PREDICTION OF OUTCOME IN TRAUMATIC BRAIN INJURY WITH COMPUTED TOMOGRAPHIC CHARACTERISTICS: A COMPARISON BETWEEN THE COMPUTED TOMOGRAPHIC CLASSIFICATION AND COMBINATIONS OF COMPUTED TOMOGRAPHIC PREDICTORS

Andrew I.R. Maas, M.D., Ph.D.

Department of Neurological Surgery,
Erasmus Medical Center,
Rotterdam, The Netherlands

Chantal W.P.M. Hukkelhoven, M.Sc.

Center for Clinical Decision Sciences,
Department of Public Health,
Erasmus Medical Center,
Rotterdam, The Netherlands

Lawrence F. Marshall, M.D.

Department of Neurological Surgery,
University of California,
San Diego, California

Ewout W. Steyerberg, Ph.D.

Center for Clinical Decision Sciences,
Department of Public Health,
Erasmus Medical Center,
Rotterdam, The Netherlands

Reprint requests:
Andrew I.R. Maas, M.D., Ph.D.,
Department of Neurosurgery,
Erasmus Medical Center
P.O. Box 2040,
3000 CA Rotterdam,
The Netherlands.
Email: airmaas@erasmusmc.nl

Received, January 21, 2005.
Accepted, May 20, 2005.

BACKGROUND AND OBJECTIVE: The Marshall computed tomographic (CT) classification identifies six groups of patients with traumatic brain injury (TBI), based on morphological abnormalities on the CT scan. This classification is increasingly used as a predictor of outcome. We aimed to examine the predictive value of the Marshall CT classification in comparison with alternative CT models.

METHODS: The predictive value was investigated in the Tirilazad trials (n = 2269). Alternative models were developed with logistic regression analysis and recursive partitioning. Six month mortality was used as outcome measure. Internal validity was assessed with bootstrapping techniques and expressed as the area under the receiver operating curve (AUC).

RESULTS: The Marshall CT classification indicated reasonable discrimination (AUC = 0.67), which could be improved by rearranging the underlying individual CT characteristics (AUC = 0.71). Performance could be further increased by adding intraventricular and traumatic subarachnoid hemorrhage and by a more detailed differentiation of mass lesions and basal cisterns (AUC = 0.77). Models developed with logistic regression analysis and recursive partitioning showed similar performance. For clinical application we propose a simple CT score, which permits a more clear differentiation of prognostic risk, particularly in patients with mass lesions.

CONCLUSION: It is preferable to use combinations of individual CT predictors rather than the Marshall CT classification for prognostic purposes in TBI. Such models should include at least the following parameters: status of basal cisterns, shift, traumatic subarachnoid or intraventricular hemorrhage, and presence of different types of mass lesions.

KEY WORDS: Computed tomography, Computed tomography classification, Outcome, Prognosis, Statistical models, Traumatic brain injury

Neurosurgery 57:1173-1182, 2005 DOI: 10.1227/01.NEU.0000186013.63046.6B www.neurosurgery-online.com

Classification of traumatic brain injury (TBI) is necessary to accurately describe patient series and requires grouping of patients according to specific characteristics. In clinical practice, the clinical severity of TBI is generally classified as severe, moderate or mild according to the level of consciousness as measured with the Glasgow Coma Scale (GCS). The increased use of early sedation, intubation and ventilation in more severe patients has decreased the value of the full GCS for purposes of classification (1, 4, 25). Alternatively, in more severe patients, TBI can be classified according to morphological criteria based on computed tomographic (CT) or magnetic resonance imaging (MRI) investigations. Although MRI may be more sensitive for detecting small white matter lesions in a later phase after TBI (9, 32), CT examination remains the investigation of choice in the acute phase.

Conventional classification of TBI with CT findings differentiates between focal and diffuse injuries (10, 22). In 1991 Marshall et al.

(24), after analysis of the Traumatic Coma Data Bank, proposed a CT classification for grouping patients with TBI according to multiple CT characteristics. This CT classification identifies six different groups of patients with TBI, based on the type and severity of several abnormalities on the CT scan. It differentiates between patients with and without mass lesions and permits a further discrimination of patients with diffuse injuries into four categories, taking into account signs of raised intracranial pressure (ICP; i.e., compressed or absent basal cisterns, midline shift). Since its introduction, this CT classification has become widely accepted for descriptive purposes, and is also increasingly being used as major predictor of outcome in TBI. Various studies have confirmed the predictive value of the CT classification (17, 21, 28), and the international guidelines on prognosis include the CT classification as a major CT predictor based on Class I evidence (5). Whether the Marshall CT classification is best suited for prediction or whether other combinations of CT parameters may be more appropriate for this specific purpose has not been investigated in detail.

The aim of the present study was to examine the prognostic performance of the Marshall CT classification in comparison with other combinations of CT predictors in TBI, by reevaluating and refining the CT characteristics used to determine this classification and by including additional CT parameters.

PATIENTS AND METHODS

The Marshall CT Classification

The Marshall CT classification is presented in *Table 1*. Discriminating features in this classification are 1) presence or absence of mass lesions, 2) presence or absence of intracranial abnormalities, 3) CT signs of raised intracranial pressure (status of basal cisterns, shift), and 4) planned evacuation of mass lesions.

To facilitate comparison with alternative classifications, we translated the Marshall CT classification into a binary tree (*Fig. 1*).

Patients

Our studies were conducted on the combined data sets of the International and North American Tirilazad trials (n = 2269). Details on the Tirilazad trials have been reported elsewhere (16, 23). Inclusion and exclusion criteria of the two trials were similar. Both trials included patients between 15 and 65 years of age with severe (GCS, 3–8) or moderate (GCS, 9–12) closed TBI. With respect to the CT characteristics the inclusion criteria varied slightly: the international study excluded moderate TBI patients with a normal CT scan, whereas the North American study excluded such patients only when the blood alcohol level exceeded 0.2 g/dl. Protocols and recommendations for management were comparable for both trials.

