RESEARCH PAPER

Performance characteristics of methods for quantifying spontaneous intracerebral haemorrhage: data from the Efficacy of Nitric Oxide in Stroke (ENOS) trial

Kailash Krishnan,[1] Siti F Mukhtar,[1] James Lingard,[1] Aimee Houlton,[1] Elizabeth Walker,[1] Tanya Jones,[1] Nikola Sprigg,[1] Robert A Dineen,[2] Panos Koumellis,[3] Alessandro Adami,[4] Ana M Casado,[5] Philip M W Bath,[1] Joanna M Wardlaw[5]

ABSTRACT

Background Poor prognosis after intracerebral haemorrhage (ICH) is related to haemorrhage characteristics. Along with developing therapeutic interventions, we sought to understand the performance of haemorrhage descriptors in large clinical trials.
Methods Clinical and neuroimaging data were obtained for 548 participants with ICH from the Efficacy of Nitric Oxide in Stroke (ENOS) trial. Independent observers performed visual categorisation of the largest diameter, measured volume using ABC/2, modified ABC/2, semiautomated segmentation (SAS), fully automatic measurement methods; shape, density and intraventricular haemorrhage were also assessed. Intraobserver and interobserver reliability were determined for these measures.
Results ICH volume was significantly different among standard ABC/2, modified ABC/2 and SAS: (mean) 12.8 (SD 16.3), 8.9 (9.2), 12.8 (13.1) cm³, respectively (p<0.0001). There was excellent agreement for haemorrhage volume (n=193): ABC/2 intraobserver intraclass correlation coefficient (ICC) 0.96–0.97, interobserver ICC 0.88; modified ABC/2 intraobserver ICC 0.95–0.97, interobserver ICC 0.91; SAS intraobserver ICC 0.95–0.99, interobserver ICC 0.93; largest diameter: (visual) interadjudicator ICC 0.82, (visual vs measured) adjudicator vs observer ICC 0.71; shape intraobserver ICC 0.88 interobserver ICC 0.75; density intraobserver ICC 0.86, interobserver ICC 0.73. Graeb score (mean 3.53) and modified Graeb (5.22) scores were highly correlated. Using modified ABC/2, ICH volume was underestimated in regular (by 2.2-2.5 cm³, p<0.0001) and irregular-shaped haemorrhages (by 4.8-4.9 cm³, p<0.0001). Fully automated measurement of haemorrhage volume was possible in only 5% of cases.
Conclusions Formal measurement of haemorrhage characteristics and visual estimates are reproducible. The standard ABC/2 method is superior to the modified ABC/2 method for quantifying ICH volume.
Clinical trial registration ISRCTN9941422.

INTRODUCTION

Spontaneous intracerebral haemorrhage (ICH) is usually a severe form of stroke with a high mortality rate (~50%) by the end of the first year.[1] Poor functional outcome is related independently to increased age, systolic blood pressure, stroke severity and reduced level of consciousness. Radiologically, haemorrhage size, haemorrhage expansion, evidence of continuing bleeding (eg, 'spot sign' on CT angiography)[2 3] and presence of intraventricular haemorrhage (IVH) are all also associated with poor outcome.[4 5] The combination of ICH volume and initial Glasgow Coma scale was reported to be the strongest independent predictor of 30-day outcome.[1] Hence, reliable measurement of haemorrhage size is an important component of early clinical management and for stratification in clinical trials.

Although any effective intervention will ultimately need to show an effect on functional outcome (eg, using the modified Rankin Scale (mRS))[6] in phase III trials, earlier developments will need to study the effect of treatment on a surrogate measure. Where the intervention aims to limit haemorrhage expansion, accurate, reproducible and validated measurement of haemorrhage size (thereby allowing measurement of the change in volume) will be critical. Hence, phase II trials of blood pressure lowering and haemostatic interventions have utilised measurement of haemorrhage expansion as a key primary outcome.[7 8]

There are multiple methods for measuring haemorrhage volume on a CT scan, ranging from qualitative visual estimation to computerised automatic measurement. Visual approaches (ie, visual size categorisation based on the largest diameter) are quick while computerised methods might provide a continuous volume measure but require access to good quality electronic image files. A common approach is the ABC/2 method,[9] which has been modified such that haemorrhage slices are excluded if the area is <25% of the largest slice and counted as half if it is 25–75% of the largest slice.[10] Alternatively, semiautomatic or fully automatic computerised methods may be used, usually based on manually outlining or thresholding areas or volumes within regions of interest (ROI). The relative advantages and disadvantages of these approaches reflect the balance between measurement time, type of data required, accuracy and availability of computer workstations, and sample

size, as studied in part.[11–14] Assessment of additional properties of haemorrhage, such as shape and density[15] and extension into the ventricles or subarachnoid space, may also be useful.[16 17]

In this study, we assessed and compared methods for measuring haemorrhage size, shape and density and IVH size. The data come from the 'Efficacy of Nitric Oxide in Stroke' (ENOS) trial,[18] which assessed the management of blood pressure in acute stroke and included patients with ICH.

METHODS
Patients
Data came from the prospective international multicentre ENOS trial, a study of blood pressure management in patients with acute ischaemic stroke or ICH, and high blood pressure.[18] The trial had regulatory and ethics approvals in each participating country, and all patients gave written consent, or a relative provided proxy consent if the patient lacked capacity. The main inclusion criteria were acute stroke with motor deficit, systolic blood pressure of 140–220 mm Hg, and treatment could start within 48 h of onset. Patients were recruited from 173 sites in 23 countries across five continents. Following the entry of baseline demographic and clinical data into a secure internet-fronted database, patients were randomised to transdermal glyceryl trinitrate (5 mg) versus none for 7 days; those taking antihypertensive agents immediately prior to their stroke were also randomised to continue versus stop these, again for 7 days.

Participants had a baseline CT or MRI brain scan as part of clinical care, usually before randomisation. Where possible, a second (research) CT or MRI scan was performed at day 7±1 (end-of-treatment). DICOM, JPEG, PNG or GIF files are sent to the coordinating centre. Any scans sent on film were digitised using a VICOM digitiser (VIDAR Diagnostic PRO Advantage, USA). The following study assesses only those patients with a CT scan. For each patient, images were received with the thinnest slices provided by CT scan machines using standardised protocols in the recruiting centres. However, during the course of the trial, more volumetric images were made available because of advances in imaging technology.

Scan adjudication
Complete CT scan series were made available for assessment by a group of seven adjudicators comprising accredited neuroradiologists or neurologists trained in CT brain imaging assessment in stroke (http://www.neuroimage.co.uk/sirs, with coordination by JMW).[19] Images were viewed over the web, blind to all clinical and treatment information except for patient age, time since onset of stroke and side of brain symptoms. Responses were entered directly into the trial database via a web-based response form. Collected information included whether the patient had an ICH and its location; an estimate of its size (sorted into ordered categories based on the longest diameter in any plane): <3, 3–4.9, 5–8, >8 cm; and the presence of other qualitative findings (mass effect,[19] atrophy,[19] white matter disease,[20] old infarct or haemorrhage) using validated scoring methods.

Scan haemorrhage quantification
Scans were visualised and analysed using OsiriX for Mac (V.3, 32 bit, http://www.osirix-viewer.com[21] on a 26 inch Apple iMac. Two trained observers (KK, SFM) measured baseline ICH volumes of 193 patients using ABC/2, modified ABC/2, semiautomatic segmentation (SAS) and automatic volume calculation (AVC, using three-dimensional (3D) rendering of a stack of two-dimensional slices). Both observers assessed scans blinded to each other's

data, and repeated a proportion of the scans (47 and 34 scans for KK and SFM, respectively) at a different time blinded to their original measurements. Thus, intraobserver and interobserver variation could be estimated. For the remaining 355 patients, one observer (KK) calculated ICH volumes using all four methods.

The ABC/2 method requires measurement of the longest diameter of the haemorrhage in the axial plane ('A'), longest axis at 90° to 'A' in the axial plane ('B'), and the number of slices ('C') showing the haemorrhage (noting that slice thickness may vary through the scan). 'C' is the product of slice thickness and number of slices showing any haemorrhage. Volume is calculated as A×B×C/2.[9] In the modified ABC/2 method, the area of haemorrhage seen on a slice has to be at least 25% of the haemorrhage area seen on the largest slice if the slice is to be considered in 'C'; if the area is more than 75% of the largest area, the slice counts as 1, and 0.5 if the measured value of the area is between 25% and 75%.[10]

For SAS and AVC, it was possible to use thresholds so that the haemorrhage ROI was measured within certain segmentation parameters. Upper and lower attenuation values (typically 40–80 HU) of haemorrhage were established manually by sampling from the haemorrhage and normal brain. Where the threshold did not exclude non-haemorrhage areas (as with the pineal gland, calcified choroid plexus, bone) in close proximity with the haemorrhage, the haemorrhage boundary was manually edited. The area of each slice that was included within these parameters was measured, added together and multiplied by scan thickness to obtain the SAS volume. AVC used an Osirix '3D growing region (entire series)' method after selecting a threshold part of the haemorrhage near to its centre. With all methods, IVH blood was included in the final ICH volume.

