Grading the Amount of Blood on Computed Tomograms After Subarachnoid Hemorrhage

A. Hijdra, MD, P.J.A.M. Brouwers, MD, M. Vermeulen, MD, and J. van Gijn, MD

According to several studies, the amount of subarachnoid blood on the initial computed tomogram of patients with aneurysmal subarachnoid hemorrhage has predictive value with respect to infarction and outcome. Of several methods for assessing the amount of subarachnoid blood, none has been subjected to a study of interobserver agreement. We describe our own method, applied in previous studies, in which the amounts of blood in 10 basal cisterns and fissures and in four ventricles are graded separately. In grading single computed tomograms of 182 consecutive patients with subarachnoid hemorrhage, the agreement between pairs of three observers, studied with $\kappa$ statistics, was relatively good for individual cisterns or fissures ($\kappa$ between 0.35 and 0.65) and ventricles ($\kappa$ between 0.47 and 0.74). The Spearman rank correlation coefficients for the sum of the scores for subarachnoid and intraventricular blood were very high. Summed scores for extravasated blood are suitable as a baseline variable in follow-up studies of patients with subarachnoid hemorrhage. (Stroke 1990;21:1156-1161)

The amount of aneurysmal subarachnoid blood on an early computed tomogram (CT scan) is associated with subsequent vasospasm and cerebral ischemia and with final outcome. This is important for the management of patients with aneurysmal subarachnoid hemorrhage and for the evaluation of new modes of treatment. Most investigators have adopted their own method of quantifying the amount of subarachnoid blood on CT scans (Table 1), but none of these studies was primarily concerned with evaluating the method itself. Some authors recorded only the presence or absence of blood in the subarachnoid space, but since >90% of patients have such blood on a CT scan made <1 day after onset, this method seems too blunt. Quantification of the amount of blood present, however, is difficult. It cannot be assumed that density at a single point in the subarachnoid space measured in Hounsfield units or on a semiquantitative scale accurately reflects the total amount of extravasated blood, and measurement of the thickness of layers of blood or of clots depends as much on the dimensions of the cisterns in question as on the amount of subarachnoid blood. Simple grading scales with a limited number of categories deal inadequately with all possible aneurysm sites. Some methods also take into account intracerebral and intraventricular blood, but these types of hemorrhage can be combined with only a small amount of subarachnoid blood. Other studies emphasize the extent of subarachnoid blood in the interhemispheric fissure and the sylvian fissures, which reflects the assumption that vasospasm is elicited by tight clots around major cerebral vessels. However, cerebral infarction after subarachnoid hemorrhage is often a multifocal or diffuse event. The initial impact of the hemorrhage could play an important part, and the total amount of blood could be as important for the development of ischemia as its presence in certain cisterns and fissures.

The consequence is that grading can be only semiquantitative and subjective, but subjectivity can be restricted by choosing a method that gives the best interobserver agreement. Some methods have been found difficult to apply outside the center in which they were developed. Based on these considerations, we propose that a method of grading the amount of subarachnoid blood on CT scans should fulfill the following criteria: 1) the total amount of subarachnoid blood must be graded, 2) the distribution and extension of blood among all basal cisterns and fissures must be reflected in the grading system, 3) the amount of subarachnoid blood must be graded independent of the amounts of intracerebral and intraventricular blood, and 4) such a grading method must be tested for interobserver agreement. We describe and evaluate such a method.
Table 1. Methods of Grading Amount of Subarachnoid Blood on Computed Tomograms

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Year published</th>
<th>Anatomic distribution</th>
<th>Measurement technique</th>
<th>Intraventricular blood graded?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Takemae et al1</td>
<td>1978</td>
<td>-</td>
<td>G</td>
<td>No</td>
</tr>
<tr>
<td>Bell et al13</td>
<td>1980</td>
<td>-</td>
<td>G</td>
<td>No</td>
</tr>
<tr>
<td>Davis et al14</td>
<td>1980</td>
<td>+</td>
<td>G</td>
<td>No</td>
</tr>
<tr>
<td>Fisher et al5</td>
<td>1980</td>
<td>-</td>
<td>M</td>
<td>Yes*</td>
</tr>
<tr>
<td>Suzuki et al7</td>
<td>1980</td>
<td>-</td>
<td>H</td>
<td>No</td>
</tr>
<tr>
<td>Sano et al8</td>
<td>1982</td>
<td>±</td>
<td>H</td>
<td>No</td>
</tr>
<tr>
<td>Taneda14</td>
<td>1982</td>
<td>-</td>
<td>G</td>
<td>No</td>
</tr>
<tr>
<td>Allen et al16</td>
<td>1983</td>
<td>-</td>
<td>G</td>
<td>Yes*</td>
</tr>
<tr>
<td>Gurusinghe and Richardson10</td>
<td>1984</td>
<td>-</td>
<td>M</td>
<td>No</td>
</tr>
<tr>
<td>Mohsen et al11</td>
<td>1984</td>
<td>-</td>
<td>G, M</td>
<td>No</td>
</tr>
<tr>
<td>Pasqualin et al12</td>
<td>1984</td>
<td>-</td>
<td>G</td>
<td>No</td>
</tr>
<tr>
<td>Fujita13</td>
<td>1985</td>
<td>-</td>
<td>H</td>
<td>No</td>
</tr>
<tr>
<td>Inagawa et al17</td>
<td>1987</td>
<td>±</td>
<td>G</td>
<td>Yes</td>
</tr>
<tr>
<td>Petruk et al18</td>
<td>1988</td>
<td>-</td>
<td>G</td>
<td>Yes</td>
</tr>
</tbody>
</table>

