Software output from semi-automated planimetry can underestimate intracerebral haemorrhage and peri-haematoma oedema volumes by up to 41%

Teddy Y Wu¹, Oluwaseun Sobowale², Robert Hurford², Gagan Sharma³, Søren Christensen⁴, Nawaf Yassi¹, Turgut Tatlisumak⁵,⁶, Patricia M Desmond³, Bruce CV Campbell¹, Stephen M Davis¹, Adrian R Parry-Jones² and Atte Meretoja¹,⁷*

¹Department of Medicine at the Royal Melbourne Hospital, University of Melbourne, Parkville, Victoria, Australia
²Centre for Vascular and Stroke Research, Manchester Academic Health Sciences Centre, Salford Royal Hospital Trust, University of Manchester, Manchester, UK
³Department of Radiology, the Royal Melbourne Hospital, University of Melbourne, Parkville, Victoria, Australia
⁴Stanford Stroke Center, Stanford University, Palo Alto, California, USA
⁵Institute of Neuroscience and Physiology, Sahlgrenska Academy at the University of Gothenburg, Gothenburg, Sweden
⁶Department of Neurology, Sahlgrenska University Hospital, Gothenburg, Sweden
⁷Department of Neurology, Helsinki University Hospital, PK admin, Haartmaninkatu 4E, P.O. Box 220, 00029 HUS, Helsinki, Finland
E-mail: atte.meretoja@unimelb.edu.au

*Corresponding author:
A/Prof Atte Meretoja

Fax:+61393428443
Abstract:

Introduction

Haematoma and oedema size determine outcome after intracerebral haemorrhage (ICH), with each added 10% volume increasing mortality by 5%. We assessed the reliability of semi-automated CT planimetry using Analyze and Osirix softwares.

Methods

We randomly selected 100 scans from 1329 ICH patients from two centres. We used Hounsfield Unit thresholds of 5-33 for oedema and 44-100 for ICH. Three raters segmented all scans using both softwares, and 20 scans repeated for intrarater reliability and segmentation timing. Volume reported by Analyze and Osirix were compared to volume estimates calculated using best practice method taking effective individual slice thickness, i.e. voxel depth, into account.

Results

There was excellent overall interrater, intrarater and intersoftware reliability, all ICCs >0.918. Analyze and Osirix produced similar haematoma (mean difference Analyze – Osirix: 1.5mL(6%) +/- 5.2mL, p=<0.001), and oedema volumes (-0.6mL(-3%) +/- 12.6mL, p=0.377). Compared to a best practice approach to volume calculation, the automated haematoma volume output was 2.6mL(-11%) too small with Analyze and 4.0mL(-18%) too small with Osirix, while the oedema volumes were 2.5mL(-12%) and 5.5mL(-25%) too small correspondingly. In scans with variable slice thickness the volume underestimations were larger, -29%/-36% for ICH and -29%/-41% for oedema. Mean segmentation time was 6:53 +/- 4:02 minutes with Analyze and 9:06 +/- 5:24 minutes with Osirix, p<0.001.

Conclusion
Our results demonstrate the method used to determine voxel depth can influence the final volume output markedly. Results of clinical and collaborative studies need to be considered in the context of these methodological differences.

Keywords:
Intracerebral haemorrhage, oedema, planimetry, validation, reliability
Introduction

Intracerebral haemorrhage (ICH) is a highly fatal stroke subtype and the only proven treatments are blood pressure control and stroke unit care [1, 2]. Primary injury results from direct mechanical pressure of the haematoma in a volume dependent manner [3]. Secondary injury ensues from peri-haematomal oedema [4] with emerging evidence of oedema’s association with poor clinical outcome [5]. Effective oedema management may improve outcome and is the target of several early phase clinical trials [6-9]. Accurate volumetric assessment of both ICH and oedema is therefore vital to the analysis of clinical and collaborative studies.