Additionally, four haemorrhage characteristics (area, perimeter, mean attenuation and SD of attenuation) were recorded for the haemorrhage from the slice with the largest area of haemorrhage. Haemorrhage shape and density indices were then calculated as:

$$\text{Shape index [ref. 22]} = \text{perimeter}^2 / 4_{\Pi} \times \text{area}$$

$$\text{Density index [ref. 23]} = \text{SD/mean (ie, coefficient of variation)}$$

The divergence in visual appearance of haemorrhage shape and density were also assessed used an ordered categorical scale (1–5) where for an extra lesion edge or degree of density variation an additional point was given on the shape scale or density scale.[15] Similarly, intraventricular blood volume was assessed visually using the Graeb score and modified Graeb ordered categorical scales.[16 17] The Graeb score is calculated by scoring the amount of blood in the lateral ventricles separately (score of 1=trace of blood to a maximum of 4=ventricle full of blood and expanded) and adding this to the scores for the third and fourth ventricles separately (1=blood present and 2=filled with blood and expanded) to derive a maximum score of 12. The modified Graeb score allocates scores for separate ventricular compartments (1≤25% filled with blood to 4≥75% filled with blood) to reflect selective regional accumulation of blood and an extra one point for expansion of each ventricle. The maximum score is 32.

Intraobserver and interobserver reliability was determined for visual categorisation of shape and density. In addition, intraobserver reliability was assessed for measured shape, density index, Graeb, modified Graeb and IVH volume over different reading sessions, separated by a minimum of 14 days. All adjudication and haemorrhage measurements were made blind to baseline and follow-up clinical information, other imaging and treatment assignment.

Statistical analysis

Data are shown as number (%), median (IQR) or mean (SD). Measurement of intraobserver and interobserver variability was assessed using the intraclass correlation coefficient (ICC).[24] A probability value of <0.05 was considered statistically significant. Analyses were performed using SPSS (V.21) and checked with Medistat running on an Apple Mac.

RESULTS

Altogether, 629 patients with ICH were enrolled into ENOS.[25 26] Of these, 548 patients had CT-confirmed ICH and a baseline scan available for measurement; 81 other patients either had ICH diagnosed on MRI or had no CT scan available, and these were excluded. Patient demographic and clinical details are shown in table 1. The mean age of the 548 patients in the present analysis was 67 (SD 12) years, 66% of patients were male, mean baseline blood pressure was 171 (19)/92 (13) mm Hg, and the median time from onset of ictus to performing neuroimaging was 4.5 (IQR 5.6) hours. When adjudicated visually by experts, 63% of haematomas were located in the middle cerebral artery territory (table 1); most haemorrhages caused a mass effect (86%) and many patients had leukoaraiosis (66%) and/or a previous stroke lesion (49%) present on their scans. The most frequent visually assessed haemorrhage length category was 3–5 cm (224, 41.3%), closely followed by <3 cm (220, 40.6%) with much fewer larger haemorrhages. The mean measured haemorrhage length was 3.4 cm (longest diameter; table 1).

ICH volume

ICH volume was significantly different between ABC/2, modified ABC/2 and SAS: 12.8 (mean) (SD 16.3), 8.9 (9.2), 12.8 (13.1) cm^3, respectively (p<0.0001) (table 2). The ICC was 'excellent' at 0.84–0.96 when comparing the observers' measurements for each of ABC/2, modified ABC/2 and SAS (table 2). The observers found haemorrhage volume to be larger, by an average of 2.7–4 cm^3 with standard ABC/2, as compared to modified ABC/2, and 1.4–4 cm^3 smaller with SAS as compared to modified ABC/2. As the mean ICH value increased, the difference between ICH volume measured by modified ABC/2 versus the other two methods increased (figures 1 and 2), but no significant difference was observed between standard ABC/2 and SAS (table 2; figure 3). The slope of the best-fit regression line for the increasing difference between modified ABC/2 versus ABC/2 was −0.44, and −0.36 for modified ABC/2 versus SAS (both p<0.0001) (figures 1 and 2).

There was good intraobserver agreement for measurements assessed on 34–47 scans (5–10% of the total of 548). For ABC/2, modified ABC/2 and SAS, the intraobserver ICC was 'excellent', ranging between 0.97–0.98, and 0.95–0.99, respectively (table 3).

There was excellent interobserver agreement (ICC) based on 193 scans at 0.88 for ABC/2, 0.91 for modified ABC/2 and 0.93 for SAS (table 4). Both ABC/2 and modified ABC/2 showed excellent correlation with SAS (p<0.0001, see online supplementary figures S1–S3). Only 23 of 548 scans were amenable to analysis using the fully AVC method as the software was unable to handle the scan image with varying slice thickness; this approach was therefore ignored in further analyses.

Across the ICH visual size categories, ICH volumes calculated by modified ABC/2 were significantly smaller compared to SAS (figure 4). As the ICH size category increased, the difference between modified ABC/2 and SAS also increased. By comparison, there was no significant difference between ICH volumes

Table 1 Baseline demographic, clinical and neuroradiological factors in 548 patients with primary intracerebral haemorrhage in the ENOS trial

Variable	Data
Demographics	
Age (years)	67.9 (12.1)
Sex, male (%)	360 (65.7)
Country, UK (%)	
Asia	117 (21.4)
Europe	67 (21.4)
Other (Africa, Australasia and North America)	42 (7.7)
UK	322 (58.8)
Clinical findings	
Premorbid modified Rankin Scale=0 (%)	418 (76.3)
Previous stroke (%)	69 (12.6)
Prior antihypertensive medication use (%)	227 (41.4)
Prior history of high BP (%)	341 (62.2)
Diabetes mellitus (%)	67 (12.2)
Ischaemic heart disease (%)	56 (10.2)
Atrial fibrillation (%)	30 (6.5)
Total anterior circulation syndrome (%)	195 (35.6)
SSS (/58)	30.1 (12.3)
National Institutes of Health Stroke Scale (/42)	12.8 (5.3)
Systolic BP (mm Hg)	171.6 (19.3)
Diastolic BP (mm Hg)	92.2 (13.3)
Heart rate (bpm)	77.7 (14.5)
Time, stroke to neuroimaging (h)	4.5 [11.6]
Adjudicated CT scan findings	
Location of haemorrhage (%)	
MCA	346 (63.1)
ACA	22 (4.0)
PCA	5 (0.9)
MCA+ACA	2 (0.4)
Borderzone	14 (2.6)
Lacunar (ie, small subcortical) stroke	142 (25.9)
Brainstem and/or cerebellum	16 (2.9)
Leukoariosis	362 (66.1)
Lesion mass effect (minimal to extreme swelling)	470 (85.9)
Previous stroke lesion	270 (49.3)
IVH	141 (25.7)
Intracerebral haematoma size category (%)	
<3 cm	220 (40.6)
3 to <5 cm	224 (41.3)
5 to 8 cm	87 (16.1)
>8 cm	11 (2.0)
Measured CT scan findings	
Volume, ABC/2 (cm^3)	12.77 (16.32)
Longest diameter (cm)	3.38 (1.4)
With IVH, n=141	
Graeb score (/12)[16]	3.52 (2.4)
Modified Graeb score (/32)[17]	5.19 (4.7)
Without IVH, n=407	
Shape (/5)[15]	3.00 (1.4)
Shape index[22]	1.22 (1.1)
Density (/5)[15]	2.54 (1.3)
Density index[23]	0.19 (0.1)

Data are number (%), median (semiquartile range), or mean (SD).
NIHSS calculated from SSS scores range from 0 (comatose with quadriplegia) to 58 (normal neurological status).
ACA, anterior cerebral artery territory; BP, blood pressure; ENOS, Efficacy of Nitric Oxide in Stroke; IVH, intraventricular haemorrhage; MCA, middle cerebral artery territory; NIHSS, the National Institutes of Health Stroke Scale; PCA, posterior cerebral artery territory; SAS, semiautomatic segmentation; SSS, Scandinavian Stroke Scale.

Table 2 Comparison by two observers of different methods for measuring haemorrhage volume (cm³) on CT scans: ABC/2 versus modified ABC/2 versus SAS

Haematoma volume		Difference	p Value	ICC
Observer 1				
n=548				
ABC/2	Modified ABC/2	3.96	<0.0001	0.84
12.77 (16.32)	8.90 (9.21)			
ABC/2	SAS	−0.01	1.00	
12.77 (16.32)	12.76 (13.06)			
SAS	Modified ABC/2	3.97	<0.0001	
12.76 (13.06)	8.90 (9.21)			
Observer 2				
n=193				
ABC/2	Modified ABC/2	2.67	<0.0001	0.96
12.05 (12.40)	9.70 (10.05)			
ABC/2	SAS	0.05	0.89	
12.05 (12.40)	11.08 (11.38)			
SAS	Modified ABC/2	1.38	<0.0001	
11.08 (11.38)	9.70 (10.05)			

Data are mean (SD), difference (Δ) in haemorrhage volume between ABC/2 and SAS and ICC.
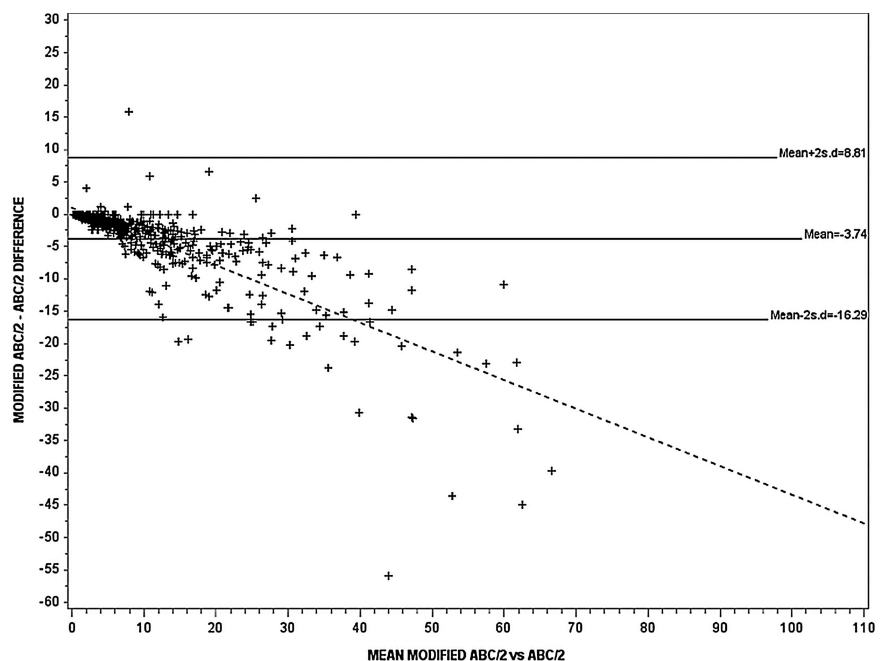ICC, intraclass correlation coefficient; SAS, semiautomated segmentation.

measured by standard ABC/2 and SAS when compared by ICH visual size categorisation (figure 4).