- anatomic distribution subordinate to some focal measurement in assignment of a certain grade; +, definite information on anatomic distribution; ±, some information on anatomic distribution; G, semiquantitative grading; M, anatomic measurement; H, Hounsfield units.

*Intraventricular blood is incompatible with more than small amounts of subarachnoid blood in this grading system.

Subjects and Methods

Each of 10 basal cisterns and fissures was graded separately on a semiquantitative scale, according to the amount of extravasated blood: 0, no blood; 1, small amount of blood; 2, moderately filled with blood; or 3, completely filled with blood. The density of the clot was not considered. Clots that had expanded the original size of a cistern or fissure were still graded as 3. The total amount of subarachnoid blood (sum score) was calculated by adding the 10 scores and ranged from 0 to 30. When an occasional cistern or fissure was considered to be inadequately visualized, we interpolated by assigning to that cistern or fissure the average score of the others. Figure 1 shows a CT scan and a corresponding diagram as an example of the grading system.

The grading scale for the amount of blood in the four ventricles was constructed in a comparable fashion, as follows: 0, no blood; 1, sedimentation of blood in the posterior part; 2, partly filled with blood; or 3, completely filled with blood. The total amount of intraventricular blood (sum score) was the total of the four scores and ranged from 0 to 12.

Before the study of interobserver agreement, in a pilot study three of the authors independently graded 10 randomly chosen CT scans and compared their results. Discrepancies between scores as well as problems in the identification of certain cisterns and

![Computed tomogram of patient 6 hours after subarachnoid hemorrhage from carotid artery aneurysm on left (L) side. *Top diagram identifies 10 basal cisterns and fissures: A, frontal interhemispheric fissure; B, sylvian fissure, lateral parts; C, sylvian fissure, basal parts; D, suprasellar cistern; E, ambient cisterns; F, quadrigeminal cistern. **Bottom diagram indicates amount of blood in each cistern and fissure. Sum score is 22 points.](http://stroke.ahajournals.org/)

Downloaded from http://stroke.ahajournals.org/ by guest on August 26, 2017
in the differentiation of subarachnoid and intracerebral blood were discussed. The cisterns, fissures, and ventricles, together with the relevant problems identified in this pilot study, follow.

In the frontal interhemispheric fissure (A in Figure 1) the falx can be mistaken for blood. When later CT scans are available, densities of the structure in question can be compared. If the densities do not differ over time, the structure is the falx. Distinguishing subarachnoid from intracerebral blood may be difficult; a hematoma can often be identified because it deviates from the midline at some point, and sometimes cerebrospinal fluid can be identified along the edge of the hematoma. The lateral parts of the sylvian fissure (B in Figure 1), an insular cistern, are mostly in the sagittal plane and perpendicular to the basal parts of the sylvian fissure (C in Figure 1), which are in the frontal plane. The basal parts of the sylvian fissure lie anterosuperior to the tip of the temporal lobe and run from the suprasellar cistern (D in Figure 1) laterally to the sphenoid angle. The interpeduncular fossa is considered part of the suprasellar cistern. Left and right parts of the suprasellar cistern are graded separately since blood is often distributed asymmetrically. The shape of the suprasellar cistern depends on the angle of the CT slice. Blood in the ambient cisterns (E in Figure 1) can be confused with blood on the tentorium. In the latter case, cerebrospinal fluid is visible between the mid-brain and the tentorial edge, and the line of blood does not curve inward into the quadrigeminal cistern (F in Figure 1) but continues in a posterolateral direction. Asymmetrical distribution of blood in the quadrigeminal cistern always depends on uneven distribution between the two ambient cisterns, and therefore the left and right parts are not graded separately. The lateral ventricles often contain a small amount of sedimented blood in the posterior horns; some CT scans show a narrow layer of blood near the foramina of Monro. One point was assigned to these findings.

