The prognosis of patients with SAH is determined in large part by the severity of the initial bleeding and the occurrence of rebleeding or cerebral ischemia. The amount of blood present on the diagnostic CT scan has been noted as a significant determinant in the assessment of injury severity and is a major factor in the development of cerebral ischemia by means of vasospasm. There have been many methods devised to assess the amount of blood present in the subarachnoid space. Most of these methods are based on scales, which qualitatively estimate the amount and location of the bleeding based on subjectively ap-
Subarachnoid hemorrhage volumetric analysis

plied criteria. These scales have been criticized owing to their lack of reliability, especially the lack of interobserver reproducibility due to sometimes confusing or misleading grading criteria. In an attempt to reach a more objective assessment of hemorrhage volume, a semiquantitative scale was designed by Hijdra et al. This scale has been found to provide better reliability and prognostic accuracy than previous qualitative scales. Despite the development of this semiquantitative scale, blood clot measurement remains subject-dependent, thus hampering appropriate comparisons between series of patients and different management strategies.

A more objective way to assess the severity of the bleeding would be to measure the actual volume of the bleeding in the different intracranial compartments, as has been tried with other intracranial masses. The quantitative assessment of bleeding volume in patients with SAH has been performed by several authors, first in a limited number of cases, and recently in a larger group of patients. However, there has been no analysis of the reliability of these volumetric measurements in terms of interrater or intermethod reproducibility, and their comparison with other scales is limited. Therefore, the objective of this study was to assess the feasibility and reliability of 2 quantitative methods to measure bleeding volume in patients who have suffered spontaneous SAH, and to compare them to other qualitative and semiquantitative scales in terms of reliability and accuracy in predicting DCI and outcome.

Methods

Patient Sample

The study included a prospective series of patients consecutively admitted to Hospital 12 de Octubre in Madrid, Spain, over a 4-year period (2006–2010). All of these patients had a diagnosis of SAH, and diagnostic CT was able to be performed in the first 24 hours after the onset of the symptoms. The diagnosis of SAH was made by the presence of a typical clinical history of spontaneous SAH and the presence of intraventricular blood or blood in the basal cisterns on the admission CT scan. The onset of the hemorrhage was considered the point at which neurological symptoms or clinical deterioration began. In awake and cooperative patients, this information was provided by the patients themselves. In unconscious, disoriented, or dysphasic patients, this information was obtained by relatives who had witnessed the moment of deterioration. Exclusion criteria for this study were age below 18 years, history of previous SAH or stroke, and nonaneurysmal SAH. All procedures were reviewed by our local institutional ethical committee and informed consent was obtained in all patients.

All patients were treated in accordance with published national guidelines for the treatment of aneurysmal SAH with aneurysms generally secured by endovascular or surgical means within 24–72 hours after the beginning of symptoms. All patients received oral or intravenous nimodipine and were treated in an ICU. No corticosteroids or prophylactic anticonvulsant therapy were used.

Data Collection, DCI Definition, and Outcome Assessment

Data corresponding to clinical characteristics such as age, sex, history of previous hypertension, and medical history were recorded at admission. Patients were evaluated in the emergency department and their clinical situation was assessed by the Glasgow Coma Scale and the WFNS scale. The clinical condition at admission was considered good (WFNS Grades I, II, or III) or poor (WFNS Grades IV or V). The presence and location of the aneurysm was identified by admission CT angiography or cerebral angiography. Delayed cerebral ischemia was defined as clinical deterioration attributable to vasospasm when other causes of deterioration could be ruled out such as hydrocephalus, rebleeding, or electrolyte abnormalities, or by the detection of a new infarct on brain CT related to vasospasm that was not visible on the admission or immediate posttreatment scan, or both. Delayed cerebral ischemia was diagnosed by the treating physician of the patient and prospectively recorded. Outcome was evaluated 6 months after the bleeding using the Glasgow Outcome Scale.

Image Analysis

All CT images were evaluated by 2 independent observers who were not aware of the clinical situation or outcome of the patients. The images were stored in DICOM format and analyzed on a personal computer using Analyze software (version 9.0, AnalyzeDirect). Two different quantitative methods were used for estimating the aneurysmal bleeding volume: ROI volume and the Cavalieri method.