There was good agreement between the adjudicators across ICH visual size categories (n=47, ICC 0.82, p<0.001) (see online supplementary table I) using the ordered categorical scale (<3, 3–4.9, 5–8, >8 cm)[19] and strong agreement between the visual size category and the observer's largest measured diameter in the axial plane 'A' (ICC 0.71, p<0.001) (see online supplementary table II).

ICH volume and shape

The most common shape was irregular (64%), followed by regular ICH. Small ICHs (5.5–8.06 cm³) were more regularly shaped than larger haemorrhages (p<0.0001) (table 5). Using modified ABC/2, haemorrhage volume was significantly lower for regular-shaped haemorrhages when compared with ABC/2 (by 2.2 cm³) and SAS (by 2.4 cm³) (table 5 and figure 5); the difference was greater with larger irregular-shaped haemorrhages (between standard ABC/2 and modified ABC/2 by 4.8 cm³ and SAS and modified ABC/2 by 4.9 cm³). When compared by shape, ICH volume calculated by standard ABC/2 did not differ from that measured by SAS (figure 5).

ICH shape and density in patients with no IVH

Intra-agreements and interagreements were both 'good' for visual assessments of haemorrhage shape: intraobserver ICC 0.88 (see online supplementary table III), interobserver ICC 0.75 (n=47) (see online supplementary table IV); and density: intraobserver ICC 0.86, interobserver ICC 0.73. Intra-observer ICC was 0.53 for calculated shape index and 0.86 for density index (see online supplementary table III).

IVH volume and severity

Intraobserver agreement for assessment of IVH volume on baseline scans from 49 patients was excellent (ICC 0.98–0.99) whether using the Graeb score, modified Graeb score or SAS (see online supplementary table III). The Graeb and modified Graeb scores were highly correlated with each other (rs=0.88, p<0.01). Both were also highly correlated with measured IVH volume using the SAS method: Graeb, (rs=0.73, p<0.01); modified Graeb (rs=0.72, p<0.01).

Comparison of ICH volume by ABC/2, modified ABC/2 to SAS

We performed a non-systematic review to compare our study with previously published work comparing ABC/2, modified ABC/2 and SAS volume measures in ICH[10–14 27–30] (see online supplementary figures S4–S7).

Some studies used the standard ABC/2 formula while others used variations of it;[10 11] Gebel et al[11] used the central haemorrhage slice to measure the largest diameter 'A' and found higher volumes computed by ABC/2 when compared to SAS. With regard to shape, three studies studied errors in warfarin-related haemorrhage due to higher frequencies of irregular shape.[12–14]

Figure 1 The Bland–Altman plot for assessment of variation in estimating haematoma volume between modified ABC/2 and standard ABC/2 (n=548), r²=0.64, p<0.0001. The continuous and dotted lines represent the regression lines. The slope of the best-fit regression line gradient was −0.44 (p<0.0001).

Figure 2 The Bland–Altman plot for assessment of variation in estimating haematoma volume between modified ABC/2 and semiautomatic segmentation (n=548), r^2=0.45, p<0.0001. The continuous and dotted lines represent the regression lines. The slope of the best-fit regression line was −0.36 (p<0.0001).
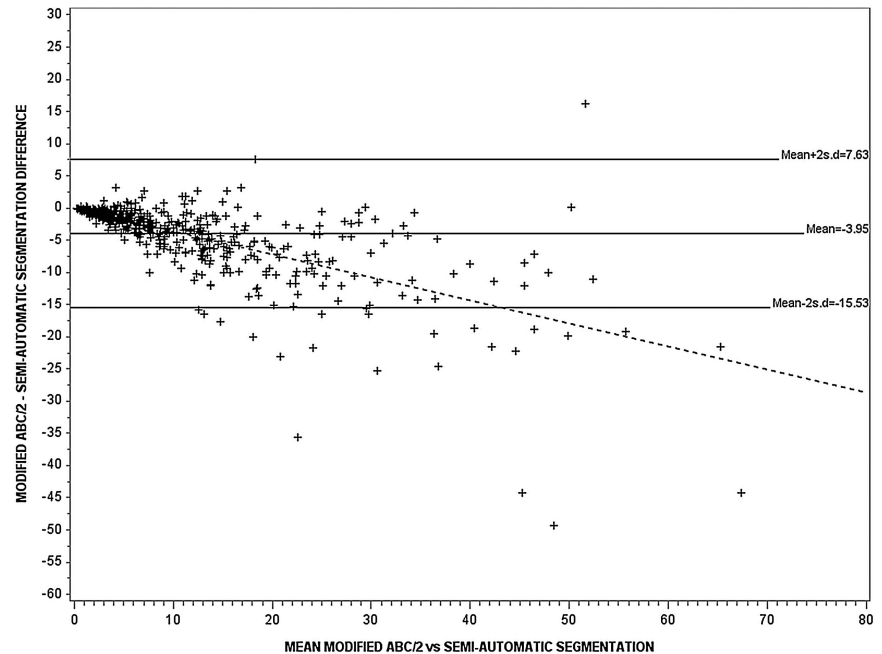
Huttner et al[12] showed that volumes of regular shaped haematomas using ABC/2 were significantly increased in regular and irregular shaped haemorrhages. However, the large slice thickness used during CT image acquisition in this study may have produced significant errors during calculation.[12]

This review comparing all three methods showed that the ABC/2 formula tended to marginally overestimate the ICH volume, but there was no absolute significant difference compared to SAS (see online supplementary figure S4); when assessed by variability in haemorrhage shape, the results follow the same sequence (see online supplementary figure S5). By comparison, ICH volumes computed using modified ABC/2 are significantly smaller when compared with SAS; the difference is greater as the haemorrhage sizes became larger and more irregular (see online supplementary figures S6 and S7).

DISCUSSION

To the best of our knowledge, this is the first study that compares several methods for assessing volume of spontaneous ICH and IVH volume on CT scanning and tests the effect of haematoma shape and regularity on these measures. The major findings are as follows: (1) Agreement within and between observers was excellent for measures of haemorrhage volume (ABC/2, modified ABC/2, SAS) and other haemorrhage parameters (visual categorisation by maximum length, description of shape, density and IVH); (2) haemorrhage volumes measured by modified ABC/2 were significantly lower than those measured by ABC/2 and SAS planimetry; the discrepancy increased as the haematoma size became larger; (3) larger haemorrhages were significantly more irregular in shape; (4) agreement was good between visual size categorisation and measured computerised

Figure 3 The Bland–Altman plot for assessment of variation in estimating haematoma volume (n=548) using ABC/2 and semiautomatic segmentation (n=548), r^2=0.03, p<0.0001. The continuous and dotted lines represent the regression lines. The slope of the best-fit regression line was 0.02 (p<0.0001).
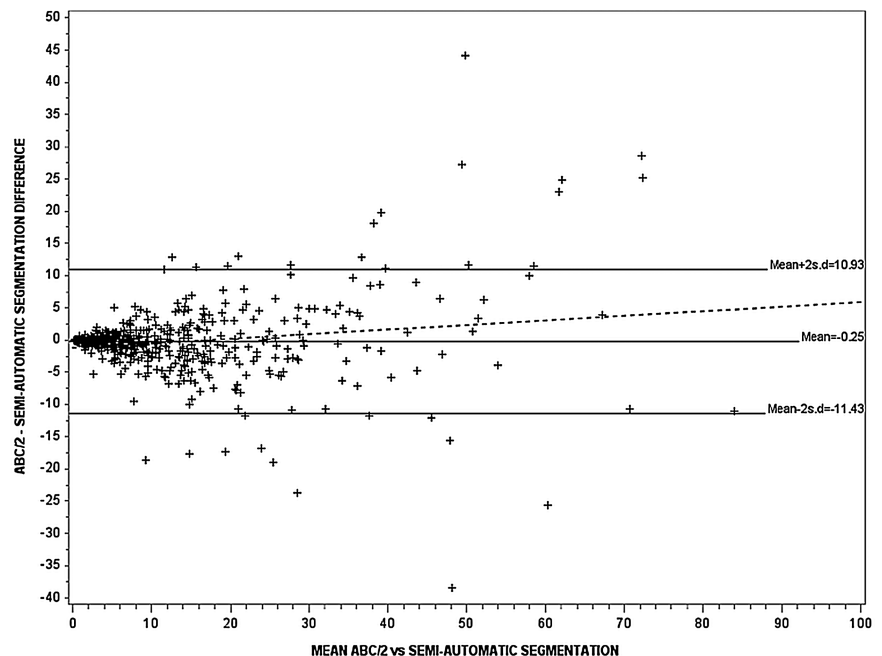
Table 3 Intraobserver comparison for two observers of haemorrhage size (n=47, 34 scans, respectively) on baseline CT scans from patients with ICH

			Difference (p)	ICC			Difference (p)	ICC
Observer	1				2			
Measurement	1	2			1	2		
Modified ABC/2 (cm³)	8.67 (8.34)	9.10 (8.34)	0.43 (0.80)	0.97	16.88 (14.60)	16.90 (14.63)	0.016 (0.97)	0.95
ABC/2 (cm³)	10.42 (10.28)	10.98 (10.79)	−0.55	0.97	20.97 (19.32)	20.15 (17.98)	0.82 (0.39)	0.96
SAS (cm³)	11.11 (10.38)	11.69 (10.99)	0.58 (0.80)	0.98	19.41 (15.32)	19.42 (15.51)	−0.11 (0.87)	0.99
Diameter (A) (cm)	3.41 (1.28)	3.37 (1.31)	0.04 (0.89)	0.98	3.89 (1.48)	4.01 (1.45)	−0.12 (0.43)	0.96

Data are mean (SD), difference in means, and ICC.
ICC, intraclass correlation coefficient; ICH, intracerebral haemorrhage; SAS, semiautomatic segmentation volume.

measurement of maximum haemorrhage diameter; (5) attempts to perform a fully automated volume measurement method failed in more than 95% of patients, as the software was unable to work with scans in which slice thicknesses varied.

