–275.5)	<0.001	1
bulging fontanelle	6.08	1.262	439.4 (37.0–5216.1)	<0.001	2
erythema	—	—	—	—	3
decreased LOC	—	—	—	—	3
late encounter group					
nausea & vomiting	3.37	1.30	29.0 (2.2–367.2)	0.009	1
LDM	3.77	1.23	43.4 (3.9–484.0)	0.002	1
increased head circumference	—	—	—	—	1
fluid tracking around shunt	—	—	—	—	1
decreased LOC	—	—	—	—	3

* OR = odds ratio; SE = standard error of the coefficient.
† The coefficient, which was determined by logistic regression, represents the increase in the log of the odds of shunt failure for the presence or absence of a given symptom while holding all other variables constant.
‡ Each coefficient was divided by 2.61 and rounded down to the nearest whole number. A patient's total score is determined by summing the weights for each symptom present. Therefore, a patient presenting during the first 5 months after shunt insertion with a bulging fontanelle and fluid tracking around the shunt would have a score of 1 + 2 = 3.

fy limitations in the scoring system, we assessed its sensitivity for shunt failure across the entire dataset. Based on the ROC curve, a score of 1 or more was chosen to predict shunt failure. In the early group, five (7%) of 73 patients with failures had symptom scores of 0 and true shunt status would not have been detected by the scoring system. Two of these patients had increased head circumference and in one the shunt reservoir could not be depressed. The other two patients had no recorded symptoms or signs. In one of them loculated ventricles and in the other a subdural effusion were revealed on CT scans at the time of the follow-up visits.

Performance in the late encounter group was not as good. Again for the entire dataset, six (33%) of the 18 patients with shunt failures had scores of 0. One patient had a headache and in another there was no refilling of the shunt reservoir. In the remaining four patients no symptoms or signs were identified, but the shunt was shown to have failed on the basis of increased ventricle size.

Discussion

In this study, follow-up data collected during the PSDT were used to determine the incidence and predictive values of individual symptoms and signs of shunt malfunction and to find a combination of those most effective for predicting malfunction.

Our data demonstrate that the common symptoms of shunt malfunction are indeed strongly predictive of shunt failure during the early period following shunt insertion, which we arbitrarily defined as less than 5 months after shunt insertion. Some factors, specifically decreased LOC and erythema around the wound, were always indicative of shunt failure (either obstruction or infection). In the late encounter group, arbitrarily defined as longer than 9 months after insertion, clinical factors were less predictive of failure, which coincides with a reduction in the rate of shunt failure. The predictive values determined in this analysis are generally higher than those reported by Watkins and colleagues.[29] Over a 5-month period those researchers studied 52 admissions (45 patients) to their institution in which the patients presented with a tentative diagnosis of shunt malfunction. In 19 (36%) of these hospital admissions, the patients were found in fact to have shunt malfunction. Assuming that at all admissions their patients were evaluated for all symptoms, the predictive values for the clinical features of malfunction were as follows: headache, PPV 30% (nine of 30), LR 0.74; irritability, PPV 47% (seven of 15), LR 1.51; decreased LOC, PPV 85% (11 of 13), LR 9.5; vomiting, PPV 36% (10 of 28), LR 1.2; and temperature, PPV 22% (two of nine), LR 0.49. These values are closer to those found in our later encounter group (Table 3). Decreased LOC was the factor most strongly predictive of failure in both their study and our early encounter group. The difference seen in prognostic values may be due to differences in the study populations, end point definitions, or methods of data collection.

Particular symptoms and signs have attracted attention in the literature. A number of authors have commented on the association or lack thereof between seizures and shunt malfunction. Although epilepsy is common in children with hydrocephalus, the literature indicates that seizures are rarely a presenting symptom in cases of shunt failure.[6,8,10,23,24] In the PSDT, seizures were a presenting symptom in only approximately 2.6% of the encounters. Johnson and colleagues[10] noted that 16 of 544 emergency encounters with patients with seizure disorders and shunted hydrocephalus culminated in shunt revisions. This yields a predictive value of 2.9%. In our study, in none of the 11 encounters in which seizures occurred was a shunt failure actually present, although the small number of patients presenting with seizures limits the strength of any conclusion based on this.