We studied interobserver agreement in the grading of admission CT scans of 182 consecutive patients with subarachnoid hemorrhage who were investigated ≤72 hours after the initial hemorrhage. An aneurysmal origin of the hemorrhage was proven by angiography or autopsy in 130 patients and was strongly suggested by CT scan evidence in 44; in the remaining eight patients the cause of the hemorrhage was nonaneurysmal (perimesencephalic hemorrhage or arteriovenous malformation). All CT scans were included, and all showed at least some blood in the subarachnoid space or the ventricles. Routine standard series without contrast enhancement were obtained. Two different machines were used, an EMI CT 1010 (Hayes, Middlesex, England) and a Philips Tomoscan 310 (Eindhoven, the Netherlands), which displayed the pictures on 160X160 and 256X256 matrices, respectively. Transparencies with a diminution factor of approximately ½ were used for assessment. Three observers separately judged the 182 CT scans.

Agreement between pairs of observers for each cistern, fissure, and ventricle was calculated with \( \kappa \) statistics. The observed agreement was corrected for chance agreement, with \( \kappa \) values ranging from 0 (total disagreement) to 1 (perfect agreement). If all degrees of disagreement are of equal importance, \( \kappa \) is expressed as \( (p_D - p_C)/(100 - p_C) \), where \( p_D \) is the percentage observed agreement and \( p_C \) is the percentage agreement expected by chance. Weighted \( \kappa \) is used when the degree of disagreement is taken into account. In our calculations, a linear weight factor of \( \frac{1}{2} \) was given to a difference of one grade in the four-category grading scale. We also calculated \( \kappa \) values for two contracted scales: \( \kappa_3 \) for the three categories 0, (1+2), and 3 and \( \kappa_2 \) for the two categories 0 and (1+2+3). \( \kappa \) values were adjusted for missed observations. Interobserver agreement for the sum scores was studied with Spearman rank correlation coefficients.

Results

Of the 1,820 subarachnoid cisterns and cerebral fissures, 83 (5%) were considered inadequately visualized for grading by one or more observers; four cisterns and fissures (0.2%) could not be graded by any observer. Only one of the 182 CT scans (0.5%) had more than two inadequately visualized cisterns or fissures. Of the 728 ventricles, 15 (2%) were considered inadequately visualized by one or more observers, and three ventricles (0.4%) could not be graded by any observer. There were no CT scans with more than one inadequately visualized ventricle.

Table 2 shows the interobserver agreement on scores for the 10 individual cisterns and fissures, represented by unweighted \( \kappa \) and weighted \( \kappa_w \) values for each pair of observers. Table 3 contains similar results for the four ventricles. \( \kappa_3 \) values for observer pairs A-C and B-C are generally lower than those of pair A-B. Within observer pairs there were no striking differences for individual cisterns, fissures, or ventricles except for relatively low agreement for the basal parts of the sylvian fissure (all pairs: \( \kappa_3 0.37-0.52 \)), the interhemispheric fissure (pairs A-C and B-C: \( \kappa_3 0.38 \) and 0.35, respectively), and the left and right part of the suprasellar cistern (pairs A-C and B-C: \( \kappa_3 0.35-0.43 \)). \( \kappa_3 \) values for the other cisterns and fissures and for the ventricles were relatively good, that is, between 0.45 and 0.74.

In general, the \( \kappa \) values increased considerably when we contracted the scale (data not shown). The mean increases in \( \kappa_3 \) and \( \kappa_2 \) values for the 10 cisterns and fissures was 0.09 and 0.17, respectively. For the four ventricles the increases were 0.03 and 0.06, respectively.

\( \kappa \) values adjusted for inadequately visualized cisterns, fissures, and ventricles (data not shown) were only slightly lower than the values shown in Tables 2 and 3 (differences 0.00-0.05 for pair A-B and 0.00-0.06 for pairs A-C and B-C).
For each CT scan we calculated the sum score for subarachnoid blood and for intraventricular blood (data not shown). The Spearman rank correlation coefficients were very high (pair A-B: 0.92 and 0.87, pair A-C: 0.91 and 0.75, and pair B-C: 0.89 and 0.74, respectively). To study the possibility that one observer had scored consistently higher than the other two, we calculated the mean differences between the sum scores of each observer pair (Table 4). Ideally, this mean difference would be 0. For each observer pair the mean difference between the sum scores was roughly one grade or less, which was very good, considering the ranges of possible sum scores.