In the ROI method, regions of hemorrhage on CT scans were delineated in each slice using a semiautomatic threshold of density. The Analyze system uses semiautomatic segmentation tools such as region growing, edge tracing, and connected thresholding tools, as well as a manual pixel selection tool, all of which were used to segment data in this study. The semiautomated method consists of selecting a pixel inside the volume to be segmented, which is called a seed. The software then automatically connects the neighboring pixels of the initial seed that have a similar intensity. The observer chooses the intensity threshold to be selected to obtain the entire region to be segmented in each slice. Therefore, from the initial point or seed, a region of similar intensity is developed that is chosen by the observer in a semiautomatic region-growing algorithm. As an adjunct to this semiautomatic method, the ROI in each slice can be modified by edge-tracing or limit-tracing tools to better adjust the ROI to the actual area to be segmented. The different hemorrhagic areas are added up slice by slice automatically by the computer to obtain the final different volumes. For a better illustration of the method a video is presented (Video 1).

VIDEO 1. Clip demonstrating the use of the ROI measurement method for the estimation of intraventricular and cisternal bleeding volume in SAH. Click here to view with Media Player. Click here to view with Quicktime.

With this methodology, the maximum time spent to segment all ROI volumes in an individual case does not exceed 15 minutes.
The Cavalieri method consists of applying a grid of points to each section, and the observer indicates (counts) the number of points inside the area to be measured. Each point has an area associated with it, and from this area a hemorrhagic area by slice can be calculated by counting the points in each slice. The total volume is calculated by multiplying the area by the slice thickness. The Analyze software implements automated algorithms for both methods. For a better illustration of this second method a video is presented (Video 2).

**Patient Demographics**

A total of 150 patients were included during this 4-year period. The median patient age was 52, and most patients (62%) were female. Forty-one percent of the patients had a poor clinical grade at admission. Most of the patients were classified as Grade 3 or 4 on the Fisher scale. Poor outcome was found in 73 patients (49%). One hundred twenty patients survived for more than 48 hours and were included in the analysis for factors related to DCI. Delayed cerebral ischemia was present in 47 patients (31%) and was related to poor outcome: 52% of patients presenting with DCI suffered a poor outcome, whereas only 25% of patients without DCI suffered a poor outcome.

**Reliability of Volumetric Methods and Qualitative and Semiquantitative Scales**

Qualitative scores showed a moderate interobserver reproducibility as expressed by their weighted $k$ indices that were always below 0.65 (Table 1), while the semiquantitative and quantitative scores had a very strong interobserver reproducibility. All quantitative measures showed a significantly higher reproducibility than qualitative scores, both when analyzing subarachnoid clot volume and the total volume of bleeding. Furthermore, reliability was very high for all quantitative measures, as expressed by the ICCs obtained both for intermethod (ROI measurements compared with those using the Cavalieri method) and interobserver agreement (Table 2). This reliability is significantly higher for quantitative measurements than for semiquantitative ones (Hijdra score), both for cisternal and total bleeding volumes.

Intermethod agreement between volumes obtained by the Cavalieri method and the ROI method was larger for lower volumes and decreased with larger volumes (Fig. 1A and D), especially when volumes were larger than 25 ml, both in cisternal bleeding volume and total bleeding volume. However, more than 95% of the differences between methods for both volumes were inside the 95% CI and showed a normal distribution (Fig. 1B and E). Agreement was reached in 95% of the observations within approximately ± 6 ml in both total and cisternal bleeding volume (Fig. 1C and F).
Subarachnoid hemorrhage volumetric analysis