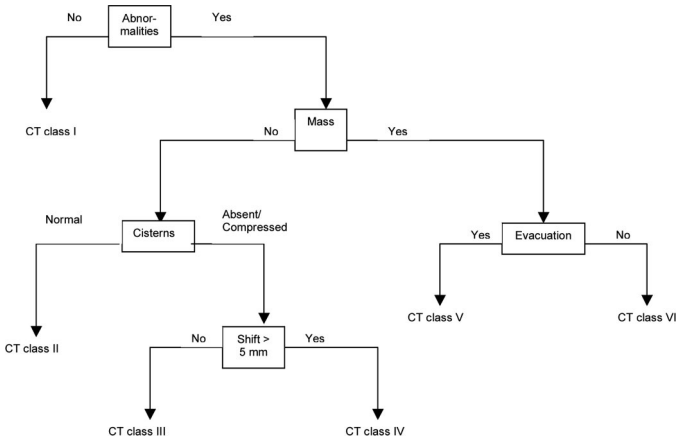

FIGURE 1. *Marshall CT classification presented in a tree structure.*

In both trials, the efficacy of Tirilazad mesylate, an aminosteroid that displays an antioxidant effect, was studied against that of placebo. We combined data from placebo and treatment groups, since in neither trial was a significant difference between the Tirilazad and the placebo treated group shown for the primary outcome measure, i.e., mortality and unfavorable outcome on the GOS.

Definitions of CT Characteristics and Outcome

We based our studies on data recorded for admission CT scans performed within the first 4 hours after injury. Full data on CT characteristics on admission were available in 2249 patients. CT data were extracted on the following items:

TABLE 1. Marshall computed tomographic classification[a]

Category	Definition
Diffuse injury I (no visible pathology)	No visible intracranial pathology seen on CT scan
Diffuse injury II	Cisterns are present with midline shift of 0–5 mm and/or lesions densities present; no high or mixed density lesion >25 cm^3 may include bone fragments and foreign bodies
Diffuse injury III (swelling)	Cisterns compressed or absent with midline shift of 0–5mm; no high or mixed density lesion >25 mm
Diffuse injury IV (shift)	Midline shift >5 mm; no high or mixed density lesion >25 cm^3
Evacuated mass lesion	Any lesion surgically evacuated
Non-evacuated mass lesion	High or mixed density lesion >25 cm^3; not surgically evacuated

[a]CT, computed tomographic.

- CT classification
- presence of abnormalities
- presence and size of midline shift
- status of basal cisterns
- presence of intraventricular blood (IVH)
- presence of traumatic subarachnoid hemorrhage (tSAH)
- presence and type of mass lesions and expected evacuation of mass lesion
- single versus multiple non-mass lesions

The characteristics "any abnormalities," "intraventricular blood," "tSAH," and expected evacuation of mass lesions were available in the data sets as binary data and scored as present or absent, without further differentiation. The other CT characteristics were classified into several categories with increasing differentiation: midline shift was classified in two ways: 1) shift $\leq$ 5 mm versus shift $>$ 5 mm and 2) no shift, shift of 1 to 5 mm, shift of 6 to 10 mm, or shift $>$ 10 mm. The status of basal cisterns was categorized in two ways: 1) normal versus abnormal (compressed and absent), and 2) normal, compressed, absent. Presence and type of mass lesions were categorized in 3 ways: 1) mass lesion present versus absent; 2) absent mass lesion, epidural mass lesion, intradural (intracerebral plus subdural) mass lesion; and 3) absent mass lesion, epidural mass lesion (EDH), subdural mass lesion (SDH), intracerebral mass lesion.

For several patients, the values of some of the predictors were missing (4.8% of the required values). These values were statistically estimated with regression models including the other predictors and subsequently imputed (14, 20). This approach is considered preferable to complete case analysis, in which patients with missing values are excluded from analysis (14). The outcome measure was mortality at 6 months postinjury.

Statistical Analysis and Performance of Models

To test whether the arrangements of the CT characteristics within the Marshall CT classification was reasonable for predictive purposes, we developed alternative models with the same variables. Subsequently, we investigated whether performance could be improved by adding CT characteristics or by separating already included characteristics into smaller categories. Each model was developed with two methods: recursive partitioning (CART) and logistic regression analysis.

We chose these two approaches for different reasons: logistic regression analysis is a standard statistical procedure, in which the relative importance of predictors is considered. Results are generally robust. On the other hand, recursive partitioning (2) has a greater clinical appeal because prediction trees are created, which are visually attractive. This approach, however, carries some risk of overfitting, especially in more complex trees. To correct for this, the trees were pruned using cross-validation. Modeling was performed with SAS software (version 6.12; SAS Institute, Inc., Cary, NC) and S-plus (version 2000; Insightful Corp., Seattle, WA), using the RPART library (http://www.stats.ox.ac.uk/pub/SWin).

Internal validity of the original CT classification and alternative models was assessed with bootstrapping procedures. Internal validity assesses whether the models perform well for a population of patients similar to those for whom the model was developed. Bootstrapping involved taking samples 100 times with replacement from the development sample. Each sample can be considered as repeating the data collection with the same number of patients and under identical circumstances as the original. In each of the 100 bootstrap samples a regression model was estimated, and evaluated on the original sample to estimate statistical optimism (7, 13, 14, 15). This validation approach leads to better predictions of outcome for patients similar to the development population (13, 14, 30).

Performance of the models was assessed with respect to discrimination, which can be quantified by the area under the receiver operating curve (AUC). For a randomly chosen pair of patients, the AUC represents the probability that a patient who dies has a higher predictive probability for mortality. The higher the AUC, the better the model discriminates. A model with an AUC of 0.50 has no discriminative power, while a model with an AUC of 1.0 reflects perfect discrimination.