**Discussion**

There is no consensus on what constitutes an acceptable $\kappa$ value,\textsuperscript{26–28} but Landis and Koch\textsuperscript{29} have proposed that values between 0.40 and 0.80 can be considered to indicate moderate to substantial agreement. In our study, the $\kappa$ values for most cerebrospinal spaces were within this range. On the other hand, one should be cautious in comparing $\kappa$ values from different studies because the values depend not only on the actual agreement between observers but also on the number of categories used in their computation.

Agreement was better for the ambient and quadrigeminal cisterns, the lateral parts of the sylvian fissure, and the ventricles than for the basal parts of the sylvian fissure, the interhemispheric fissure, and the suprasellar cistern. The difference suggests that low $\kappa$ values cannot be attributed solely to problems in applying the semiquantitative four-category grading system. The lower interobserver agreement in grading the basal parts of the sylvian fissure, the suprasellar cistern, and the interhemispheric fissure should probably be attributed to anatomic and tech-
numerical factors. Because the first two structures lie just above the base of the skull, it is difficult to identify small amounts of blood and on some CT slices these cisterns are inadequately visualized. Grading of the interhemispheric fissure can also be complicated by a calcified falx or by the presence of blood in the frontal lobe(s).

The number of inadequately visualized cisterns, fissures, and ventricles was relatively low and hardly affected the $\kappa$ values. As expected, $\kappa_w$ values and $\kappa_s$ and $\kappa_i$ values indicated better interobserver agreement. The error introduced into the sum score when an interpolated grade is assigned to an inadequately visualized cistern or fissure is small and therefore acceptable as long as only one or two of the 10 structures are inadequately visualized. The range of sum scores calculated from many more than four categories (31 for subarachnoid blood and 13 for intraventricular blood) is too large to calculate $\kappa$ statistics. However, the Spearman rank correlation coefficients for the sum scores were very high and, together with the low mean differences between the sum scores of observer pairs, represent good interobserver agreement.

In summary, our method has some drawbacks for the study of the exact anatomic distribution of subarachnoid blood, but it can be used without restriction as a measure for the total amount of subarachnoid and intraventricular blood. In multicenter studies with many observers, the use of three or two categories per cistern or fissure will probably give satisfactory interobserver agreement. We obtained experience with the four-category grading system in some of our earlier studies,\textsuperscript{30–32} in which the sum scores proved to be important prognostic factors. The use of the entire range of possible sum scores as prognostic categories will be impracticable. Broader categories can be chosen, the number of which depends on the kind of study for which the grading system is used.

Acknowledgments

We thank Mrs. I. van der Tweel and Dr. H. Frericks for their help with statistical calculations.

Table 4. Differences Between Subarachnoid and Intraventricular Sum Scores for Three Observer Pairs

<table>
<thead>
<tr>
<th>Observer pair</th>
<th>Subarachnoid blood</th>
<th>Intraventricular blood</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Difference</td>
<td>95% Confidence interval</td>
</tr>
<tr>
<td>A-B</td>
<td>0.86</td>
<td>0.45 to 1.26</td>
</tr>
<tr>
<td>A-C</td>
<td>-0.24</td>
<td>-0.69 to 0.22</td>
</tr>
<tr>
<td>B-C</td>
<td>-1.10</td>
<td>-1.59 to 0.60</td>
</tr>
<tr>
<td>A-B</td>
<td>0.06</td>
<td>-0.07 to 0.19</td>
</tr>
<tr>
<td>A-C</td>
<td>-0.27</td>
<td>-0.46 to -0.09</td>
</tr>
<tr>
<td>B-C</td>
<td>-0.33</td>
<td>-0.52 to -0.15</td>
</tr>
</tbody>
</table>

References

21. Adams HP, Kassell NF, Torner JC, Salls A: CT and clinical correlations in recent aneurysmal subarachnoid hemorrhage:


KEY WORDS • interobserver agreement • tomography, x-ray computed • subarachnoid hemorrhage
Grading the amount of blood on computed tomograms after subarachnoid hemorrhage.
A Hijdra, P J Brouwers, M Vermeulen and J van Gijn

*Stroke*. 1990;21:1156-1161
doi: 10.1161/01.STR.21.8.1156

*Stroke* is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 1990 American Heart Association, Inc. All rights reserved.
Print ISSN: 0039-2499. Online ISSN: 1524-4628

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://stroke.ahajournals.org/content/21/8/1156

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in *Stroke* can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to *Stroke* is online at:
http://stroke.ahajournals.org//subscriptions/