The volume of tissue with oedema and ICH can be quantified on computed tomography (CT) using Hounsfield unit (HU) threshold semi-automated segmentation [10-14]. From a technical perspective, the most accurate measure of volume of a segmented region is the cumulative total volume of the voxels contained within this region. In turn the individual voxel volume is determined by the product of voxel area and voxel depth. The voxel width and height, which determine the voxel area are easily accessible in the Digital Imaging and Communications in Medicine (DICOM) header. The voxel depth that should be used for volume calculation is the distance between slice centres. This information is generally not directly available in the DICOM header but must be calculated based on other spatial information in the header. The method with which voxel depth is calculated by imaging analysis software has not been explored in detail in published studies. Several softwares are available for segmentation but no reports have directly assessed the reliability of planimetric measurements between softwares.
The aims of this study were two fold. First we assessed the reliability and time taken for segmentation using semi-automated planimetry between Analyze 12.0 (Biomedical Imaging Resource; Mayo Clinic) and Osirix 6.5 (Pixmeo; Geneva; Switzerland). Secondly we analysed the impact of different methods for voxel depth estimation on the final estimated volumes. We hypothesised that the segmented volumes and times taken for segmentation are different between Analyze and Osirix, and that the volumes are influenced by the method of voxel depth determination.

Methods

Image selection

A convenient sample of 100 cases was chosen for this study. The baseline scans of randomly selected patients from the combined database of 1329 consecutive ICH patients from Helsinki University Hospital, Finland, and Salford Royal Hospital, United Kingdom were used after institutional approval. Fifty scans were randomly chosen from each centre. The Helsinki ICH study is a retrospective analysis of consecutive ICH patients admitted to Helsinki University Central Hospital between January 2005 and March 2010 [15]. Patients from Salford Royal Hospital were ICH patients treated at the centre between January 2013 and May 2015.

Image processing and segmentation

The de-identified images were transferred in the DICOM format to a central workstation. The DICOMs were then converted into Neuroimaging Informatics Technology Initiative (NIfTI) format using a DCM to NIfTI conversion tool before loading on Analyze. The NIfTI allows the individual DICOM images to be saved and loaded from one single file. For Osirix the DICOMs were loaded directly.

Segmentation steps
The semi-automated method reported by Volbers et al was used [10]. A limit boundary was placed around the ICH and oedema complex following which a seed point was placed within the region of interest (ROI). For oedema segmentation, the lower HU limit is fixed at 5 with the upper boundary manually adjusted to a maximum of 33 HU range using visual inspection and comparison to the contralateral hemisphere for background leucoaraiosis. For ICH segmentation the HU range was kept within 44-100 HU and manual editing of ROI was allowed at the rater’s discretion.

Video demonstrations of the segmentation process on each of the softwares are available on the online supplement.

**Rater workflow**

All 100 scans were segmented using Analyze and Osirix by each rater (T.Y.W, O.S, R.H) independently and blinded to patient’s clinical details. T.Y.W is a neurologist with 6 years of experience in stroke, O.S is a neurosurgery registrar with 6 years of experience in stroke and R.H is a clinical academic fellow with 3 years of experience in stroke imaging. Twenty scans (10 from Helsinki, 10 from Salford) were chosen randomly and repeat segmentation was performed in these for intrarater assessment. The intrarater assessment was performed after a minimal interval of 7 days. The time taken to perform segmentation for each software was measured for the intrarater scans.

**ROI Volume calculation**

The volume estimates were determined by automated software output (‘Sampling Options’ in Analyze and ‘ROI manager’ in Osirix) and then by subjecting the software outputs to the “best practice” volume estimation. The best practice volume was calculated using an in-house script developed using Matlab (The MathWorks,
Inc, MA, USA) (supplementary table). To this end, the ROI object files from Analyze were used while ROI segmented on Osirix was exported in DICOM format using an in-house plugin (supplementary file). The Matlab script determined the number of voxels within the ROI, which was then multiplied by the voxel volume. The voxel volume was derived from the product of voxel width and voxel height (DICOM header 0028, 0030) and voxel depth. The voxel depth may or may not be equivalent to the slice thickness presented in the DICOM header (0018, 0050) which is not well defined in the DICOM standard and used differently between scanner manufacturers. Therefore, the in-house script uses the best practice method for determining inter-slice distance, which is to determine the normal vector to the image slice orientation (identical for all slices) using “Image Orientation Patient” (IOP, DICOM header 0020, 0037).