The sources of error that effect variation between ABC/2 and SAS merit consideration. First, the largest diameter of haemorrhage was measured in the axial plane but this may not be the largest ICH diameter, which may also have contributed to differences between 'A' and the visual size categorisation. Sucu et al[31] suggested using the maximum length and width not necessarily on the same slice and found better correlation between ABC/2 and SAS. However, this adaptation applies only to chronic subdural haematomas which differ from spontaneous haemorrhages by extending to the cranial vault and being crescentic.[31] Second, when using ABC/2, the scan plane is measured through AB/2, the formula for the area of a triangle. However, a triangle is not necessarily the most appropriate description of a haematoma. Last, standard ABC/2 approximates the haematoma volume as an ellipsoid with all three axes extending in three perpendicular directions[9]; to compensate for the varying slice thickness, 'C' is derived by multiplying the number of slices of which the haematoma is seen by the slice thickness in centimetres. For the modified ABC/2, 'C' does not include slices if the area of the haemorrhage is less than 25% of the largest area,[10] an approach that has no theoretical justification. Hence, this method does not estimate all three Cartesian coordinates of the ellipsoid. Since the modified ABC/2 method underestimated the ICH volume (by 2–4 cm³), and small differences in volume equate to the variation in outcome,[32] the modified ABC/2 method cannot be recommended. Further, modern CT scanners provide thin slices and the ability to directly measure 'C', thereby eliminating the need for approximating slice areas.[33][34]

Although computer-assisted methods are considered the gold standard for volume measure,[1][11][12][35] the SAS method was time consuming and at times technically challenging. For

instance, it was difficult to set segmentation parameters to identify a haemorrhage that was adjacent to the bone; and the threshold may miss an area of haemorrhage. Additionally, the threshold method will not account for oedema associated with the ICH, which increases the mass effect; thus, the space-occupying effect of the ICH may be much larger than its measured hyperattenuated area. As segmentation is semiautomatic, it is feasible to adjust manually for errors if they are apparent, although any adjustment may be subjective. All images require visual checking and manual correction; failure to do this would result in erroneous measurements. The problem is compounded by more irregular spontaneous haemorrhages, for example, due to amyloid, in anticoagulant-associated bleeds and in traumatic haematomas, where typically lesions are irregular, of varied attenuation and often next to the bone.

In this series of patients with spontaneous ICH, a high proportion (64%) of irregular-shaped haemorrhages were found. The effectiveness of the ABC/2 method has been validated in regular, oval-shaped haemorrhages but researchers have doubted its accuracy in complex irregular haemorrhages (eg, as seen with warfarin or in large or amyloid-related haemorrhage) and those with intraventricular extension.[12][35] It has been suggested that adjusting the denominator from 2 to 3 in ABC/2 in irregularly shaped haematomas may produce more accurate measures, but this concept is yet to be supported.[12][14] Combining the present and published findings,[29] it may be postulated that as the haemorrhage became more irregular, the surface area and volume changes more than the largest diameter. Hence, the area of the largest haemorrhage slice may be more representative of the total haemorrhage size than its diameter. As a result, the true

Table 4 Interobserver comparison for two (observer 1 and 2; n=193 scans) of haemorrhage volume (cm³) on CT scans from patients with intracerebral haemorrhage (ICH)

	Observer			
	1	2	Difference (Δ)	ICC
ABC/2	10.58 (10.20)	12.05 (12.40)	−1.40	0.88
Modified ABC/2	8.42 (8.53)	9.70 (10.05)	−1.3	0.91
SAS	12.02 (12.05)	11.08 (11.38)	0.95	0.93

Data are mean (SD), difference (Δ) in haemorrhage volumes and ICC.
ICC, intraclass correlation coefficient; SAS, semiautomatic segmentation.

Table 5 Comparison of volumes (cm³) using ABC/2, modified ABC/2 and semiautomated segmentation (SAS) (n=548) by haematoma shape

	Haematoma volume		Difference (Δ)	p value
Regular (1,2) n=198	ABC/2 8.06 (10.43)	Modified ABC/2 5.56 (6.31)	2.51	<0.0001
	ABC/2 8.06 (10.43)	SAS 7.77 (8.75)	0.29	0.20
	SAS 7.77 (8.75)	Modified ABC/2 5.56 (6.31)	2.21	<0.0001
Irregular (3,5) n=350	ABC/2 15.48 (18.37)	Modified ABC/2 10.66 (10.07)	4.81	<0.0001
	ABC/2 15.48 (18.37)	SAS 15.62 (14.22)	−0.14	0.82
	SAS 15.62 (14.22)	Modified ABC/2 10.66 (10.07)	4.96	<0.0001

Data are mean (SD), difference (Δ) in volume and p value.

Figure 4 Box plots of intracerebral haemorrhage (ICH) volumes (n=548) by visually estimated size and corresponding volume measured by ABC/2, modified ABC/2 and semiautomatic segmentation (SAS). The difference between modified ABC/2 versus standard ABC/2 and modified ABC/2 versus SAS increases as the size category increases and is present in all four size categories (<3 (p<0.0001), 3–5 (p<0.001), 5–8 (p<0.0001) and >8 (p<0.0001)). There was no significant difference between standard ABC/2 and SAS in all four size categories.
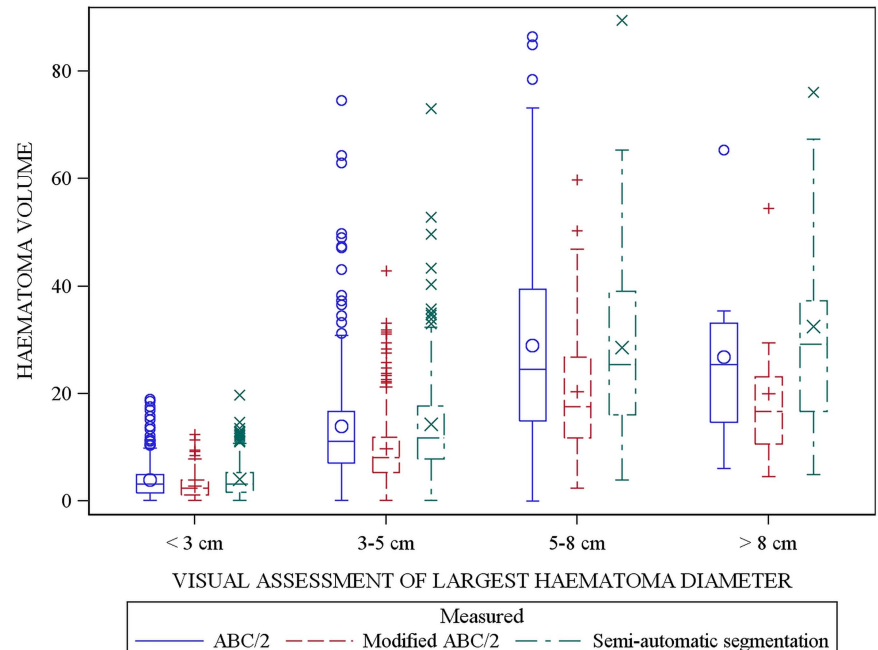
volume of an irregular haemorrhage may be better estimated using SAS. This finding is important since irregular shape haemorrhages are more likely to expand and affect morbidity and mortality.[15]

Our study is novel in design as it compares two commonly used methods ABC/2 and the modified formula (25%/75% distinction) with computer-assisted SAS and pragmatic visual scoring. Moreover, haemorrhage shape was incorporated in the analysis so that the source of error within each individual measurement method was assessed further. Our results show good compatibility between ABC/2 and computer-assisted SAS, irrespective of haematoma shape and poor approximation using modified ABC/2.

There are several potential explanations for variation in inter-observer performance. Thinner slices may produce more accurate volume measurement,[33 34] although the scans were identical for each observer. Alternatively, each additional slice to measure

increases the potential for errors. Furthermore, observers may have chosen slightly different window settings. Additionally, although the observers were all trained in using OsiriX, accuracy in measurement does depend on operator experience.[33 36]

When comparing haemorrhage size on the basis of its longest measured diameter ('A' in ABC/2) with an adjudicated visual categorisation, there was 'strong' agreement, within the adjudicators and between the observers and adjudicators, with ICC 0.72. Hence, it is unsurprising that the visual category versus the largest measured axial diameter has a slightly lower agreement than some of the other measured values. Note that the visual categorisation assigns an ordinal value based on the largest size in any plane, not just the axial plane, and thus is not the same as a linear measure of largest axial diameter.