<table>
<thead>
<tr>
<th>Scale/Score</th>
<th>κ Value</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>qualitative</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fisher scale</td>
<td>0.64</td>
<td>0.50–0.76</td>
</tr>
<tr>
<td>modified Fisher scale</td>
<td>0.59</td>
<td>0.47–0.70</td>
</tr>
<tr>
<td>Claasen scale</td>
<td>0.61</td>
<td>0.49–0.73</td>
</tr>
<tr>
<td>quantitative and semiquantitative*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAH only</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cisternal Hijdra score</td>
<td>0.80</td>
<td>0.73–0.85</td>
</tr>
<tr>
<td>Cavalieri-ROI method</td>
<td>0.88†</td>
<td>0.81–0.94</td>
</tr>
<tr>
<td>cisternal bleeding volume (Cavalieri)</td>
<td>0.88†</td>
<td>0.82–0.94</td>
</tr>
<tr>
<td>total bleeding volume</td>
<td></td>
<td></td>
</tr>
<tr>
<td>total Hijdra score</td>
<td>0.82†</td>
<td>0.77–0.87</td>
</tr>
<tr>
<td>total bleeding volume (ROI)</td>
<td>0.92‡†</td>
<td>0.88–0.96</td>
</tr>
<tr>
<td>total bleeding volume (Cavalieri)</td>
<td>0.91†‡†</td>
<td>0.88–0.94</td>
</tr>
</tbody>
</table>

* Obtained by analyzing quartile score distribution between observers.
† Significant difference compared with all qualitative measures (p < 0.01).
‡ Significant difference compared with total Hijdra scores (p < 0.05).

Interobserver agreement for cisternal bleeding volume (Fig. 2) and total bleeding volume (Fig. 3) as measured using the ROI method shows a very similar behavior. Differences between observers were lower for smaller volumes and higher for larger volumes of bleeding (Figs. 2A, 2D, 3A, and 3D). Also, differences between observers showed a normal distribution (Figs. 2B, 2E, 3B, and 3E), and 95% of agreement between observers was reached at a difference of approximately ± 6 ml (Figs. 2C, 2F, 3C, and 3F). This value represents 11% of the range of cisternal bleeding volume obtained and only 4% of the range of total bleeding volume obtained. After analyzing the interobserver reliability of the Hijdra semiquantitative score for both cisternal Hijdra scores (Fig. 2G–I) and total Hijdra scores (Fig. 3G–I), although differences were evenly distributed along all scores and showed a normal distribution, 95% of agreement was obtained around the value of 10 points, which was a third of the actual range of scores.

Relationship of Bleeding Volume to the Occurrence of DCI and Outcome: Comparison With Other Scales

The relationship between the main epidemiological characteristics and bleeding volumes with outcome and development of DCI is presented in Table 3. Clinical factors typically related to poor outcome also appeared to be related in this series, such as poor clinical grade, hypertension, and diabetes. Among included patients, the distribution of demographic factors was not statistically different between patients developing DCI or not. Larger bleeding volumes, including both cisternal and total bleeding volume, were related to a poorer outcome and a higher risk of developing DCI. This relationship between bleeding volume and outcome was further studied by determining the ORs for poor outcome and the development of DCI for each quartile of total bleeding volume (ROI) obtained (Table 4). The proportion of patients suffering DCI or a poor outcome increased with each total bleeding volume (ROI) quartile, and therefore there were increased odds of developing DCI or achieving a poor outcome as total bleeding volume (ROI) increased. This relationship was maintained after adjusting for the main clinical factors related to outcome. The discriminative capacity for both outcome and DCI for the different radiological scales and bleeding volumes was studied by plotting the ROC curve (Fig. 4) and analyzing the different AUCs (Table 5). Quantitative analysis of total bleeding volume achieved the highest AUC, and the discriminative ability of this value was higher than that observed for qualitative scales, for predicting both the development of DCI and outcome.

Discussion

This study shows that it is possible to quantify the amount of cisternal and total bleeding volume in patients suffering from SAH. This quantification has an acceptable reliability between both different quantifying methods and between independent observers, and finally, these bleeding volume estimations are useful for predicting both the development of DCI and outcome.

