Mass volume measurement in severe head injury: accuracy and feasibility of two pragmatic methods

Nino Stocchetti, Marco Croci, Diego Spagnoli, Felicetta Gilardoni, Federico Resta, Angelo Colombo

Abstract

**Objective**—To assess the clinical feasibility and the accuracy of two pragmatic methods in comparison with a conventional computer based method of measurement of masses from CT.

**Methods**—Nineteen CT scans of 11 patients with severe head injury, showing 34 traumatic lesions, were examined. The volume of every lesion was digitally measured, then a panel of three examiners independently repeated the measurement using the ellipsoid and the Cavalieri method in random order.

**Results**—All the lesions were identified by all the readers and the mean volume measured by each examiner differed by less than 1.5 ml. The average reading time for each scan was 4 minutes for the ellipsoid and 7 minutes for the Cavalieri method. The average volume of the lesions was 34.2 (SD 35) ml with the digital system, and 38.4 (SD 41) ml and 34.8 (SD 36) ml for the ellipsoid and the Cavalieri readings respectively. The average difference between the applied technique and the digital system was 0.57 (SD 9.99) ml for the Cavalieri direct estimator and 0.20 (SD 15.48) ml for the ellipsoid method. The 95% confidence interval for this difference fell between −2.75 and 3.89 ml for the Cavalieri, and between −4.94 and 5.35 ml for the ellipsoid method. There were 19 lesions >25 ml; the ellipsoid method identified 16 of them, whereas 17 were classified with the Cavalieri method. When considering individual lesions rather than the average volume, discrepancies were detected with both methods. The ellipsoid method was less precise, especially when extracerebral lesions were measured.

**Conclusions**—Both pragmatic methods are inferior to computer based reading, which is the choice when accurate volume estimation is necessary. However, if a digital volumetric determination of the lesions using a CT computer is not possible, the two pragmatic methods offer an alternative.

(J Neurol Neurosurg Psychiatry 2000;68:14–17)

Keywords: head injury; computed tomography; mass lesions

Due to the physical properties of the dura and the skull, every acute increase of volume in the intracranial space may lead to intracranial hypertension. This is especially relevant after traumatic brain injury, when the development of post-traumatic intracranial masses is the most dangerous complication.

In head injury, the volume of masses has been identified as one of the prognostic factors, and a CT grading system accepted worldwide incorporates volume measurements. When a cerebral contusion or an extra-axial collection increases, ample variations in volume are easily detected without meticulous measurement. When subtle changes develop over time, on the other hand, the precision provided by more scientific measurement techniques offers a more valuable basis for clinical decision making.

The calculation of a given area, and the reconstruction of a delimited volume, are quite feasible with modern CT apparatus, especially when a dedicated workstation is part of the machinery. Unfortunately, volume determinations are seldom performed during routine CT readings, although the digital measurement of the lesions is the most accurate system available today. Finally, digital measurement proves impossible or cumbersome when multicentre trials are conducted, and central reading of CT is planned.

There is, therefore, a need for pragmatic methods of volume measurement on the CT printouts, provided that those methods are accurate and feasible.

The ellipsoid method (EM) was developed to calculate the volume of arterovenous malformations. It is based on the concept that the volume of an ellipsoid is about one half the volume of the parallelepiped into which it is placed. By measuring three diameters of a given lesion in the arterial phase of an angiogram, a parallelepiped is reconstructed, and its volume, divided in half, is close to the real volume of the malformation. By extending this geometrical concept from angiography to CT scan, calculation of space occupying lesions becomes possible (fig 1 A).

More recently, the Cavalieri direct estimator (CDE) has been introduced. It breaks down the lesion on CT into a corresponding number of points, through special grids; the volume of a lesion appearing on the scan is the product of the sum of the points that fall on the lesion, the area associated with each point, and the distance between scan slices (fig 1 B). The grid can be constructed by photocopying a template provided in the original article or by preparing a uniformly spaced point grid (by computer or by hand) to be copied onto an overhead transparency.
As volume measurement is clinically important, but rarely performed in the neuroradiological practice, a posteriori methods, feasible even outside the neuroradiology department, may be useful or necessary. We have therefore assessed the performances and the feasibility of two pragmatic methods (the EM and the CDE) in measuring the volume of mass lesions in cases of patients with severe head injury.