Our study also shows that qualitative descriptors of haemorrhage characteristics of haemorrhage shape, density, presence of

Figure 5 Box plots of intracerebral haemorrhage (ICH) volumes (n=548) by visually assessed shape and corresponding volume assessed by ABC/2, modified ABC/2 and semiautomatic segmentation (SAS) show: (1) larger ICHs are irregular in shape (p<0.0001); (2) the mean difference between modified ABC/2 versus standard ABC/2 and modified ABC/2 versus SAS increased as the haematoma shape became more irregular (p<0.0001). No significant difference was observed between standard ABC/2 and SAS.
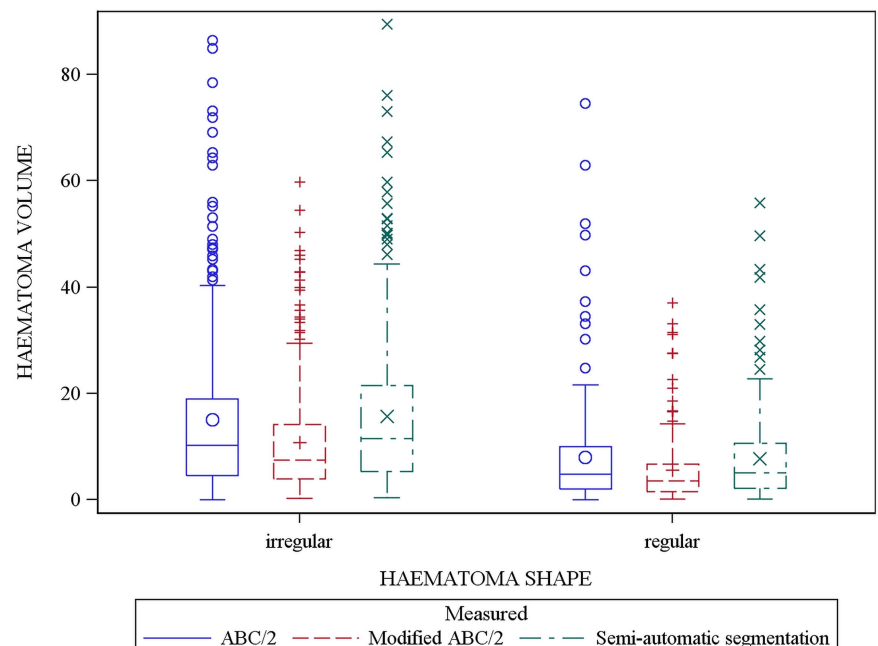
IVH using the Graeb and modified Graeb scores can be reliably measured and are reproducible.

Of note, when measuring density heterogeneity, we used only the coefficient of variation (cv), whereas Barras used four additional measures (although these were apparently inferior to cv).[23] Calculation of shape index using parameters obtained from the windowed haemorrhage had a low ICC of 0.53. When this result is held in context with similar computed volumes, it could be interpreted that the observer chose different threshold settings which explain the source of variability.

This study has several strengths including the use of multiple assessors, multiple methods for assessing haemorrhage characteristics, and a large data set. To minimise measurement error, observers were trained to use the OsiriX software for measuring brain CT scan parameters and in recognition of haemorrhages. Similarly, CT adjudicators were experienced neuroradiologists. Multiple observers were used to allow measurement of intraobserver and interobserver variation. The data set involved patients from five continents and so the findings have excellent external validity.

Nevertheless, the study has three significant limitations. First, the data come from a randomised controlled trial in acute stroke. Trial exclusion criteria can limit the type of patients (and therefore variation in haemorrhage) so that the data set studied here did not include patients with normal/low blood pressure, GCS<8, or without motor signs. As a result, patients with very large haematoma were under-represented. Second, the time from stroke onset to neuroimaging was relatively long (median 4.2 h), reflecting that the ENOS protocol allowed enrolment up to 48 h after stroke. Hence, ICH on later scans may already have developed some peri-ICH oedema. Finally, just over 25% of the patients had measurable IVH, and therefore the results do not represent a population where IVH is more frequent. The failure to get a fully automated volume measure in 95% of scans may be less of a problem with other software that can handle scans of variable slice thickness.

In conclusion, the modified ABC/2 formula significantly underestimates haemorrhage volume when compared to standard ABC/2 and computer assisted semiautomated segmentation (SAS) the difference increased as haematoma volume became larger and more irregular in shape. Most relevant for clinicians, our results show that the standard ABC/2 method offers more accurate quantification of ICH volume by the bedside. Although SAS is consistent and reliable, the method is slow, reliant on manual correction, advanced software and likely to cause delay in clinical decision-making often in emergency situations. Further research is required to create a faster computerised volumetric model for accurate and reliable measurement of haematomas in the clinical setting. Visual assessment using an ordered categorical scale has the strength of providing an 'instant' marker of haemorrhage size in the acute situation and can be applied in the absence of measurement tools and historical cut film data, and where very large study sizes or very moved images preclude computer based ABC/2 or SAS measurements. This study importantly shows that haemorrhage characteristics can be reliably measured as accuracy is pivotal to clinical trials in which ICH volume change may be a surrogate end point.

Author affiliations
[1]Stroke, Division of Clinical Neuroscience, University of Nottingham, Nottingham, UK
[2]Radiological Sciences Research Group, Division of Clinical Neuroscience, University of Nottingham, Nottingham, UK
[3]Department of Neuroradiology, Nottingham University Hospitals, Queen's Medical Centre, Nottingham, UK
[4]Stroke Centre, Ospedale Sacro Cuore Via Sempreboni, Verona, Italy
[5]Division of Neuroimaging Sciences, Centre for Clinical Brain Sciences, Western General Hospital, Edinburgh, UK

REFERENCES

1. Broderick JP, Brott TG, Duldner JE, et al. Volume of intracerebral hemorrhage. A powerful and easy-to-use predictor of 30-day mortality. *Stroke* 1993;24:987–93.
2. Wada R, Aviv RI, Fox AJ, et al. CT angiography "spot sign" predicts hematoma expansion in acute intracerebral hemorrhage. *Stroke* 2007;38:1257–62.
3. Almandoz JED, Yoo AJ, Stone MJ, et al. Systematic characterization of the computed tomography angiography spot sign in primary intracerebral hemorrhage identifies patients at highest risk for hematoma expansion the spot sign score. *Stroke* 2009;40:2994–3000.
4. Mendelow AD, Gregson BA, Fernandes HM, et al. Early surgery versus initial conservative treatment in patients with spontaneous supratentorial intracerebral haematomas in the International Surgical Trial in Intracerebral Haemorrhage (STICH): a randomised trial. *Lancet* 2005;365:387–97.
5. Mayer SA, Brun NC, Broderick J, et al. Safety and feasibility of recombinant factor VIIa for acute intracerebral hemorrhage. *Stroke* 2005;36:74–9.
6. Lees KR, Bath PMW, Schellinger PD, et al. Contemporary outcome measures in acute stroke research: choice of primary outcome measure. *Stroke* 2012;43:1163–70.
7. Anderson CS, Huang Y, Wang JG, et al. Intensive blood pressure reduction in acute cerebral haemorrhage trial (INTERACT): a randomised pilot trial. *Lancet Neurol* 2008;7:391–9.
8. Mayer SA, Brun NC, Begtrup K, et al. Efficacy and safety of recombinant activated factor VII for acute intracerebral hemorrhage. *N Engl J Med* 2008;358:2127–37.
9. Kwak R, Kadoya S, Suzuki T. Factors affecting the prognosis in thalamic hemorrhage. *Stroke* 1983;14:493–500.
10. Kothari RU, Brott T, Broderick JP, et al. The ABCs of measuring intracerebral hemorrhage volumes. *Stroke* 1996;27:1304–5.
11. Gebel JM, Sila CA, Sloan MA, et al. Comparison of the ABC/2 estimation technique to computer-assisted volumetric analysis of intraparenchymal and subdural hematomas complicating the GUSTO-1 trial. *Stroke* 1998;29:1799–801.
12. Huttner HB, Steiner T, Hartmann M, et al. Comparison of ABC/2 estimation technique to computer-assisted planimetric analysis in warfarin-related intracerebral parenchymal hemorrhage. *Stroke* 2006;37:404–8.
13. Freeman WD, Barrett KM, Bestic JM, et al. Computer-assisted volumetric analysis compared with ABC/2 method for assessing warfarin-related intracranial hemorrhage volumes. *Neurocrit Care* 2008;9:307–12.
14. Sheth KN, Cushing TA, Wendell L, et al. Comparison of hematoma shape and volume estimates in warfarin versus non-warfarin-related intracerebral hemorrhage. *Neurocrit Care* 2010;12:30–4.
15. Barras CD, Tress BM, Christensen S, et al. Density and shape as CT predictors of intracerebral hemorrhage growth. *Stroke* 2009;40:1325–31.
16. Graeb DA, Robertson WD, Lapointe JS, et al. Computed tomographic diagnosis of intraventricular hemorrhage- etiology and prognosis. *Radiology* 1982;143:91–6.
17. Morgan T, Dawson J, Spengler D, et al. The modified Graeb score an enhanced tool for intraventricular hemorrhage measurement and prediction of functional outcome. *Stroke* 2013;44:635–41.
18. The ENOS Trial Investigators. Glyceryl trinitrate vs. control, and continuing vs. stopping temporarily prior antihypertensive therapy, in acute stroke: rationale and

design of the Efficacy of Nitric Oxide in Stroke (ENOS) trial (ISRCTN99414122). *Int J Stroke* 2006;1:245–9.