Piatt[19] evaluated the usefulness of "pumping the shunt" in determining the likelihood of malfunction and noted a sensitivity of 18 to 20% and a specificity of 63 to 93%. In

our analysis, considering the early encounter group, sensitivity was similar (11% for reservoir depression and 17% for reservoir refill); however, specificity was higher (99%) for the two parameters, meaning that there were very few false-positive test results. As a result the PPV for the reservoir failing to depress or refill was much higher in our study than in Piatt's study (80–86% compared with 17–21%). The predictive value of the absence of abnormality indicating shunt patency (NPV) was similar (79–80% compared with 65–81%). The patients in Piatt's study were older (median age 87 months) and had a lower overall shunt failure rate (14%). Our evaluation of this clinical sign in relation to others was limited by a lack of reports among a substantial number of patients, either because the information was not obtainable or because individual clinical practice did not include the assessment.

Ashkenazi and colleagues[1] reported a case series of 15 patients presenting with fever as the first sign of shunt malfunction, but did not provide sufficient information to determine the sign's predictive value. Watkins, et al.,[29] reported that fever was slightly protective and its presence lowered the probability of a shunt failure (overall rate of failure 36%, rate of failure associated with fever 22%). However, none of their patients suffered from shunt infection. In our study population, fever was not associated with shunt failure in the absence of infection, but occurred in 17 of 25 patients with infection in the early encounter group, raising the predictive value for shunt failure to the 89% reported in Table 3.

It would have been desirable to compare overall failure rates and predictive values in patients examined during routine follow-up visits compared with those seen on an emergency basis. However, the study data forms did not allow us accurately to distinguish the setting of the encounter.

Success of the Clinical Decision Rule

In attempting to rule out a shunt failure, the absence of an individual symptom did little to reduce the likelihood of shunt failure. By combining symptoms in a weighted scoring system, we improved our predictive ability. The presence of only one of several factors in the clinical scoring system was associated with a high enough rate of shunt failure in either the early or late encounter group to justify liberal use of imaging studies in patients with such a presentation. The absence of all symptoms in the scoring system produced a lower posttest probability of shunt failure, especially in the early encounter group, but not all shunt failures were detected by it. A number of these patients experienced ventricular enlargement without clinical symptoms, highlighting the value of routine postoperative imaging studies in this patient population. Whether such misclassifications would result in actual harm, or merely delay diagnosis, cannot be determined from this type of secondary data analysis. Nevertheless, we believe that misclassification rates of 7% and 33% for the early and late encounter groups, respectively, are too high, without a clear understanding of their effects.

Limitations of the Study

Information gathered in rigorous clinical trials, such as that which has been used in this analysis, must be carefully assessed for its relevance to the routine practice of medicine. If the predictive values seem above expectation, it must be remembered that the data were gathered on very young patients presenting to neurosurgeons for both routine and unscheduled follow-up examinations after a first shunt placement. Different providers, seeing a broader cross section of pediatric disease, might be expected to have a lower prevalence of children with shunt malfunction and, thus, lower predictive values for individual symptoms. However, the symptoms and signs themselves should retain their individual diagnostic properties (sensitivity, specificity, and LR).

Despite the fact that we evaluated a relatively large number of total encounters for rare symptoms, the num-