Application in Clinical Practice

Presentation of a classification according to a prediction tree is readily understandable for a clinical audience. Interpretation of a logistic regression model is more complicated. To facilitate application of these models we created a score chart to estimate the outcome probability based on the values of the regression coefficient, which were re-scaled and rounded to whole numbers.

RESULTS

Individual CT Characteristics and Outcome

The distribution of CT characteristics and outcome is presented in *Table 2*. Mortality was lower in North American patients (19%) than in international patients (24%), partly reflecting a different distribution of severe versus moderate patients included in both trials. More than 90% of the patients had abnormalities on the admission CT scan: 84% showed evidence of parenchymal or extracerebral lesions, 45% had abnormal basal cisterns, 53% tSAH and 21% had intraventricular blood. Mass lesions were present in 39% of the population, and of these 80% had an associated intracerebral lesion. 74% of all mass lesions were evacuated and of these 84% were evacuated within 4 hours of injury.

Midline shift, basal cisterns, intraventricular blood, and traumatic SAH were identified as significant predictors of mortality (*Table 3*). In the multivariable analysis, the full differentiation of lesions was not clearly associated with differences in mortality, but the differentiation between epidural and intradural lesions was highly relevant (*Table 3* and *4*).

TABLE 2. Distribution of computed tomographic parameters and mortality in the international and North American samples[a]

CT parameters	Total (n = 2249) No. (%)	International sample (n = 1112) No. (%)	North American sample (n = 1137) No. (%)
Abnormalities			
No	173 (8)	52 (5)	121 (11)
Yes	2076 (92)	1060 (95)	1016 (89)
Missing	0	0	0
Traumatic SAH			
No	1030 (47)	519 (48)	511 (46)
Yes	1171 (53)	573 (52)	604 (54)
Missing	42	20	22
Intraventricular blood			
No	1746 (79)	863 (79)	883 (79)
Yes	473 (21)	235 (21)	238 (21)
Missing	30	14	16
Basal cisterns			
Normal	1194 (54)	576 (52)	618 (56)
Compressed	709 (32)	375 (34)	334 (30)
Absent	296 (13)	143 (13)	153 (14)
Missing	50	18	32
Midline shift			
None	1426 (64)	654 (59)	772 (70)
1–5 mm	362 (16)	208 (19)	154 (14)
6–10 mm	233 (11)	121 (11)	112 (10)
> 10 mm	190 (9)	118 (11)	72 (6.5)
Missing	38	11	27
Lesion			
No	360 (16)	117 (11)	243 (22)
Yes	1872 (84)	995 (89)	877 (78)
One, not mass	691 (37)	429 (29)	262 (30)
Multiple, not mass	585 (31)	329 (33)	256 (29)
Mass lesion	729 (39)	370 (37)	359 (41)
Epidural[b]	204 (28)	129 (35)	75 (21)
Subdural[b]	418 (57)	205 (55)	213 (59)
Intracerebral[b]	584 (80)	277 (75)	307 (86)
Missing	21	4	17
CT classification			
I	173 (8)	52 (5)	121 (11)
II	833 (37)	425 (38)	408 (36)
III	426 (19)	219 (20)	207 (18)
IV	88 (4)	46 (4)	42 (4)
V	539 (24)	289 (26)	250 (22)
VI	190 (8)	81 (7)	109 (10)
Missing	0	0	0
Outcome			
Mortality			
Yes	491 (22)	270 (24)	221 (19)
No	1758 (78)	842 (76)	916 (81)
Missing	0	0	0

[a] CT, computed tomographic.
[b] More than one type of mass lesion on the CT scan was possible.

TABLE 3. Multivariable analysis of computed tomographic characteristics, pooled Tirilazad patients[a]

Characteristics	Mortality (%)	OR (95% CI)[b]
Abnormalities	6.4	Reference
No	23	1.0 (0.5–2.0)
Yes		
Shift		
No shift	17	Reference
0–5 mm	26	1.4 (1.0 – 1.9)
6–10 mm	36	1.6 (1.1–2.4)
> 10 mm	49	2.0 (1.3–3.1)
Basal cisterns		
Normal	15	Reference
Compressed	27	2.0 (1.5–2.7)
Absent	55	5.7 (4.0–8.0)
Intraventricular blood		
No	19	Reference
Yes	31	2.0 (1.5–2.6)
tSAH		
No	12	Reference
Yes	30	2.0 (1.5–2.5)
Lesions		
No	12	Reference
Single non-mass	15	0.9 (0.6–1.4)
Multiple non-mass	23	1.3 (0.9–1.9)
Epidural mass	17	0.5 (0.4–0.9)
Subdural mass	40	1.4 (0.9–2.1)
Intracerebral mass	35	1.1 (0.7–1.7)

[a] OR, odds ratio; CI, confidence interval; tSAH, traumatic subarachnoid hemorrhage.
[b] Missing values were entered.

Prognostic value of CT classification versus alternative groupings of individual components

Figure 2 presents the classification of our patient population according to the Marshall CT classification in a prediction tree format with mortality figures per class. The percentage mortality in patients with no abnormalities (CT Class 1) and in patients with diffuse injuries without radiological signs of raised ICP (CT Class 2) was low (6.4 and 11% resp.).

The highest mortality rate, 44%, was observed in patients with absent or compressed basal cisterns and a midline shift larger than 5 mm (CT Class 4). Mortality rates for patients with mass lesions were 30% for those with evacuated mass lesions and 34% for those with non evacuated mass lesions. Analysis of the discriminatory properties of the Marshall CT classification showed an AUC of 0.669.

Figure 3 presents a prediction tree constructed with recursive partitioning, using the same characteristics and the same number of terminal nodes as used in the Marshall CT classification. We found that a primary division according to status of the basal cisterns yields the strongest discrimination. Subsequently for patients with present basal cisterns a split on abnormalities and for patients with absent or compressed basal cisterns a split on shift greater than 5 mm caused the maximum reduction in heterogeneity. Discriminatory analysis showed an AUC 0.705, considerably higher than found for the original CT classification.