19 Wardlaw JM, Sellar R. A simple practical classification of cerebral infarcts on CT and its interobserver reliability. *Am J Neuroradiol* 1994;15:1933–9.

20 Sandercock P, Lindley R, Wardlaw J, *et al*. Third international stroke trial (ist-3) of thrombolysis for acute ischaemic stroke. *Trials* 2008;9:37.

21 Rosset A, Spadola L, Ratib O. OsiriX: an open-source software for navigating in multidimensional DICOM images. *J Digit Imaging* 2004;17:205–16.

22 Pabst W, Gregerova E. *Characterisation of particle and particle systems*. Prague: ICT, 2007;27–29.

23 Barras CD, Tress BM, Christensen S, *et al*. Quantitative CT densitometry for predicting intracerebral hemorrhage growth. *Am J Neuroradiol* 2013;34:1138–44.

24 Shrout P, Fleiss J. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull* 1979;86:420–8.

25 ENOS Investigators. Baseline characteristics of the 4011 patients recruited into the "efficacy of Nitric Oxide in Stroke (ENOS) trial. *Int J Stroke* 2014;9:711–20.

26 Bath PMW, Woodhouse L, Scutt P, *et al*. Management of high blood pressure in acute stroke: efficacy of Nitric Oxide in Stroke (ENOS), a partial-factorial randomised controlled trial. *Lancet* 2014. Published online 22 Oct 2014 http://dx.doi.org/10.1016/ S0140-6736(14)61121-1

27 Wang CW, Juan CJ, Liu YJ, *et al*. Volume-dependent overestimation of spontaneous intracerebral hematoma volume by the ABC/2 formula. *Acta Radiol* 2009;50:306–11.

28 Maeda AK, Aguiar LZ, Martins C, *et al*. Hematoma volumes of spontaneous intracerebral hemorrhage: the ellipse (ABC/2) method yielded volumes smaller than those measured using the planimetric method. *Arquivod de Neuro-Psiquiatria* 2013;71:540–4.

29 Yang W, Feng Y, Zhang Y, *et al*. Volume quantification of acute infratentorial hemorrhage with computed tomography: validation of the formula 1/2ABC and 2/3SH. *PLoS ONE* 2013;8:e62286.

30 Zhao KJ, Liu Y, Zhang RY, *et al*. A precise, simple, convenient and new method for estimation of intracranial hematoma volume-the formula 2/3sh. *Neurol Res* 2009;31:1031–6.

31 Sucu HK, Gokmen M, Gelal F. The value of XYZ/2 technique compared with computer-assisted volumetric analysis to estimate the volume of chronic subdural hematoma. *Stroke* 2005;36:998–1000.

32 Brott T, Broderick J, Kothari R, *et al*. Early hemorrhage growth in patients with intracerebral hemorrhage. *Stroke* 1997;28:1–5.

33 Divani AA, Majidi S, Luo X, *et al*. The ABCs of accurate volumetric measurement of cerebral hematoma. *Stroke* 2011;42:1569–74.

34 Prionas ND, Ray S, Boone JM. Volume assessment accuracy in computed tomography: a phantom study. *J Appl Clin Med Phys* 2010;11:3037.

35 Kosior J, Idris S, Dowlatshahi D, *et al*. Quantomo: validation of a computer-assisted methodology for the volumetric analysis of intracerebral haemorrhage. *Int J Stroke* 2011;6:302–5.

36 Hussein HM, Tariq NA, Palesch YY, *et al*. Reliability of hematoma volume measurement at local sites in a Multicenter Acute Intracerebral Hemorrhage Clinical Trial. *Stroke* 2013;44:237–9.

Table I. Inter-adjudicator comparison of visually estimated longest haemorrhage diameter across four size categories (<3 cm, 3-5 cm, 5-8 cm, >8 cm) in 47 patients; intra-class correlation (ICC)=0.82, p<0.001.

Observer 1	Observer 2				Total
	<3	3-5	5-8	>8	
<3	36	3	0	0	39
3-5	1	3	1	0	5
5-8	0	0	3	0	3
>8	0	0	0	0	0
Total	37	6	4	0	47

1

Table II. Comparison of visually estimated maximum haemorrhage category versus computerised measurement (n=548). Data are number (%), mean (standard deviation) or median[interquartile range]; intra-class correlation (ICC) = 0.71 (p<0.001). Modal diameter 0.59 cm.

Size category (cm)	Axial diameter 'A', minimum (cm)	Axial diameter 'A', mean (SD) (cm)	Axial diameter 'A', maximum (cm)
<3	0.66	2.22 (0.74)	5.10
3 to <5	0.59	3.76 (0.92)	7.29
5 to 8	2.12	5.07 (1.09)	7.53
>8	3.46	5.44 (1.27)	7.52
3 to <5 [<3]	0.59	3.38 (1.39)	7.53

Table III. Intra-observer comparison of haemorrhage shape, Graeb and modified-Graeb scores, and intraventricular haemorrhage. Data are mean (standard deviation) and intra-class correlation coefficient (ICC).

	Measure 1	Measure 2	ICC
Without IVH, n=82			
Shape (/5)	3.68 (1.18)	3.74 (1.14)	0.88
Shape index	1.26 (0.37)	1.24 (0.34)	0.53
Density (/5)	2.94 (1.07)	2.87 (0.99)	0.86
Density index	0.22 (0.09)	0.21 (0.08)	0.86
With IVH, n=49			
Graeb score (/12)	4.10 (2.47)	4.04 (2.43)	0.99
Modified Graeb score (/32)	6.94 (4.94)	7.04 (4.96)	0.99
Volume (by SAS, cc^3)	5.36 (8.33)	5.69 (9.32)	0.98

SAS: semi-automatic segmentation

Table IV. Inter-observer comparison of visually estimated haemorrhage shape and density. Data are mean (standard deviation) and intra-class correlation coefficient (ICC).

	Observer 1	Observer 2	ICC
N=48			
Shape (/5)	2.70 (1.23)	2.96 (1.28)	0.75, p<0.001
Density (/5)	1.62 (0.60)	1.92 (0.70)	0.73, p<0.001