Since the introduction of CT in the study of SAH, the amount of bleeding has been measured by different qualitative scales.15 These scales used different definitions to classify patients into grades, which described different risks of developing vasospasm and/or poor outcome. The more commonly used scale, the Fisher scale, was found to provide a good correlation with both outcome and the development of DCI. However, this relationship has not been entirely validated by other groups or series of patients, and criticisms of its interobserver reproducibility are common in the literature.3,15,33 There have been sev-
eral issues raised concerning this scale. Some grades do not show clear differences in predicting outcome or DCI, such as between Grades 1 and 2 or between Grades 3 and 4.19,20,22 The presence of a large amount of intraventricular hemorrhage cannot be used as a differentiating factor between patients presenting with thick clots in the cisterns, and thus are classified as Grade 3.13 Also, Grade 4 does not discriminate between small or large intraventricular bleeding volumes, or between intraventricular hemorrhage and intracerebral hematoma. Furthermore, there is a common misconception between Grades 3 and 4 and the original definition of the scale, as many clinicians classify a Fisher Grade 3 CT scan as Fisher Grade 4 only due to the presence of some intraventricular hemorrhage, without taking into account that Grade 3 should be chosen instead of Grade 4 if a patient has profuse cisternal bleeding (localized clots and/or vertical layers of blood 1 mm or greater in thickness), regardless of intraventricular hemorrhage or a hematoma.

Thus, other authors have proposed different modifications of the Fisher scale such as the modified Fisher scale8 and Claassen scale,3 taking into consideration the presence and amount of intraventricular bleeding. Both of these scales have better prognostic capacity for determining the risk of developing vasospasm and poor outcome than the Fisher scale, but they also rely on subjective definitions of the amount of bleeding, and thus do not demonstrate better interobserver reliability than the Fisher scale.15 A semiquantitative approach to the problem of grading the amount of both subarachnoid and intraventricular bleeding was designed by Hijdra et al.9 In this approach, 10 different cisterns are evaluated using a score that describes the amount of blood in each of them, from no blood (0 points) to completely filled with blood (3 points). All scores are summed to establish a final cisternal score, which ranges from 0 to 30. Intraventricular hemorrhage is also quantified in each ventricle in a similar way, determining an intraventricular score that ranges from 0 to 12. Both scores can be summed to obtain a semiquantitative total bleeding score. This approach has proved to be useful in predicting both outcome and the risk of developing DCI in patients with SAH11,20,33 and also presents a better interobserver reliability than qualitative scales,11,33 although the complexity of its scoring process makes it cumbersome to use in everyday clinical practice for some.32

The idea of quantifying the amount of bleeding present on CT scans of patients suffering SAH using computer-based techniques is not new. Broderick et al.2 in 1994, and then Friedman et al.7 in 2002, found that using digitized CT scans, a volume of cisternal or total volume of bleeding of 20–21 ml strongly predicted the development of DCI. Reilly et al.23 in 2004 determined that not only the quantitative analysis of the initial subarachnoid clot but also the percentage of clot clearance per day have a good relationship to the development of DCI, both as independent factors. More recently, Ko et al.14 used personal computer–based software to quantify volumes in DICOM images, and showed that the quantitation of SAH is feasible, and that both cisternal and total bleeding volume are related to both the development of DCI and outcome. Sato et al.29 developed a method to quantify SAH on 3D CT, which yields volume data in less time than with 2D techniques. In this method, hemorrhage volume within the different compartments is determined based on blood Hounsfield units. These investigators found a good correlation between the blood burden and the development of symptomatic vasospasm, although just 4 of their 50 patients developed DCI. Despite these quantification studies, none of them has been able to determine if there is...
subarachnoid hemorrhage volumetric analysis

interobserver or intermethod agreement when using this technology in determining SAH volumes.

The analysis of the reliability of computerized volumetric measures has seldom been taken into consideration, because of both the difficulty in assessing reliability with quantitative measures and also because of the very infrequent use of these techniques in clinical practice. One of the first attempts to determine the reliability of computerized volumetric methods using the ICC and the methodology described by Bland and Altman was conducted by Stochetti et al.30 in head injury lesions. Their results in hematoma volume assessment showed that interobserver agreement is high in low-volume measures whereas it decreases in large lesions. A similar trend occurred in our experience, for both the agreement between methods and between observers. The 95% limits of agreement for the intermethod agreement between the computerized and Cavalieri method ranged from −2.75 ml to 3.89 ml for traumatic lesions, similar to that obtained both for cisternal and total bleeding volume in our paper (−4.2 ml and 4.8 ml for cisternal volume and −5.1 ml and 5.8 ml for total bleeding volume). Hematomas are normally spherical or ellipsoidal lesions that are easy to delimit, and their volume is determined by simple formulas in which their 3 maximal diameters are multiplied, reflecting a very similar result to that obtained by sophisticated computerized methods. However, cisternal bleeding is much more difficult to delimit and its shape determines that no volumetric formula can be applied to estimate its volume. Therefore, one could expect that, although sophisticated, these methods could present a similarly modest interobserver or intermethod agreement as qualitative methods, due to the difficulty in the spatial definition of bleeding in cisterns. The analysis applied in this paper attempts to determine if this is the case, and shows that intermethod and interobserver agreement is good for quantitative measures, for both cisternal and total bleeding volumes. We also found that the Hijdra score shows a better performance in this aspect than other scales, as shown recently by Ibrahim et al.11