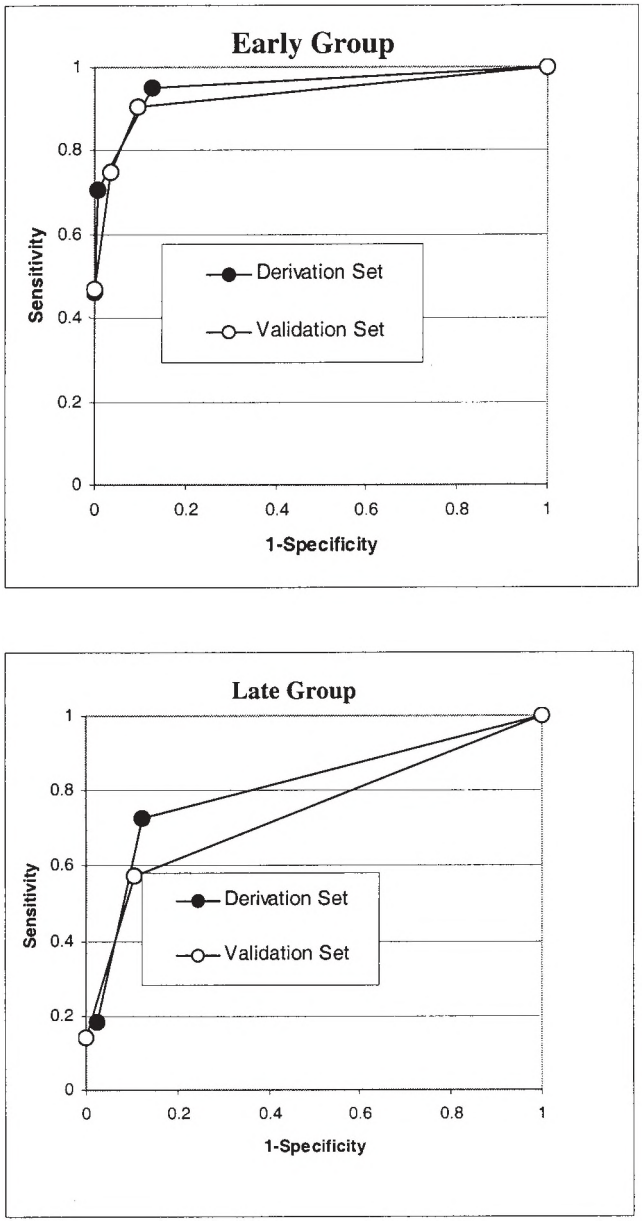

FIG. 1. Graphs demonstrating ROC curves for scoring systems. Comparisons between the derivation and validation sets are shown for both early and late encounter groups.

bers remained small for rare symptoms, producing wide CIs. Similarly, the combination of a smaller late encounter sample, due to prior shunt failures, and a reduction in the incidence of failures combined to reduce the number of shunt failures available for assessment in the late encounter group, with similar impact on the CIs.

Conclusions

Our analysis determined the predictive values of common symptoms and signs of shunt failure in very young children after their first shunt insertion. This was based on a secondary analysis of data from the recently completed PSDT. Individual symptoms and signs had high PPVs, but offered only limited ability effectively to rule out a shunt malfunction. Combining symptoms and signs to produce a weighted scoring system improved predictive ability. However, effectively ruling out the presence of a shunt malfunction remained problematic, particularly in the late encounter group. Using actual prevalence rates in the early and late follow-up periods of 26% and 12%, respectively, the likelihood of shunt malfunction in patients with a minimum shunt score was reduced to 4% and 8%, respectively. However, when even a single symptom or sign from the scoring system was present, the probability of a shunt failure was approximately 40%, justifying the liberal use of imaging studies or other objective assessments of shunt function in these patients.