4

Figure 1. Graph for assessment of variation between two observers in estimating haematoma volume using ABC/2 and semi-automatic segmentation (n=193), r_2 =0.93, p<0.0001. The continuous and the dotted lines represent the regression lines. The slope of the best-fit regression line was 1.05 (p<0.0001).
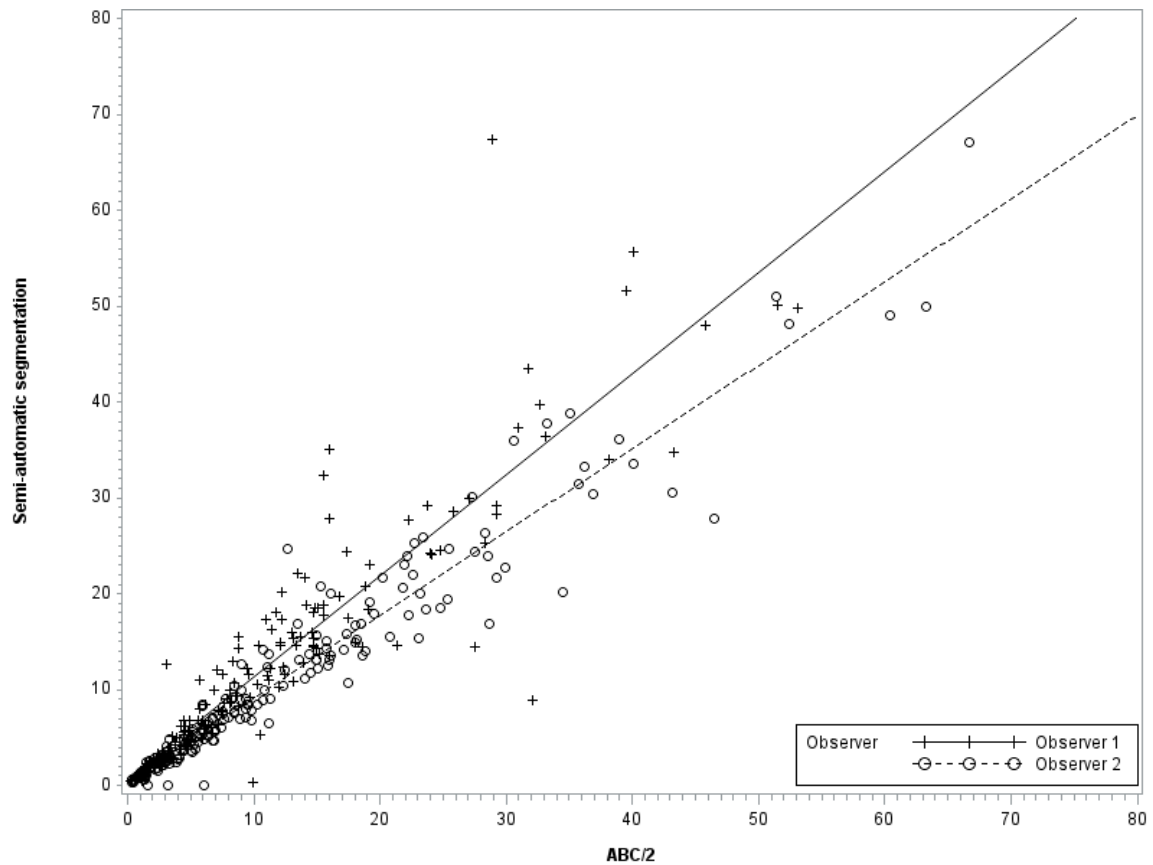
Figure 2. Graph for assessment of variation in estimating haematoma volume using ABC/2 and modified ABC/2 (n=193), r_2 =0.94, p<0.0001. The continuous and the dotted lines represent the regression lines. The slope of the best-fit regression line was 0.78 (p<0.0001).
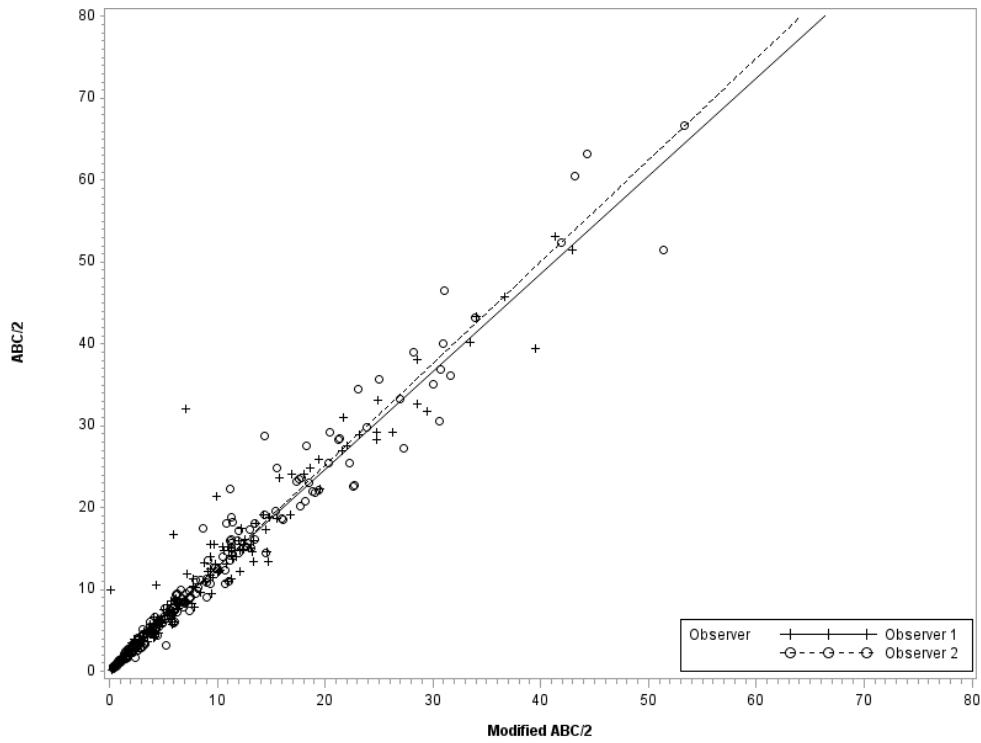
Figure 3. Graph for assessment of variation in estimating haematoma volume using modified ABC/2 and semi-automatic segmentation, $r_2 = 0.86$, $p < 0.0001$. The continuous and the dotted lines represent the regression lines. The slope of the bestfit regression line was 1.31 ($p < 0.0001$).
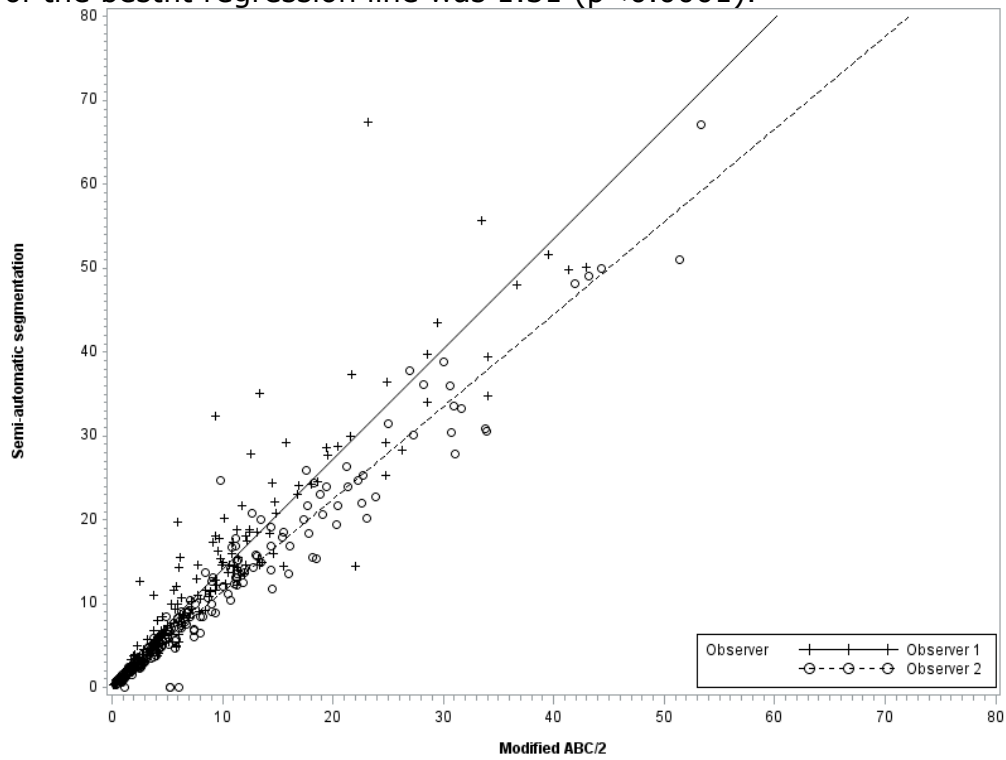
Figure 4. Forest plot of studies comparing ICH volume measurement by ABC/2 and computer- assisted SAS in spontaneous and anticoagulant related ICH. The squares indicate the point estimates and the width of the horizontal lines is the 95% confidence interval of the estimate. The diamond at the bottom represents the point estimate as well as the 95% confidence intervals of the overall effect within the categories.

Study or Subgroup	ABC/2 Mean	SD	Total	Planimetry Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
1.1.1 Spontaneous ICH								
Zhao 2009	45.98	39.97	186	43.16	37.55	186	10.4%	2.82 [-5.06, 10.70]
Yang 2013	3.67	7.04	147	4.83	7.59	147	29.8%	1.16 [2.83, 0.51]
Wang 2009	43.6	0	40	33.8	0	40		Not estimable
Gebel 1998	68.7	0	244	63.3	0	244		Not estimable
ENOS 2014	12.77	16.32	548	12.76	13.06	548	29.5%	0.01 [-1.74, 1.76]
Subtotal (95% CI)			1165			1165	69.7%	-0.52 [-1.72, 0.67]
Heterogeneity: Tau² = 0.00; Chi² = 1.60, df = 2 (P = 0.45); I² = 0%								
Test for overall effect: Z = 0.86 (P = 0.39)								
1.1.2 Anticoagulation ICH								
Huttner 2006	40.83	3.9	50	36.6	3.5	45	30.3%	4.23 [2.74, 5.72]
Subtotal (95% CI)			50			45	30.3%	4.23 [2.74, 5.72]
Heterogeneity: Not applicable								
Test for overall effect: Z = 5.57 (P < 0.00001)								
1.1.3 Spontaneous ICH vs Anticoagulation ICH								
Subtotal (95% CI)			0			0		Not estimable
Heterogeneity: Not applicable								
Test for overall effect: Not applicable								
Total (95% CI)			1215			1210	100.0%	1.23 [-1.85, 4.31]
Heterogeneity: Tau² = 7.57; Chi² = 25.42, df = 3 (P < 0.0001); I² = 88%								
Test for overall effect: Z = 0.78 (P = 0.43)								
Test for subgroup differences: Chi² = 23.81, df = 1 (P < 0.00001), I² = 95.8%								

Figure 5. Forest plot of studies comparing variation between ABC/2 and computerassisted SAS measurements by haematoma shape. The squares indicate the point estimates and the width of the horizontal lines is the 95% confidence interval of the estimate. The diamond at the bottom represents the point estimate as well as the 95% confidence intervals of the overall effect within the categories.

Study or Subgroup	ABC/2 (original)			Planimetry			Weight	Mean Difference IV, Random, 95% CI	Mean Difference IV, Random, 95% CI
	Mean	SD	Total	Mean	SD	Total			
1.2.1 Regular									
ENOS 2014	8.06	10.43	198	7.77	8.75	198	15.6%	0.29 [−1.61, 2.19]	
Huttner 2006	42.5	6.6	28	39.9	4.8	28	12.5%	2.60 [−0.42, 5.62]	
Yang 2013	2.37	3.25	51	3.03	4.57	51	16.5%	−0.66 [−2.20, 0.88]	
Subtotal (95% CI)			277			277	44.6%	0.34 [−1.25, 1.92]	
Heterogeneity: Tau² = 0.87; Chi² = 3.60, df = 2 (P = 0.17); I² = 44%									
Test for overall effect: Z = 0.42 (P = 0.68)									
1.2.2 Irregular									
ENOS 2014	15.48	18.37	350	15.62	14.22	350	14.1%	−0.14 [−2.57, 2.29]	
Huttner 2006	38.9	8.2	28	34.3	6.6	28	10.3%	4.60 [0.70, 8.50]	
Yang 2013	5.47	9.06	96	6.14	8.78	96	13.9%	−0.67 [−3.19, 1.85]	
Subtotal (95% CI)			474			474	38.3%	0.88 [−1.84, 3.59]	
Heterogeneity: Tau² = 3.55; Chi² = 5.33, df = 2 (P = 0.07); I² = 63%									
Test for overall effect: Z = 0.63 (P = 0.53)									
1.2.3 Seperated									
Huttner 2006	43.3	9.7	27	32.9	7.9	27	8.5%	10.40 [5.68, 15.12]	
Yang 2013	9.79	8.33	21	9.75	7.07	21	8.6%	0.04 [−4.63, 4.71]	
Subtotal (95% CI)			48			48	17.1%	5.21 [−4.94, 15.37]	
Heterogeneity: Tau² = 47.92; Chi² = 9.35, df = 1 (P = 0.002); I² = 89%									
Test for overall effect: Z = 1.01 (P = 0.31)									
Total (95% CI)			799			799	100.0%	1.51 [−0.36, 3.37]	
Heterogeneity: Tau² = 4.86; Chi² = 26.35, df = 7 (P = 0.0004); I² = 73%									
Test for overall effect: Z = 1.58 (P = 0.11)									
Test for subgroup differences: Chi² = 0.94, df = 2 (P = 0.63), I² = 0%									

−50 −25 0 25 50
Planimetry larger ABC/2 larger

Figure 6. Forest plot of studies comparing ICH volume by modified ABC/2 and computer-assisted SAS in spontaneous and anticoagulation related ICH. The squares indicate the point estimates and the width of the horizontal lines is the 95% confidence interval of the estimate. The diamond at the bottom represents the point estimate as well as the 95% confidence intervals of the overall effect within the categories.