Likewise, our data show that quantitative blood burden is a better predictor of DCI and outcome 6 months after SAH compared with the qualitative and semiquantitative scales. Moreover, there has been no overlap in the ORs for DCI and outcome in different hemorrhage volume groups, which suggests a strong association between

Fig. 2. Analysis of interobserver agreement for cisternal bleeding volume for the ROI method (A–C), Cavalieri method (D–F), and semiquantitative cisternal Hijdra score (G–I). CH = cisternal Hijdra score.
Ko et al. also found this robust association between blood burden and DCI, which is the modified Fisher scale equivalent in predicting DCI. On the other hand, they found that the quantitative analysis of the blood clot was a better predictor of outcome 3 months after SAH.

Taken together, blood clot quantitative measurement is not only a good research tool but also a good predictor of outcome and the development of DCI.

**TABLE 3: Factors related to outcome and the development of DCI**

<table>
<thead>
<tr>
<th>Variable</th>
<th>All Patients (n = 150)</th>
<th>Good (n = 77)</th>
<th>Poor (n = 73)</th>
<th>p Value</th>
<th>DCI (n = 120)</th>
</tr>
</thead>
<tbody>
<tr>
<td>median age (IQR)</td>
<td>52 (45–65)</td>
<td>51 (45–62)</td>
<td>54 (46–71)</td>
<td>0.14</td>
<td>53 (46–62)</td>
</tr>
<tr>
<td>male (%)</td>
<td>57 (38)</td>
<td>32 (42)</td>
<td>24 (33)</td>
<td>0.2</td>
<td>20 (43)</td>
</tr>
<tr>
<td>poor grade (%)</td>
<td>60 (40)</td>
<td>14 (18)</td>
<td>46 (63)</td>
<td>&lt;0.01</td>
<td>17 (36)</td>
</tr>
<tr>
<td>hypertension (%)</td>
<td>69 (46)</td>
<td>30 (39)</td>
<td>38 (52)</td>
<td>0.08</td>
<td>21 (45)</td>
</tr>
<tr>
<td>diabetes (%)</td>
<td>18 (12)</td>
<td>6 (8)</td>
<td>12 (16)</td>
<td>0.09</td>
<td>6 (13)</td>
</tr>
<tr>
<td>smoking (%)</td>
<td>42 (28)</td>
<td>25 (33)</td>
<td>16 (22)</td>
<td>0.16</td>
<td>15 (32)</td>
</tr>
<tr>
<td>median cisternal bleeding vol w/ ROI (IQR)</td>
<td>13.2 (4.6–23.1)</td>
<td>8.1 (2.6–16.2)</td>
<td>18 (10–26)</td>
<td>&lt;0.01</td>
<td>16 (11–21)</td>
</tr>
<tr>
<td>median total bleeding vol w/ ROI (IQR)</td>
<td>17.8 (6.4–33.1)</td>
<td>9.2 (2.7–18.6)</td>
<td>28 (18–50)</td>
<td>&lt;0.01</td>
<td>19 (11–33)</td>
</tr>
</tbody>
</table>

* Shading indicates statistical significance. IQR = interquartile range.
Subarachnoid hemorrhage volumetric analysis

TABLE 4: Risk for the development of DCI or poor outcome associated with hemorrhage volume