**Materials and methods**

Eleven patients with severe head injury admitted to the neuroscience intensive care unit of a university hospital in 1997 were studied. They were eight men and three women, mean age 53 (SD 19) years, with a median Glasgow coma score after resuscitation of 6. Each patient underwent CT at admission, and repeated controls thereafter. Among the scans of these patients, 19 containing mass lesions were selected. A lesion was defined as a hyperdense or mixed (hyperdense and hypodense) area. Each scan was digitally processed (CT Scan Pace, General Electric, Maryland, USA) to determine the volume of every lesion. Such measurements were taken by a neuroradiologist, who traced every lesion on the screen and calculated the area and volume of interest. That was accepted as the reference gold standard. The results of these measurements were not disclosed to the panel of examiners.

Three members of the intensive care unit staff were trained in the identification and measurement of areas and volumes on the CT slices by using the two pragmatic methods. As part of the training, the examiners performed collective readings of several scans of patients not participating in the study, and repeated measurements of lesions using the CDE and the EM under the guidance of a neuroradiologist. Grids for the CDE were prepared and provided to the examiners. Relevant references in the literature were reviewed. After this training period, which required about 2 weeks, the 19 CT scans and appropriate examination forms were submitted to each panellist separately. The CDE and EM were used in a random order.

Both the volumes measured in each scan and the reading time for each method were recorded in the forms. The data forms were then collected, summarised, and compared with the gold standard.

To assess the agreement between the two pragmatic methods and the gold standard the statistical method devised by Bland and Altman was used. A two way analysis of variance (ANOVA) was used for measuring the interobserver variability. The Pearson χ² test with Yates continuity correction was used for comparing percentages.

**Results**

The 19 CT scans studied comprised 23 contusions/lacerations, 10 subdural hematomas, and four extradural hematomas. The range of the lesions directly computed at the CT scan varied from 3 ml to 124 ml; the sum of all lesions produced a total volume of 1260.65 ml, with an average volume for every lesion of 34.23 (SD 35.88) ml.

**READER AGREEMENT AND READINGS**

The readings performed by the three examiners led to consistent measurements. The results of their calculations were very close; the mean volume measured by each examiner differed by <1.5 ml. The ANOVA analysis confirmed this finding, excluding any significant difference among the readers in the use of both methods. All 37 lesions were identified by all readers, providing 222 pragmatic measurements to be compared with the gold standard.

The measurements calculated by the three readers with each method for every lesion were averaged; this average represented the volume calculated with the method. The range of the lesions measured with the EM varied from 0.9...
indicates one or more lesions detected at the CT scan. All values are in ml.

corresponding to a perfect agreement between the two methods of measurement. Each point
in which the gold standard is reported. The zero line represents the line of equality,
volume measured with the computerised method. This di-

Figure 2 The y axis shows the difference between the volume read with the CDE and the volume measured with the computerised method. This difference is plotted against the x axis, in which the gold standard is reported. The zero line represents the line of equality, corresponding to a perfect agreement between the two methods of measurement. Each point indicates one or more lesions detected at the CT scan. All values are in ml.

Figure 3 The y axis shows the difference between the volume read with the EM and the volume measured with the computerised method. This difference is plotted against the x axis, in which the gold standard is reported. The zero line represents the line of equality, corresponding to a perfect agreement between the two methods of measurement. Each point indicates one or more lesions detected at the CT scan. All values are in ml.

to 147.4 ml; the sum of all lesions produced a
total volume of 1273.87 ml, with an average volume
for every lesion of 34.43 (SD 40.98) ml. The range of the lesions measured with the
CDE varied from 4.8 ml to 150 ml; the sum of all
lesions produced a total volume of 1287.42
ml, with an average volume for every lesion of
34.8 (SD 36.6) ml.