The location of each slice “Image Position Patient” (IPP, DICOM header 0020, 0032) is then projected onto the slice normal vector producing one value per slice expressing the location of the slice on the slice normal as a scalar value (Zpos). This calculation is not influenced by gantry tilt which can overestimate the voxel depth by approximately 5% (supplementary figure 1).

The Zpos was then used to calculate the voxel depth using the equation,

\[ \text{Voxel depth} = \frac{\text{Zpos}(n+1) - \text{Zpos}(n)}{2} + \frac{\text{Zpos}(n-1) - \text{Zpos}(n)}{2} \]

Where \( n \) is the image slice, \((n-1)\) the image slice below and \((n+1)\) the slice above. For the first and last slice of the scan we considered the thickness to be the same as the second slice and the penultimate slice respectively. This calculation approximates the slice distance not just between slices with identical distance between them, but also in the interface where slice thickness changes from for example 7.5mm to 5mm, where an overlap (figure 1a) can be present. Our in-house method adjusts for any slice
overlap and is direction insensitive (figure 1b). We will demonstrate using a case illustration the impact of varying voxel depth on volume output. We developed this approach because Z position within IPP is direction and gantry tilt sensitive and can over-estimate the volume when slice over-lap is present.

Statistical analysis

Intraclass correlation coefficients (ICC) were calculated using 2-way random effects model and used to measure interrater and intrarater reliability for ICH and oedema. An ICC was considered moderate agreement if 0.41 to 0.60, substantial agreement if 0.61 to 0.80, and excellent if 0.81 to 1.00 [16]. Bland-Altman plots were used to assess for systematic bias. Paired T test or one-way ANOVA were used to compare means where appropriate. A p value of <0.05 was considered significant. All statistics were performed using SPSS 22 (IBM, Armonk, NY).

Reporting standards

This study was reported in accordance with the Guidelines for Reporting Reliability and Agreement Studies [17].

Results

Patient characteristics

Of the 100 scans included in this study, 87 were performed within 24 hours of ictus. The median age was 69, 48% were men, baseline median National Institutes of Health Stroke Scale score was 10, median time to scan was 2.8 hours and ICH location was lobar in 36%, deep in 43% and infratentorial in 10%.

Volumetric output and agreement statistics

The mean +/- standard deviation haematoma volume segmented on Analyze was 23.4 +/- 29.6 mL and 20.9 +/- 25.6 mL for oedema. The overall interrater ICCs were 0.994 for ICH and 0.952 for oedema, the intrarater ICCs were 0.998 for ICH and 0.983 for
oedema. The mean volumes for haematoma and oedema performed on Osirix were 21.9 mL +/- 27.0 and 21.5 +/- 24.4 mL respectively. The interrater ICCs were 0.997 for ICH and 0.944 for oedema and intrarater ICCs were 0.996 for ICH and 0.918 for oedema. The intersoftware ICCs were 0.991 for ICH and 0.932 for oedema (table 1). ICH volume was 1.5 mL ((6%) +/- 5.2 mL, p=<0.001) larger in Analyze and oedema was 0.6 mL ((3%) +/- 12.5 ml p=0.377) larger in Osirix. The difference in the ICH or oedema volumes obtained between raters was not statistically significant (table 2). The Bland-Altman plots for interrater agreement (figure 2, supplementary figure 2), presented as volume difference between raters 1 - 2, raters 2 - 3 and raters 1 - 3 demonstrated a bias of 0.6 mL, -1.1 mL and -0.5 mL respectively for ICH and 0.5 mL, -0.4 mL and 0.1 mL for oedema segmented on Analyze. The interrater bias for Osirix was 1.4 mL, -1.4 mL and 0 mL for ICH and 6.7 mL, -1.8 mL and 4.9 mL for oedema.

The mean segmentation time was 6:53 +/- 4:02 min with Analyze and 9:06 +/- 5:24 min with Osirix with a difference of 2:13 +/- 3:10 min p < 0.001 in favour of Analyze.