<table>
<thead>
<tr>
<th>Total Bleeding Volume (ROI)</th>
<th>All Patients (n = 150)</th>
<th>Patients Surviving &gt; 48 hrs (n = 120)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Outcome</td>
<td>Crude</td>
</tr>
<tr>
<td></td>
<td>Good</td>
<td>Poor</td>
</tr>
<tr>
<td>&lt;6.4 ml</td>
<td>31</td>
<td>6</td>
</tr>
<tr>
<td>6.4–17.8 ml</td>
<td>26</td>
<td>12</td>
</tr>
<tr>
<td>17.8–33.1 ml</td>
<td>14</td>
<td>24</td>
</tr>
<tr>
<td>&gt;33.1 ml</td>
<td>6</td>
<td>31</td>
</tr>
<tr>
<td>p value for trend</td>
<td>&lt;0.01</td>
<td>&lt;0.01</td>
</tr>
</tbody>
</table>

* Adjusted for age, poor clinical grade, hypertension, diabetes, and smoking.

of DCI and prognosis in patients suffering an SAH. The use of semiautomatic techniques might be considered as cumbersome and slow, but this limitation may be solved with the development of automatic systems for quantifying lesion volume.

This study has several limitations. First, this study is a retrospective analysis based on a prospectively collected registry. Therefore, some degree of bias is inevitable. To partially address this issue, the 2 observers were blinded to clinical outcome and vasospasm development when analyzing the CT scans. Also, the effect of treatment on DCI has not been evaluated. Although no specific effort was made to evacuate a subarachnoid clot during aneurysm surgery other than in the cisterns that needed to be opened for clip placement, it must be noted that the clot clearance rates may not follow the natural history of SAH and may also differ from the coiled aneurysm because there is no clot removal in these cases. Clot evacuation during surgery may have some effect on the occurrence of DCI and final outcome.\(^3\) Another limitation of the methods introduced for obtaining the bleeding volume is that they are not fully automated and therefore they can be viewed as burdensome. This fact could determine that they will be reserved just for scientific purposes and will not be generally adopted. Analyze software allows a fully automated region-growing algorithm based on the automatic growth of a region of pixels of similar intensity. This method was attempted with several patients. However, the automated segmentations produced did not adjust properly to the bleeding and therefore this method was discarded. Also, in our experience, the time spent on the ROI method is limited. However, it is desirable that a fully automated algorithm for volume calculation based on pixel intensity could be developed.

There are multiple strengths of this study. As opposed to previous studies that only included patients with a Fisher grade of 3, or 2 or 3 different scales, we have included every possible consecutive patient with SAH.
different scales and quantitative measures, and to determine the amount of blood burden using several scales. To the best of our knowledge, this is the first study across all SAH grades and have measured the blood clot using qualitative, semiquantitative, and quantitative scales. To the best of our knowledge, this is the first study to determine the amount of blood burden using several different scales and quantitative measures, and to compare them in terms of reliability and accuracy in predicting DCI and outcome.

**Conclusions**

The use of quantitative measures may reduce interobserver variability in comparison with categorical scales. They are feasible using dedicated software, and they show a better prognostic capability related to outcome and DCI than conventional categorical scales. A better option for determining the bleeding load of patients when volumetric quantitation is not possible is the use of the semiquantitative Hijdra score, which shows both acceptable interobserver agreement and prognostic capability.

**Disclosure**

This work was supported by the Spanish Ministries of Health under grant FIS PI 10/01939 and by Fundación Mutua Madrileña grant 2007/082.

Author contributions to the study and manuscript preparation include the following: Conception and design: Lagares, Gómez, Lobato. Acquisition of data: Lagares, Jiménez-Roldán, Ramos, Munarriz. Analysis and interpretation of data: Lagares, Jiménez-Roldán, Munarriz. Drafting the article: Jiménez-Roldán, Alén, Gómez. Critically revising the article: all authors. Reviewed submitted version of manuscript: all authors. Approved the final version of the manuscript on behalf of all authors: Lagares. Statistical analysis: Lagares.

**References**

Subarachnoid hemorrhage volumetric analysis


Supplemental online information:

Please include this information when citing this paper: published online September 21, 2012; DOI: 10.3171/2012.8.JNS12100.

Address correspondence to: Alfonso Lagares, M.D, Ph.D., Department of Neurosurgery, Hospital 12 de Octubre, Universidad Complutense de Madrid, Avda de Cordoba s/n, Madrid, Spain 28041. email: algadoc@yahoo.com.