References

1. Ashkenazi E, Umansky F, Constantini S, et al: Fever as the initial sign of malfunction in non infected ventriculoperitoneal shunts. **Acta Neurochir 114:**131–134, 1992
2. Barquet N, Domingo P, Caylà JA, et al: Prognostic factors in meningococcal disease. Development of a bedside predictive model and scoring system. **JAMA 278:**491–496, 1997
3. Benzel EC, Mirfakhraee M, Hadden TA: Evaluation of CSF shunt function: value of functional examination with contrast material. **AJNR 12:**143–147, 1991
4. Drake JM, Kestle J: Rationale and methodology of the multicenter pediatric cerebrospinal fluid shunt design trial. Pediatric Hydrocephalus Treatment Evaluation Group. **Childs Nerv Syst 12:**434–447, 1996
5. Drake JM, Kestle JRW, Milner R, et al: Randomized trial of cerebrospinal fluid shunt valve design in pediatric hydrocephalus. **Neurosurgery 43:**294–305, 1998
6. Faillace WJ, Canady AI: Cerebrospinal fluid shunt malfunction signaled by new or recurrent seizures. **Childs Nerv Syst 6:** 37–40, 1990
7. Fouyas IP, Casey ATH, Thompson D, et al: Use of intracranial pressure monitoring in the management of childhood hydrocephalus and shunt-related problems. **Neurosurgery 38:** 726–732, 1996
8. Hack CH, Enrile BG, Donat JF, et al: Seizures in relation to shunt dysfunction in children with meningomyelocele. **J Pediatr 116:**57–60, 1990
9. Jamjoom AHB, Wilson PJE: Misleading clinical syndromes of CSF shunt malfunction. **Br J Neurosurg 2:**391–394, 1988
10. Johnson DL, Conry J, O'Donnell R: Epileptic seizure as a sign of cerebrospinal fluid shunt malfunction. **Pediatr Neurosurg 24:**223–228, 1996
11. Kast J, Duong D, Nowzari F, et al: Time-related patterns of ventricular shunt failure. **Childs Nerv Syst 10:**524–528, 1994
12. Kirkpatrick M, Engleman H, Minns RA: Symptoms and signs of progressive hydrocephalus. **Arch Dis Child 64:**124–128. 1989
13. Mirfakhraee M, Benzel EC, Crofford MJ, et al: Metrizamide shuntography for evaluation of shunt malfunction in hydrocephalus. **AJNR 6:**815–822, 1985
14. Morgan MK, Johnston IH, Spittaler PJ: A ventricular infusion technique for the evaluation of treated and untreated hydrocephalus. **Neurosurgery 29:**832–836, 1991
15. Nelson LM, Bloch DA, Longstreth WT Jr, et al: Recursive partitioning for the identification of disease risk subgroups: a case-control study of subarachnoid hemorrhage. **J Clin Epidemiol 51:**199–209, 1998
16. Noetzel MJ, Baker RP: Shunt fluid examination: risks and benefits in the evaluation of shunt malfunction and infection. **J Neurosurg 61:**328–332, 1984
17. O'Hayon BB, Drake JM, Ossip MG, et al: Frontal and occipital horn ratio: a linear estimate of ventricular size for multiple imaging modalities in pediatric hydrocephalus. **Pediatr Neurosurg 29:**245–249, 1998
18. Piatt JH Jr: Cerebrospinal fluid shunt failure: late is different from early. **J Neurosurg 82:**363A, 1995 (Abstract)
19. Piatt JH Jr: Physical examination of patients with cerebrospinal fluid shunts: is there useful information in pumping the shunt? **Pediatrics 89:**470–473, 1992
20. Piatt JH Jr: Pumping the shunt revisited. A longitudinal study. **Pediatr Neurosurg 25:**73–77, 1996
21. Pople IK: Doppler flow velocities in children with controlled hydrocephalus: reference values for the diagnosis of blocked cerebrospinal fluid shunts. **Childs Nerv Syst 8:**124–125, 1992
22. Sackett D, Haynes R, Guyatt G, et al: **Clinical Epidemiology: A Basic Science for Clinical Medicine, ed 2.** Boston: Little, Brown and Company, 1991
23. Saukkonen AL, Serlo W, von Wendt L: Epilepsy in hydrocephalic children. **Acta Paediatr Scand 79:**212–218, 1990
24. Sekhar LN, Moossy J, Guthkelch AN: Malfunctioning ventriculoperitoneal shunts. Clinical and pathological features. **J Neurosurg 56:**411–416, 1982
25. Simel DL, Samsa GP, Matchar DB: Likelihood ratios with confidence: sample size estimation for diagnostic test studies. **J Clin Epidemiol 44:**763–770, 1991
26. Sood S, Kim S, Ham SD, et al: Useful components of the shunt tap test for evaluation of shunt malfunction. **Childs Nerv Syst 9:**157–162, 1993
27. Uvebrant P, Sixt R, Bjure J, et al: Evaluation of cerebrospinal fluid shunt function in hydrocephalic children using 99mTc-DTPA. **Childs Nerv Syst 8:**76–80, 1992
28. Vernet O, Farmer JP, Lambert R, et al: Radionuclide shuntogram: adjunct to manage hydrocephalic patients. **J Nucl Med 37:**406–410, 1996
29. Watkins L, Hayward R, Andar U, et al: The diagnosis of blocked cerebrospinal fluid shunts: a prospective study of referral to a paediatric neurosurgical unit. **Childs Nerv Syst 10:** 87–90, 1994
30. Zar JH: **Biostatistical Analysis, ed 3.** Upper Saddle River NJ: Prentice Hall, 1996

Manuscript received March 29, 2000.
Accepted in final form October 2, 2000.
Support for this study was provided to Dr. Kestle by the Medical Research Council of Canada, a Primary Children's Hospital Foundation grant, and Medtronic PS Medical/Cordis Corp.

Address reprint requests to: Hugh J. L. Garton, M.D., M.H.Sc., Section of Neurosurgery, Department of Surgery, University of Michigan Medical Center, 2128 Taubman HSC, 1500 East Medical Center Drive, Ann Arbor, Michigan 48109. email: hgarton@umich.edu.