Comparing in-house method with Analyze automated output

The ICCs between our in-house method and Analyze’s automated output were 0.986 and 0.980 respective for ICH and oedema. The automated output underestimated ICH by 2.6 mL (-11%) p<0.001, and oedema by 2.5 mL (-12%) p<0.001( table 3, figures 3, 4). When analysis was restricted to scans with variable slice thickness (n=50), the underestimation increased to 6.4 mL (-29%) p<0.001 for ICH and 6.0 mL (-29%) p<0.001 for oedema. The automated output from Analyze overestimated ICH by 1.3 mL (+5%) p<0.001, oedema by 1.0 mL (+5%) p<0.001, for scans with uniform thickness (n=50).
**Comparing in-house method with Osirix automated output**

The analysis was performed on 98 patients as automated output failed in two patients with ROIs limited to a single CT image slice. The ICCs between the in-house method and Osirix automated output were 0.983 for ICH and 0.921 for oedema (table 3). The automated output underestimated ICH by 4.0 mL (-18%) \( p<0.001 \) and oedema by 5.5 mL (-25%) \( p<0.001 \). The underestimation in scans with variable slice thickness was 7.5 mL (-36%) \( p<0.001 \) for ICH and 9.4 mL (-41%) \( p<0.001 \) for oedema. The difference for scans with uniform slice thickness was -0.3 mL (-1%) \( p=0.045 \) for ICH and -1.7 mL (-8%) \( p<0.001 \) for oedema. The Bland-Altman plots for these comparisons are on figures 3 and 4.

**Influence of calculated voxel depth / slice thickness on volume output**

The detailed workflow for ROI volume calculation for a case with variable slice thickness and a case with uniform thickness is represented in the supplementary online spreadsheet. The non-gantry adjusted voxel depth from the infratentorial region was used as the effective voxel depth in both softwares to calculate volume in the automated output when scans were loaded as a single series (figure 1c). When compared to our in-house method, there was an underestimation of 12% and 24% using automated output from Analyze and Osirix respectively in the case with variable thickness. For the case with uniform thickness Analyze overestimated by 1% while Osirix underestimated volume by 7%.

**Discussion**

The main findings of our study are twofold. Firstly, there is excellent interrater, intrarater and intersoftware reliability in semi-automated segmentation of ICH and oedema. Secondly, our results demonstrate that the automated output from the two
softwares consistently underestimate volumes in scans with variable slice thickness by up to 41%.

**Reliability of semi-automated planimetry**

We demonstrated excellent interrater and intrarater ICCs in both Analyze and Osirix using a validated semi-automated planimetry approach [10]. The reliability of ICH measurement in our study is consistent with that reported in literature (interrater and intrarater ICCs >0.900) [18,14,12]. Divani et al demonstrated in a simulated cadaveric ICH model that Analyze software and Medical Imaging Processing, Analysis and Visualization (Center for Information Technology, National Institutes of Health, Bethesda, MD) software produced less error than ABC/2 method when compared to the actual volume of simulated blood injected but the authors did not directly compare the volumes produced by the different softwares [19]. To our knowledge our report is the first to directly compare ICH and oedema volume outputs from different softwares.

The oedema interrater ICCs are also in agreement with the largest oedema study to date - the pooled analysis of the Intensive Blood Pressure Reduction in Acute Cerebral Haemorrhage Trials (INTERACT) using the same semi-automated approach (n=1138, MIStar 3.2 software; oedema interrater ICC 0.91 for INTERACT 1, 0.93 for INTERACT 2, ICH ICCs not reported) [5]. Other threshold based planimetry segmentation including methods proposed by McCourt et al (interrater ICCs for oedema at 24 hours 0.99) [13] using an upper HU limit for oedema of 23 and the edge detection technique by Urday et al (interrater and intrarater ICCs both 0.99) [11] have also used Analyze and demonstrated excellent reliability. Osirix was also used in the Minimally Invasive Surgery and rt-PTA in ICH Evacuation (MISTIE) phase II study
for assessment of ICH and oedema volumes but rater reliability assessment was not reported [20].