Alternative models with additional variables

We investigated whether models could be developed with better discriminative properties by adding additional CT predictors not originally included in the Marshall CT classification or by further separation of already included CT characteristics.

The added benefit of additional parameters was assessed versus the basic model presented in *Figure 3*. Results are summarized in *Table 4*. The discriminative ability could be improved considerably by adding tSAH and intraventricular blood, and by further differentiating the basal cisterns, midline shift and mass lesions into several categories. Dropping the characteristic "any abnormalities" had negligible influence on the AUC. No statistically significant interactions were observed between the selected characteristics ($P = 0.42$).

The results of this multivariable analysis showed the potential for developing an alternative model with added characteristics. Including all discriminating variables in a logistic regression model yielded an AUC of 0.769, and of 0.794 in the prediction tree. Such a model is, however, complex and resulted in 15 terminal nodes in the prediction tree. Searching for an appropriate compromise between good discrimination and easy clinical applicability we chose a more simple model based on the following characteristics: midline shift (subdivided into 0–5 mm, > 5 mm), basal cisterns (subdivided into absent, compressed and present), mass lesion (subdivided into epidural and intradural), traumatic subarachnoid hemorrhage and/or intraventricular blood. The apparent validity of this model was 0.750 and on internal validation we obtained an AUC of 0.748.

Clinical application

For clinical application, we translated the logistic regression model into a score chart, with which the probability for mortality according to the CT characteristics can be estimated by adding the scores for individual patients (*Table 5*). We chose to add plus 1 to the sum score to make the grading numerically consistent with the grading of the motor score of the GCS and with the Marshall CT classification. *Table 6* shows the application of this score chart for classifying the study population according to prognostic risk. The difference in observed mortality rates between patients from the lowest and patients from the highest risk group is 61%, which is considerably larger than the maximal difference in mortality in the Marshall CT classification (38%, *Fig. 2*), *Table 7* illustrates the better discrimination for prognostic risk assessment of the CT prediction score in comparison to the Marshall CT classification, particularly in patients with mass lesions.

TABLE 4. Added discriminative value of extra computed tomographic characteristics and further differentiation of computed tomographic characteristics[a]

Models	Discrimination (AUC)	
	Logistic regression	Recursive partitioning
Basic model[b]	0.703	0.705
Added CT characteristic		
Evacuation mass lesion	0.714	0.712
Dropped CT characteristics		
Any abnormalities	0.703	0.701
Added CT characteristics		
Non-mass lesions (single or multiple)	0.714	0.712
Blood (tSAH and/or intraventricular blood)	0.730	0.737
Separation of already included CT characteristics		
Mass as epidural versus intradural	0.720	0.719
Mass as epidural versus subdural versus intracerebral	0.722	0.721
Cisterns (normal, compressed, absent)	0.726	0.727
Shift (no, 1–5 mm, 6–10 mm, > 10 mm)	0.710	0.716
All[c]	0.769	0.794

[a] AUC, area under the operating curve; CT, computed tomographic; tSAH, traumatic subarachnoid hemorrhage.
[b] The basic model contains all characteristics included in the Marshall classification, except evacuation mass lesion.
[c] Basic model plus single non-mass lesions, multiple non-mass lesions, mass as epidural versus subdural versus intracerebral, cisterns as normal versus compressed versus absent, shift as 1–5 mm versus 6–10 mm versus > 10 mm, blood as tSAH versus intraventricular blood.

DISCUSSION

We confirmed the predictive value of the Marshall CT classification in a large series of patients (n = 2249), but showed that a better discrimination can be obtained by making fuller use of the individual CT characteristics underlying the Marshall CT classification. Discrimination could be further improved by adding intraventricular and traumatic subarachnoid hemorrhage and by a more detailed differentiation of mass lesions and basal cisterns (AUC, 0.77). We do not wish to detract from the general validity and appeal of the Marshall CT classification when used for descriptive purposes. This classification however was not developed from the perspective of prognosis, and the question whether the categorization of variables in the Marshall CT classification is appropriate for predictive purpose is relevant. For instance, in the Marshall CT classification, radiological signs of raised ICP (status of basal cisterns and presence of shift) are only used for further differentiation of patients with diffuse injuries while these parameters may also be expected to be of prognostic value particularly in patients with mass lesions. Indeed, *Table 7* shows a better prognostic discrimination of the proposed prognostic CT score over the Marshall CT classification, particularly in patients with mass lesions.

The presence of tSAH has been shown to be a strong predictor both for outcome and mortality in TBI (8, 11, 12, 18, 26, 27, 29) but is not included in the Marshall CT classification. The predictive value of tSAH in TBI is confirmed in our study and we have additionally shown that including this parameter in a predictive model significantly increases discrimination; we also found IVH to be an independent predictor, in contrast to other studies in which the relation of IVH to poorer outcome was mainly caused by the association with other predictors (6, 8, 19). Further, the Marshall CT classification does not permit any distinction on type of mass lesion. Many studies have shown that prognosis in patients with an EDH is much better than in those with a sub-