AGREEMENT OF THE TWO METHODS WITH THE
GOLD STANDARD
The agreement between the CDE and EM and the
computed reading was acceptable on average,
but questionable in some cases.
The average difference between the applied
technique and the reference value (termed bias, a measure of systematic error with the
statistical method of Bland and Altman) was
0.57 (SD 9.99) ml for the CDE and 0.20 (SD
15.48) ml for the EM. The 95% confidence
interval (95% CI) for this bias fell between
-2.75 and 3.89 ml for the CDE, and between
-4.94 and 5.35 ml for the EM. The maximum
difference between the pragmatic readings and
the reference value was in the order of 25 ml
with the CDE and reached 35 ml with the EM.
A graphical description of the agreement
between each pragmatic method and the gold
standard is illustrated in figs 2 and 3. In these
plots the distance between the line of equality,
lying on the zero difference between the pragmatic method and the reference reading,
expresses the discrepancy between each
measurement and the true value. There are
discrepancies among the two pragmatic meth-
ods and the digital recording in individual
cases. With the CDE these discrepancies are
few and more evident at large volumes; with the
EM the discrepancies are more frequent and
present at lower volumes.
The limits of agreement, an additional
mathematical descriptor of the agreement,
were between -19.41 and 20.55 ml with the
CDE and varied between -30.76 and 31.16 ml
with the EM.

MASSES GREATER THAN 25 ML
Since a volume greater than 25 ml is
considered a threshold for surgical indication,
the capability of appropriately identifying this
cut off value may indicate the clinical reliability of pragmatic methods. Nineteen lesions >25
ml were found from the CT analysis. Seventeen
lesions >25 ml were identified with the CDE
and 16 were detected with the EM. There is
not a statistical significant difference between
these two findings.

EXTRACEREBRAL AND INTRACEREBRAL LESIONS
Volume may become more difficult to measure
in lesions of a particular shape. Extracerebral
lesions, such as subdural haematomas for
example, may be more difficult to measure than
a rounded contusion in the brain parenchyma.
The performances of the two pragmatic meth-
ods in determining the volume of intracerebral
or extracerebral lesions were, in fact, different.
The CDE was more accurate than the EM,
especially in cases of extracerebral masses. The
main data are reported in the table.

READING TIME AND LEARNING EFFECT
The average reading time to complete the
measurement of a CT scan was 7 minutes 44
seconds (SD 42 seconds) for the CDE method
and 4 minutes 27 seconds (SD 38 seconds) for
the EM (p<0.001). As a reference, the average
reading time using the digital system was 2
minutes 54 seconds (SD 26 seconds).

In both methods the reading time for each
CT was higher at the beginning of this research
and decreased as the study proceeded. Such
decrease was probably due to a learning effect,
as the accuracy of the measurement did not
change over time. The examiners became more
experienced and read the CT faster, but not at
the price of reducing the precision of their
analysis. The reduction in reading time was
37% with the CDE method, and 48% with the
EM.

Intracerebral versus extracerebral lesions

<table>
<thead>
<tr>
<th></th>
<th>Mean Volume (SD)</th>
<th>Bias (SD)</th>
<th>Limits of agreement</th>
<th>95% CI for the bias</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gold standard:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intracerebral</td>
<td>23</td>
<td>28.7 (30)</td>
<td>-1.91 (0.07)</td>
<td>-21.4 to +17.6</td>
</tr>
<tr>
<td>Extracerebral</td>
<td>14</td>
<td>43.3 (43)</td>
<td>-0.20 (3.3)</td>
<td>-13.8 to +23</td>
</tr>
<tr>
<td>Cavalieri direct estimator:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intracerebral</td>
<td>23</td>
<td>27.5 (25)</td>
<td>-1.91 (0.07)</td>
<td>-21.4 to +17.6</td>
</tr>
<tr>
<td>Extracerebral</td>
<td>14</td>
<td>47.9 (48)</td>
<td>4.64 (9.2)</td>
<td>-13.8 to +23</td>
</tr>
<tr>
<td>Ellipsoid method:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intracerebral</td>
<td>26.7 (28)</td>
<td>-1.19 (14.3)</td>
<td>-29.8 to +27.4</td>
<td>-7.4 to +5</td>
</tr>
<tr>
<td>Extracerebral</td>
<td>45.8 (55)</td>
<td>2.49 (17.5)</td>
<td>-32.6 to +37.6</td>
<td>-7.7 to +12.6</td>
</tr>
</tbody>
</table>
Mass volume measurement in severe head injury