Although excellent ICCs and similar volumes were demonstrated between Analyze and Osirix, there is more interrater oedema variability with Osirix. The variability in oedema volume in Osirix could have resulted from the need to ‘black out’ brain regions (online supplementary video) outside of the peri-haematomal region, which may interfere the rater’s ability to distinguish oedema from leucoaraiosis without visibility of the contralateral hemisphere.

**Implications for clinical studies**

We observed 3-4 mL or 11-18% underestimation of ICH volume and 3-6 mL or 12-25% underestimation of oedema volume using the automated software outputs. These are substantial errors in measurement, as 1 mL increase in ICH volume has been associated with 5% increase in death and dependency [21] and a 10% increase in ICH volume with 5% increase in overall mortality [22]. The >6 mL difference in ICH between our in-house output and automated output in scans with variable slice thickness may influence the outcome of clinical studies. The 6 mL difference is defined in some ongoing ICH clinical trials as the marker of significant haematoma expansion [23]. Furthermore there is emerging evidence of the volume dependent effect of peri-haematomal oedema on outcome [5,24,25]. The largest analysis of oedema on outcome comes from the pooled analysis of the INTERACT studies, demonstrating absolute oedema growth at 24 hours was independently associated with increased odds of death or dependency at 90 days (OR 1.17 (1.02-1.33), p=0.025) [5].

In this analysis, per mL of oedema growth is associated with 3% increase of poor outcome and a 10 mL growth increased the risk of poor outcome by 40%. We have identified significant volume underestimation in patients with variable slice thickness.
and given the volume dependent effect of ICH and oedema on outcome, the results of future studies will need to be considered in the context of these findings.

We have also demonstrated a 2-minute (24%) difference in segmentation time per scan in favor of Analyze over Osirix. Analyze is available commercially for an annual license renewal fee or outright purchase while the basic version of Osirix is a downloadable freeware. The cost of software and time taken for segmentation may influence the choice of software by researchers.

*Failure of automated output in variable slice thickness*

Our results indicate that software generated volume output is unreliable in scans with variable slice thickness. Analyze underestimates volume on average by nearly 30% while Osirix by approximately 40%. Our analysis indicates that the software uses the voxel depth of the infratentorial slices for volume calculation when the scans are loaded as one single series (figure 1c, supplementary excel file). A solution to this issue is loading the supratentorial and infratentorial regions separately which can be performed using the ‘DICOM tool’ function in Analyze and manually in Osirix. This approach still requires software to adjust for effect of gantry tilt on voxel depth which can additionally overestimate by including the overlap region at the interchange between supratentorial and infratentorial regions. This approach would also increase segmentation and processing time and impact on research output productivity. The automated output from Osirix uses the average area between adjacent image slices multiplied by voxel depth to calculate volume (supplementary excel file), and consequently ROIs present on a single image slice could not be derived from the automated output and is an additional source of underestimation. Previous studies have also included patients with variable slice thickness [10,14,25]. In the Efficacy of Nitric Oxide in Stroke (ENOS) study the investigators were unable to obtain
automated output from Osirix in scans with variable slice thickness and had to multiply the surface area by scan thickness to derive volume [14] while other studies did not specifically address how voxel depth was calculated [10,25]. Our in-house method remains an estimate of the factual volume but minimises potential for error.

**Limitations**

This study is limited by inclusion of mainly early scans with the majority of the scans performed within 24 hours. Oedema increases in volume and evolves over a period of two weeks while haematoma resorption takes place over the same time. Our study could not address whether semi-automated segmentation is also reliable for late oedema. We also did not have MRI images to correlate with CT based segmentation as MRI is not routinely performed for stroke diagnosis at the study centres. The reliability of the threshold based approach against MRI has been already demonstrated previously [10,11]. Thirdly our in-house calculation requires additional technical expertise and manipulation beyond the planimetric segmentation which may not be available to all research centres. A new version of Osirix (7.5) was made available after completion of the study but the approach to segmentation is unchanged. Since completion of this work we have modified the in-house Osirix plugin (developed by S.C.) to enable automated volumetric output using the best practice method which is available for download with this manuscript. Our work will help others in the future by providing codes and plugin to achieve the corrections needed to take true slice thickness into account.