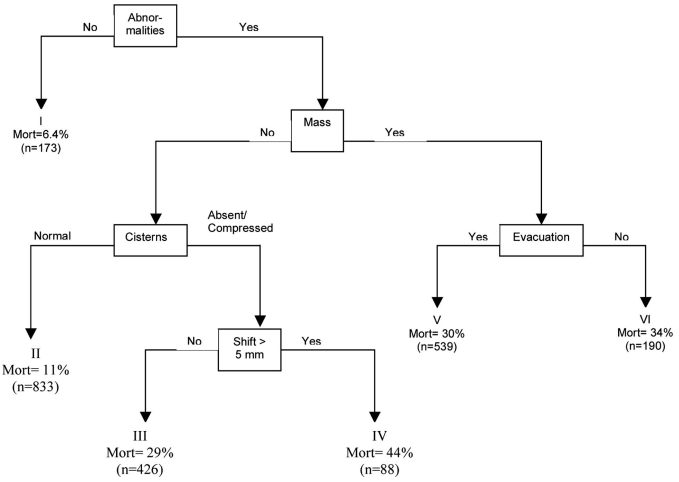

FIGURE 2. *Mortality related to Marshall CT classification.*

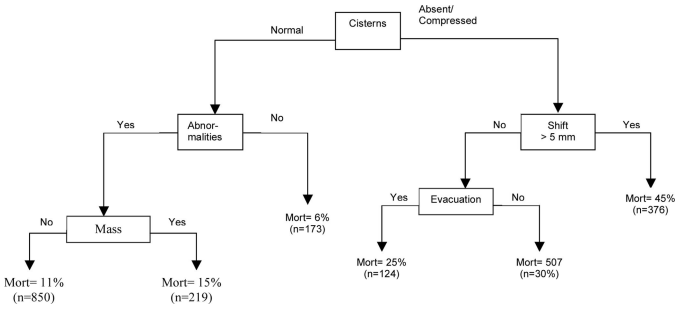

FIGURE 3. *CT prediction tree constructed with recursive partitioning, using the same characteristics and the same number of terminal nodes as used in the Marshall CT classification.*

TABLE 5. Prognostic score chart for the probability of mortality in patients with severe or moderate traumatic brain injury according to their computed tomographic characteristics[a]

Predictor value	Score
Basal cisterns	
Normal	0
Compressed	1
Absent	2
Midline shift	
No shift or shift ≤ 5 mm	0
Shift > 5 mm	1
Epidural mass lesion	
Present	0
Absent	1
Intraventricular blood or tSAH	
Absent	0
Present	1
Sum score[b]	+1

[a] SAH, traumatic subarachnoid hemorrhage.
[b] The sum score can be used to obtain the predicted probability of mortality from the formulae below. We chose to add plus 1 to make the grading numerically consistent with the grading of the motor score of the GCS and with the Marshall CT classification. The corresponding probabilities are calculated with the formula: Probability (mortality) = $1/[1+ e^{-(-2.60+ 0.80 *Sumscore)}]$

TABLE 6. Computed tomographic classification by prediction score

Score	No. of patients	Actual mortality no. (%)
1	36	0 (0)
2	600	41 (6.8)
3	773	122 (16)
4	465	121 (26)
5	261	138 (53)
6	114	69 (61)

dural or intracerebral hematoma (5, 10). Bricolo et al. (3) postulated that mortality should approach zero in patients with an uncomplicated EDH. As shown in *Table 6*, we found zero mortality in such patients (mass lesion with a prognostic score of 1). A further problem with the original CT classification is that it differentiates between patients with evacuated versus non evacuated mass lesions. Many have argued that this reflects a clinical decision and does not in itself constitute a CT parameter, and in clinical practice this has led to confusion and it has been proposed not to include this differentiation (28). Nevertheless, we do note a 4% difference in mortality between patients with evacuated versus non evacuated mass lesions. Further in depth adjusted analysis, however, will be required to determine whether the baseline characteristics of these two groups were similar or not.

TABLE 7. Marshall computed tomographic classification versus Rotterdam computed tomographic score[a]

Rotterdam CT Score	Marshall CT classification						Total
	1	2	3	4	5	6	
1	0	0	0	0	35	1	36
2	173	336	0	0	65	26	600
3	0	492	107	5	95	74	773
4	0	5	249	19	136	56	465
5	0	0	70	37	134	20	261
6	0	0	0	27	74	13	114
Total	173	833	426	88	539	190	2249

[a] CT, computed tomographic.

Previous studies have shown that the Marshall CT classification is a strong predictor in TBI (5, 17, 21, 28) with high inter- and intraobserver reliability (33). Wardlaw et al. (34), however, found in a retrospective analysis of 425 patients of varying severity that the Marshall CT classification did not remain a significant independent outcome predictor on multivariate analysis when clinical features were included, in contrast to tSAH and a newly suggested, ill defined variable describing "overall appearance." For the present study we did not include clinical characteristics in our models, but in a previous study describing a prediction model for TBI, we found that both the Marshall CT classification and tSAH remained as statistically significant predictors in multivariate analysis, following adjustment for clinical variables (17).

Consistent with other prediction studies in TBI (31), we found that performance of the models was more dependent on the variables included than on the statistical approach. We found no clear statistical benefit in the use of a prediction tree compared to logistic regression models. We considered the use of a prediction tree in the current analysis appropriate as the Marshall CT classification can be readily presented as a prediction tree, which may be appealing for the clinician. Furthermore, a tree can capture and correct easily for interaction, i.e., different relations between predictors in different subgroups. Interaction, if present, is detected by a better discriminative ability of the tree as compared to a logistic model. We did not observe such differences in our studies. The clinical appeal of a prediction tree method is, however, also dependent on the number of terminal nodes. The limited number of nodes (n = 6) in the Marshall CT classification and in the basic model make this type of presentation appropriate. When additional variables were added to our model we found an optimal number of 15 terminal nodes which significantly decreases the clinical appeal. For this reason we would prefer the logistic regression model because discriminative properties are similar. We realize that a logistic regression model may have less clinical appeal and therefore suggest translating it into a score chart as proposed in *Table 6*. Although this score chart performed well in our study population, assessment of

its general applicability will require validation in other data sets.