Discussion

The volume of intracranial lesions, and its change over time, is of obvious interest in the diagnosis and management of patients with head injury. Specific features of space occupying lesions, including their volume, are associated with the outcome of patients with traumatic brain injury. Based on data from the traumatic coma data bank, Marshall et al. showed that the existence of specific findings from CT, including the presence of masses exceeding 25 ml, was associated with increasing degrees of mortality and morbidity. This scale is now widely used both in clinical practice and in the scientific literature on head injury. Interestingly, this scale requires the measurement of the volume of intracranial lesions, as the threshold of 25 ml is part of the grading.

In some clinical trials recently performed on head injury, a central reading of the CT scan was carried out, and intracranial volume measurements performed. There are clinical problems, such as the evolution of a cortical contusion, that require volume measurement. The overall mass effect, which is considered for clinical management, depends on the size of a given lesion and on the associated swelling and oedema as well, but the size of the lesion itself is crucial.

For these reasons an accurate estimate or, possibly, the measurement of the volume of cerebral masses has become an integral part of the treatment and of the classification of severe head injury.

Unfortunately, the determination of the volume of masses is not performed as an element of the routine reading of CT scans. In a recent multicentre study on head injury involving 18 neurotraumatological centres, all 18 centres claimed to have used the Marshall scale, but the volume of lesions was seldom measured in one centre, and never measured in the other 17 (personal unpublished data).

The data collected in this study show that both pragmatic methods are on average acceptably precise when performed by readers who underwent a short training period. Accuracy was achieved by all readers, as no difference was found among them in their calculation of the volume. The systematic error of the pragmatic methods was <1 ml, and the mean volume of all lesions measured with the EM and CDE was very close to the volume measured digitally. When considering individual lesions rather than the average volume, however, discrepancies were detected with both methods and some discrepancies were relevant especially with the EM.

The CDE is more accurate than the EM. That is proved by the Bland and Altman analysis, as suggested by the comparison of figs 2 and 3. The bias is lower with the EM because this method determines both overestimates and underestimates, leading to a narrow final mean value. By comparing statistical measures of dispersion, such as SD, or specific parameters of the Bland and Altman analysis, such as the limits of agreement, it seems that the CDE provides measurements closer to the reference value.

Similarly, this method performed better in cases of extracranial lesions. The EM was developed to assess the volume of vascular malformations and it is not surprising that it does not excel in the head injury field. The EM shows its weakness mainly when extracerebral collections are concerned. In fact, the shape of these lesions is less suitable for calculation by the ellipsoid theory. By contrast, the CDE method may accurately take into account even the irregular shape of thin lesions.

Such greater accuracy of the CDE technique is counterbalanced by the need for a longer execution time; this system takes twice as long to perform as the EM. The necessary reading time, however, is initially 8 minutes but, after a brief training, it decreases to an average of 5 minutes 11 seconds for each CT. With this reduction due to practice, the reading time of each scan with the CDE remains fully acceptable in the clinical setting.

Both pragmatic methods are inferior to computer based reading, which is the choice when accurate volume estimation is necessary. However, if some sequential volumetric determination of the lesions using a CT computer is not possible, the two pragmatic methods offer an alternative.

They may be mastered after short training, do not require expensive equipment, and provide a different level of accuracy. The CDE performs better than the EM but takes longer. It is recommended whenever accuracy of measurement is crucial for possible therapeutic decisions.

7 Clarebuck RE, Sipos EP. The efficient calculation of neurosurgically relevant volumes from computed tomographic scans using Cavalier’s direct estimator. Neurosurgery 1997; 41:339–43.