3.1 ABC/2 (modified) vs planimetry

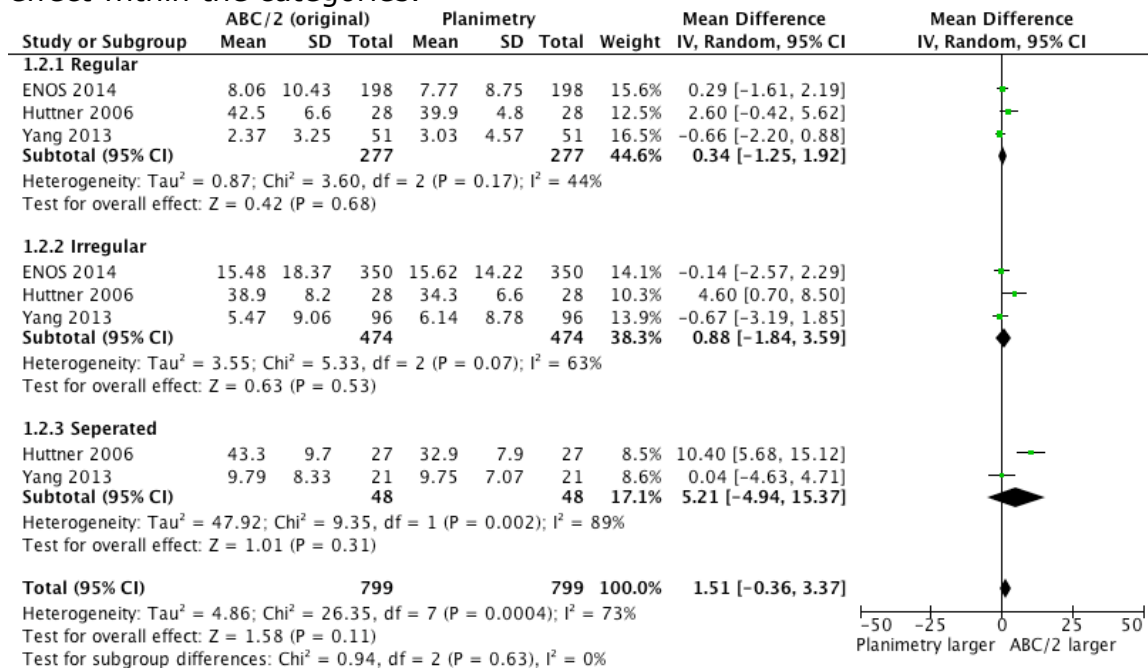| Study or Subgroup	ABC/2 (modified)			Planimetry			Weight	Mean Difference IV, Random, 95% CI	Mean Difference IV, Random, 95% CI
	Mean	SD	Total	Mean	SD	Total			
3.1.1 Spontaneous ICH									
ENOS 2014	8.8	9.21	548	12.76	13.06	548	33.3%	−3.96 [−5.30, −2.62]	
Kosior 2011	48	41.6	30	33.1	26.1	30	5.1%	14.90 [2.67, 32.47]	
Kothari 1996	27.5	2.9	118	26	2.6	118	34.1%	1.50 [0.80, 2.20]	
Maeda 2013	12.8	11.37	20	15.04	12.46	20	17.0%	−2.24 [−9.63, 5.15]	
Sheth 2010	30	37	50	35	38	50	6.8%	−5.00 [−19.70, 9.70]	
Subtotal (95% CI)			766			766	96.3%	−0.80 [−5.19, 3.59]	
Heterogeneity: Tau² = 14.09; Chi² = 53.74, df = 4 (P < 0.00001); I² = 93%									
Test for overall effect: Z = 0.36 (P = 0.72)									
3.1.2 Anticoagulation ICH									
Freeman 2008	72.2	0	7	75.4	0	7		Not estimable	
Sheth 2010	44	51	50	53	56	50	3.7%	−9.00 [−29.99, 11.99]	
Subtotal (95% CI)			57			57	3.7%	−9.00 [−29.99, 11.99]	
Heterogeneity: Not applicable									
Test for overall effect: Z = 0.84 (P = 0.40)									
Total (95% CI)			823			823	100.0%	−1.11 [−5.40, 3.19]	
Heterogeneity: Tau² = 13.97, Chi² = 54.50, df = 5 (P < 0.00001); I² = 91%									
Test for overall effect: Z = 0.51 (P = 0.61)									
Test for subgroup differences: Chi² = 0.56, df = 1 (P = 0.45), I² = 0%									

−100 −50 0 50
Modified ABC/2 (larger) Planimetry (l...

Figure 7. Forest plot of studies comparing variation between modified ABC/2 and computer-assisted SAS by haematoma shape. The squares indicate the point estimates and the width of the horizontal lines is the 95% confidence interval of the estimate. The diamond at the bottom represents the point estimate as well as the 95% confidence intervals of the overall effect within the categories.
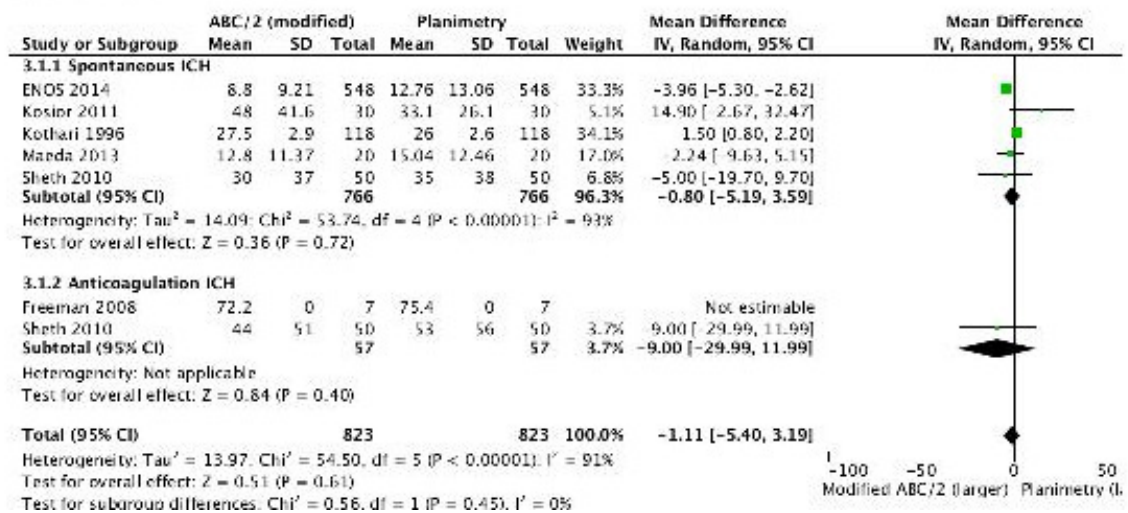
Study or Subgroup	ABC/2 (modified)			Planimetry			Weight	Mean Difference IV, Random, 95% CI	Mean Difference IV, Random, 95% CI
	Mean	SD	Total	Mean	SD	Total			
3.2.1 Regular									
ENOS 2014	5.56	6.31	198	7.77	8.75	198	51.8%	−2.21 [−3.71, −0.71]	
Freeman 2008	22	0	3	28.7	0	3		Not estimable	
Subtotal (95% CI)			201			201	51.8%	**−2.21 [−3.71, −0.71]**	
Heterogeneity: Not applicable									
Test for overall effect: Z = 2.88 (P = 0.004)									
3.2.2 Irregular									
ENOS 2014	10.66	10.07	350	15.62	14.22	350	48.2%	−4.96 [−6.79, −3.13]	
Freeman 2008	110	0	4	110.5	0	4		Not estimable	
Subtotal (95% CI)			354			354	48.2%	**−4.96 [−6.79, −3.13]**	
Heterogeneity: Not applicable									
Test for overall effect: Z = 5.33 (P < 0.00001)									
Total (95% CI)			555			555	100.0%	**−3.53 [−6.23, −0.84]**	

Heterogeneity: Tau^2 = 3.05; Chi^2 = 5.20, df = 1 (P = 0.02); I^2 = 81%
Test for overall effect: Z = 2.57 (P = 0.01)
Test for subgroup differences: Chi^2 = 5.20, df = 1 (P = 0.02), I^2 = 80.8%

−50   −25   0   25   50
Planimetry larger   Modified ABC/2 larger