A number of limitations of our study should be recognized. First, our studies were performed on a large patient series including only patients with severe and moderate injury. Results cannot, therefore, be extrapolated towards patients with mild injuries. Secondly, we focused our studies on analysis of data from the initial CT examination performed within 4 hours after injury. Other studies (21, 28) have shown that the "worst" CT scan obtained during the clinical course has greater predictive value. Also within the current data set we found that the final CT classification, based on the worst CT following admission, yielded better discrimination (AUC, 0.692 for Marshall CT classification and 0.716 for basic model). However, the intent of our studies was to investigate the use of the CT classification and CT predictors toward a prognostic classification of TBI on admission. Such classification is considered useful to establish the baseline characteristics and prognostic risk of TBI patients on admission. Third, the predictive analysis presented was conducted versus 6 months mortality. For these studies we chose mortality rather than the GOS dichotomized into unfavorable versus favorable as this constitutes a hard and objective endpoint without any missing outcome data. As a sensitivity analysis we additionally calculated the discriminative properties of the Marshall CT classification, the basic model in which the individual parameters of the CT classification were rearranged and the extended model versus unfavorable outcome and found similar results.

In summary, we conclude that the Marshall CT classification has strong predictive power, but greater discrimination can be obtained if the individual CT parameters underlying the CT classification are included in a prognostic model. Consequently, for prognostic purposes, we recommend the use of individual characteristics rather than the CT classification. Performance of CT models for predicting outcome in TBI can be significantly improved by including more details of variables and by adding other variables to the model. We suggest that such models should include the following characteristics: status of basal cisterns, shift, tSAH and/or IVH and presence of mass lesions with differentiation between EDH versus intradural lesions. For more easy clinical application, models can be translated into a score chart.

REFERENCES

1. Balestreri M, Czosnyka M, Chatfield DA, Steiner LA, Schmidt EA, Smielewski P, Matta B, Pickard JD: Predictive value of Glasgow coma scale after brain trauma: Change in trend over the past ten years. **J Neurol Neurosurg Psychiatry** 75:161–162, 2004.
2. Breiman L, Friedman JH, Olshen RA: *Classification and Regression Trees.* Belmont, Wadsworth, 1984.
3. Bricolo AP, Pasut LM: Extradural hematoma: toward zero mortality: A prospective study. **Neurosurgery** 14:8–11, 1984.
4. Buechler CM, Blostein PA, Koestner ARN, Hurt K, Schaarsr M, McKernan J: Variation among Trauma Centers Calculation of Glasgow Coma Scale Score: Result of a national survey. **J Trauma** 45:429–432, 1998.
5. Chesnut RM, Ghajar J, Maas AR: Guidelines for the Management and prognosis of severe traumatic brain injury part II: Early indicators of prognosis in severe traumatic brain injury. **J Neurotrauma** 17:556–627, 2000.
6. Cordobes F, de la Fuente M: Intraventricular hemorrhage in severe head injury. **J Neurosurg** 58: 217-222, 1983.
7. Efron B,Tibshirani R: *An introduction to the bootstrap.* New York, Chapman and Hall, 1993.
8. Eisenberg HM, Gary HE Jr, Aldrich EF, Saydjari C, Turner B, Foukles MA, Jane JA, Marmarou A, Marshall LF, Young HF: Initial CT findings in 753 patients with severe head injury. A report from the NIH traumatic coma data bank. **J Neurosurg** 73:688–698, 1990.
9. Firsching R, Woischneck D, Klein S, Reissberg S, Dohring W, Peters B: Classification of severe head injury base on magnetic resonance imaging. **Acta Neurochirurgica (Wien)** 143:263–271, 2001.
10. Gennarelli TA, Spielman GM, Langfitt TW, Gildenberg PL, Harrington T, Jane JA, Marshall LF, Miller JD, Pitts LH: Influence of the type of intracranial lesion on outcome from severe head injury. **J Neurosurg** 56:26–32, 1982.
11. Greene KA, Jacobowitz R, Marciano FF: Impact of traumatic subarachnoid hemorrhage on outcome in nonpenetrating head injury. Part II: Relationship to clinical course and outcome variables during acute hospitalization. **J Neurotrauma** 41:964-971, 1996.
12. Greene KA, Marciano Ff, Johnson BA, Jacobowitz R, Spetzler RF, Harrington TR: Impact of traumatic subarachnoid hemorrhage on outcome in nonpenetrating head injury. Part I: A proposed computerized tomography grading scale. **J Neurosurg** 83:445–452, 1995.
13. Harrell FE: *Regression modeling strategies: With applications to linear models, logistic regression and survival analysis.* New York, Springer-Verlag, Inc., 2001.
14. Harrell FE, Lee KL, Mark DB: Multivariable prognostic models: Issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. **Stat Med** 15:361-387, 1996.
15. van Houwelingen JC, le Cessie S: Predictive value of statistical models. **Stat Med** 8:1303–25, 1990.
16. Hukkelhoven CW, Steyerberg EW, Farace E, Habbema JD, Marshall LF, Maas AI: Regional differences in patient characteristics, case management, and outcomes in traumatic brain injury: Experience from the Tirilazad trials. **J Neurosurg** 97:549–557, 2002.
17. Hukkelhoven CW, Steyerberg EW, Habbema JD: Predicting outcome after traumatic brain injury: development and validation of a prognostic score base on admission characteristics. **J Neurotrauma** (in press).
18. Kakarieka A, Braakman R, Schakel EH: Clinical significance of the finding of subarachnoid blood on CT scan after head injury. **Acta Neurochir (Wien)** 129:1-5, 1994.
19. Lee JP, Lui TN, Chang CN: Acute post-traumatic intraventricular hemorrhage analysis of 25 patients with emphasis on final outcome. **Acta Neurol Scand** 84:85–90, 1991.
20. Little R: Regression with missing X's: A review. **J Am Stat Assoc** 87:1227–1237, 1992.
21. Lobato RD, Cordobes F, Rivas JJ, de la Fuente M, Montero A, Barcena A, Perez C, Cabrera A, Lamas E: Outcome from severe head injury related to the type of intracranial lesion: A computerized tomography study. **J Neurosurg** 59:762–774, 1983.
22. Lobato RD, Gomez PA, Alday R, Rivas JJ, Dominguez J, Cabrera A, Turanzas FS, Benitez A, Rivero B: Sequential Computerized Tomography Changes and Related Final Outcome in Severe Head Injury Patients. **Acta Neurochir (Wien)** 139:385–391, 1997.
23. Marshall LF, Maas AR, Bowers MS: A multicenter trial on the efficacy of using Tirilazad Mesylate in cases of head injury. **J Neurosurg** 89:519–525, 1998.
24. Marshall LF, Marshall SB, Klauber MR, Clark MvB, Eisenberg HM, Jane JA, Luerssen TG, Marmarou A, Foulkes MA: A new classification of head injury based on computerized tomography. **J Neurosurg** 75:S14–S20, 1991.
25. Moskopp D, Stähle C, Wassmann H: Problems of the Glasgow Coma Scale with early intubated patients. **Neurosurg Rev** 18:253-257, 1995.
26. Ono J, Yamaura A, Kubota M, Okimura Y, Isobe K: Outcome prediction in severe head injury: Analyses of clinical prognostic factors. **J Clin Neurosci** 8:120–123, 2001.
27. Selladurai BM, Jayakumar R, Tan YY, Low HC: Outcome prediction in early management of severe head injury: An experience in Malaysia. **Br J Neurosurg** 6:549–557, 1992.

28. Servadei F, Murray GD, Penny K: The value of the "worst" computed tomographic scan in clinical studies of moderate and severe head injury. **Neurosurgery** 46: 70-77, 2000.
29. Servadei F, Murray GD, Teasdale GM, Dearden M, Iannotti F, Lapierre F, Maas AJ, Karimi A, Ohman J, Persson L, Stocchetti N, Trojanowski T Unterberg A: Traumatic Subarachnoid Hemorrhage: Demographic and clinical study of 750 patients from the European Brain Injury Consortium Survey of Head Injuries. **Neurosurgery:** 50:261–269, 2002.
30. Steyerberg EW, Eiikemans MJ, Harrell FE Jr, Habbema JD: Prognostic modeling with logistic regression analysis: A comparison of selection and estimation methods in small data sets. **Stat Med** 19:1059–1079, 2000.
31. Titterington DM, Murray GD, Murray LS, Spiegelhalter DJ, Skene AM, Habbema JDF, Gelpke GJ: Comparison of discrimination techniques applied to a complex data set of head injured patients. **J R Stat Soc [Ser A]** 144:145–175, 1981.
32. Uchino Y, Okimura Y, Tanaka M, Saeki N, Yamaura A: Computed Tomography and Magnetic Resonance Imaging of Mild Head Injury—is it appropriate to classify patients with Glasgow Coma Scale Score of 13 to 15 as mild injury? **Acta Neurochir (Wien)** 143:1031–1037, 2001.
33. Vos PE, van Voskuilen AC, Beems T, Krabbe PF, Vogels OJ: Evaluation of the traumatic coma data bank computed tomography classification for severe head injury. **J Neurotrauma** 18:649–655, 2001.
34. Wardlaw JM, Easton VJ, Statham P: Which CT features help predict outcome after head injury? **J Neurol Neurosurg Psychiatry** 72:188–192, 2002.

Acknowledgments

We express our gratitude to all of the study participants and the principal investigators of the Tirilazad trials whose work made this report possible. We further wish to thank Marja van Gemerden for her administrative assistance. Grant support was provided by National Institutes of Health Grant no. NIH-NS 42691.

COMMENTS

Maas et al. have conducted a very thorough analysis of the correlation of specific computed tomographic (CT) characteristics with 6-month mortality in a population of patients with severe and moderate traumatic brain injury (TBI). As one might expect, they found that making greater use of individual CT characteristics allowed them to improve on the already sizeable predictive value of the original Marshall CT classification scheme, which assessed CT scans by placing them into specific categories. Unlike the authors of many such predictive schemes, Maas et al. went to great lengths to create a simple scoring chart to facilitate the use of their system. The discussion is especially frank in its acknowledgment of the limitations of this work. However, a word of caution is in order concerning the application of these results. Information such as that provided by this model is quite useful in helping to form a general idea about a patient's prognosis, but it is only one item of data in the overall assessment of a patient's clinical condition, underlying health status, rate of progression of disease, and numerous other factors that must be considered when a patient's overall "prognosis" is assessed.

Alex B. Valadka
Houston, Texas

In this important publication, the authors have improved upon the Marshall classification of CT scan in TBI. They have rearranged the underlying individual CT scan characteristics and included new characteristics, namely traumatic subarachnoid and intraventricular hemorrhage. They have subsequently applied this new "Marshall/Maas Classification" to a large population of head injured patients taken from the previous Tirilazad trial (2269 patients)(1, 2). Using statistical modeling techniques, the authors have shown that this new classification is a powerful predictor of outcome in TBI.

Thus, in contrast with many other disease states, severe brain injury outcomes can be accurately predicted using four variables. These variables are age, Glasgow Coma Scale on admission (especially motor score), CT characteristics, and presence of ischemic and hemodynamic secondary insults. Using these four predictors, numerous papers have used statistical modeling techniques to predict up to 80% or more of the outcome accuracy. It remains to be seen whether this new modification of the CT characteristics will boost this predictive accuracy even higher. Clearly, the algorithms presented in this important article need to be applied both prospectively and retrospectively to other large patient cohorts to find this out.

The most important implication of this study, however, is that it is possible to test outcome in head injury against the individual predictors for each individual patient. In the classical design of head injury trials, randomization and adequate power have been used as the two tools to allow for a treatment effect to be detected between placebo and treatment groups.

If each patient's outcome could be measured against their predicted outcome, however, this suggests that much smaller numbers of patients could be used in trials, and also that it would not be necessary to stratify for severity as has been done in most previous trials. This, in turn, opens up the possibility that head injury trials could be performed more rapidly and more cost-effectively. Only by the development of surrogate markers such as the CD-4 count, and the viral load has it been possible for the acquired immunodeficiency syndrome/human immunodeficiency virus research community to develop the wide spectrum of therapeutic agents available for that disease. This article constitutes significant progress along the road of better trial design in head injury for the future.

M Ross Bullock
Richmond, Virginia

1. Hukkelhoven CWP, Steyerberg EW, Farace E, Habbema JD, Marshall LF, Maas AI: Regional differences in patient characteristics, case management, and outcomes in traumatic brain injury: experience from the Tirilazad trials. **J Neurosurg** 97:549–557, 2002.
2. Marshall LF, Maas AI, Marshall SB, Bricolo A, Fearnside M, Iannotti F, Klauber MP, Langarrigue J, Lobato R, Persson L, Pickard JD, Plek J, Servadei F, Wellis GN, Morris GF, Means ED, Musch B: A multicenter trial on the efficacy of using tirilazad mesylate in cases of head injury. **J Neurosurg** 89:519–525, 1998.

The ability to predict outcome early after a severe TBI enables treating physicians to appropriately advise family members regarding expected duration of treatment, likely short- and long-term outcomes, future needs for rehabilitation, and other clinical, social and financial issues that require extensive planning. Early outcome prediction also helps clinicians and hospitals anticipate and plan for the future medical needs of the patient. For more than a decade, the Marshall CT classification scheme has been an important adjunct to the initial clinical assessment in providing the most accurate early outcome prediction possible. But, those of us who have used this CT classification have had concerns that it did not account for CT findings that seemed to be prognostically important. Maas et al. have focused on those shortcomings using a sophisticated series of statistical analyses applied to a very large database. In so doing, they confirmed that distinguishing between epidural and subdural hematomas, and including traumatic subarachnoid hemorrhage and intraventricular hemorrhage as additional predictive variables, significantly improves the prognostic value of the initial CT scans as compared with the

Marshall Classification. They developed a scoring system that seems to be more closely associated with outcomes than the Marshall scheme when applied to the more than 2000 patients in the Tirilazad database. Although it remains impossible to predict outcomes with complete accuracy, the new classification system developed by these authors is a welcome improvement to the Marshall system and will significantly improve our prognostic abilities.

Donald W. Marion
Boston, Massachusetts

In this clinical study, the authors evaluate the predictive value of the Marshall CT classification scheme in patients with moderate-to-severe TBI with 6-month mortality as an outcome measure and compare it with that of several alternative CT models developed by rearrangement of various CT characteristics. The article highlights the significance of radiological examination in the assessment of patients with severe head injuries in whom the Glasgow Coma Scale score has become less applicable for TBI grading as a result of the increasingly frequent use of early intubation, sedation, and mechanical ventilation in the treatment of this patient population. It also points out the fact that, although the Marshall CT classification has been widely accepted as an accurate descriptor and predictor (in multiple studies) of prognosis in TBI, no one has investigated the possibility that other combinations of CT parameters may, in fact, be better suited for the purpose of prognostication, which is what the authors aim to accomplish in this study. Using a database of the combined Tirilazad trials, they show that the inclusion of various characteristics, such as the presence or absence of subarachnoid and intraventricular hemorrhage, the quality of the basal cisterns and the location of mass lesions, in a novel classification scheme results in a statistically significant increase in the discriminatory capabilities of an outcome predictor model. By translating this complex logistic regression model into a score chart similar to the motor score of the Glasgow Coma Scale, they attempt to create a more practical clinical classification scheme and demonstrate its superiority to the Marshall model in a "head to head" comparison.

We found several areas of concern in this article, a number of which the authors recognize themselves, including the following: The database excludes patients with mild TBI, a significant subset of individuals for whom this classification would not be applicable. The emphasis on the importance of basal cisterns and midline shift in the model requires some consistency in the accuracy with which these CT scans are read and interpreted. Is there any difference in the predictive value when films are evaluated by physicians with dissimilar levels of experience, i.e., neuroradiology attendings versus radiology residents versus neurosurgical residents? Also, the study uses the "initial CT examination performed within 4 hours after injury." As the authors mention, it has been shown elsewhere that the worst CT scan obtained during the clinical course has greater predictive value. It would be interesting to see if the results change when this criterion is used and if the 4-hour cutoff really has any merit in the overall TBI population. To make their model more clinically relevant, the authors convert the values of the regression coefficient to a simple six-point score estimating outcome probability (rescaled and rounded to whole numbers). In doing this, they likely sacrifice some accuracy and would have to answer the question of whether or not this score applies to TBI patients in general; it is absolutely necessary to evaluate the validity of this scale in a prospective fashion. We also agree with the decision to exclude clinical features in these models, as the aim here is to develop a classification scheme based solely on the features of the CT examination.

Jason H. Huang
Nathan Ranalli
Eric L. Zager
Philadelphia, Pennsylvania

The Marshall CT classification has served a very useful purpose as a predictor of head injury severity and outcome and has been widely embraced since its introduction in 1990. However, it is important to bear in mind that it was devised from data collected between 1984 and 1987 through the Traumatic Coma Data Bank, and is based on a total of 753 patients.

Nevertheless, until this contribution by Maas et al., there has not been a concerted effort to revisit or revise this classification to achieve greater reliability and sensitivity. This novel modeling methodology is based on more than 2000 scans and seems to have accomplished these goals.

Only two concerns are expressed over this important contribution. First, the scan abnormalities are heavily weighted to mass lesions (84%), with 74% of these undergoing evacuative surgery. This would appear somewhat skewed given the current spectrum of what is presently being seen in trauma centers in the US and may unintentionally bias the modeling. Second, the authors' findings need to be corroborated on another large head injury database or through prospective investigations

Jack E. Wilberger
Pittsburgh, Pennsylvania