**Conclusion**
Our results demonstrate that semi-automated volumetric segmentation of haematoma and oedema provides consistent interrater, intrarater and intersoftware results and can be performed in a timely manner. The method used to determine voxel depth can substantially influence volumetric measurement and this is of critical importance to the accuracy of multi-centre studies.
Acknowledgments

TYW is supported by grants from the Neurological Foundation of New Zealand and the Royal Melbourne Hospital Neuroscience Foundation. AP-J is supported by a National Institute for Health Research Clinician Scientist Award. The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the National Institute for Health Research or the Department of Health. AM is supported by grants from National Health and Medical Research Council (Australia), the Academy of Finland and the Finnish Medical Foundation.

Compliance with ethical standards

We declare that all human and animal studies have been approved by the Helsinki University Hospital and have therefore been performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments. We declare that given no identifiable patient data is presented, Helsinki University Hospital waived informed consent for this observational registry study.

Conflict of interest:

We declare that we have no conflict of interest.
References:


**Figure legend**

**Figure 1** Different methods of voxel depth calculation

Four representative consecutive brain slices with a gantry tilt of 16.5°. Slice 9 and 8 represent junction between thicker supratentorial (7.5mm) and thinner infratentorial (5.0mm) slices. Voxel depth derived from mid position of consecutive slices result in overestimation of the overlap region (a). In-house method using half distance of inter-slice midpoints (b) results in no over estimation by inclusion of overlap region.

Software automated output utilises the infratentorial voxel depth for the entire series (c) for volume calculation resulting in underestimation.

**Figure 2** Bland-Altman plots for haematoma segmentation

Bland-Altman plots with intraclass correlation coefficients (ICC) for semi-automated planimetry of hematoma segmentation on Analyze and Osirix. The solid line represents the mean difference and dotted lines the limits of 95% agreement. Analyze outliers: the two outliers with approximately 120 mL ICH both had large right frontal hematoma with heterogeneous density. The other outlier had moderate movement artifact together with moderate ventricular hemorrhage. Osirix outliers: The outliers with less than 50 mL ICH had large volume ventricular hemorrhage and the other outliers had heterogeneous ICH density.

**Figure 3.** Bland-Altman plots for haematoma volume difference between software output and in-house method

Bland-Altman plots comparing haematoma volume produced by software output from Analyze (top panels) by uniform (n=50) or variable (n=50) image slice thickness. Haematoma volume comparison with Osirix (bottom panels) were available for 49
patients with uniform slice thickness and 49 patients with variable slice thickness as automated output failed in 2 patients with region of interest limited to one single image slice. The solid line represents the mean difference and dotted lines the limits of 95% agreement.

**Figure 4.** Bland-Altman plots for oedema volume difference between software output and in-house method

Bland-Altman plots comparing oedema volume produced by software output from Analyze (top panels) by uniform (n=50) or variable (n=50) image slice thickness. Oedema volume comparison with Osirix (bottom panels) were available for 49 patients with uniform slice thickness and 49 patients with variable slice thickness as automated output failed in 2 patients with region of interest limited to one single image slice. The solid line represents the mean difference and dotted lines the limits of 95% agreement

**Supplementary figure 1** Effect of gantry tilt on voxel depth calculation

Two representative image slices of 7.5mm thickness with a gantry tilt of 16.5°. Slice thickness is derived from mid-point to mid-point distance. a) Gantry tilt was not adjusted when image position patient (DICOM header, 0020, 0032) coordinates were used resulting in an inter-slice distance of 7.82 mm, solid line. B) Gantry tilt was adjusted using our in-house method and the inter-slice distance was 7.5mm, dotted line

**Supplementary figure 2** Bland-Altman plots for oedema segmentation
Interrater Bland-Altman plots with intraclass correlation coefficients (ICC) for semi-automated planimetry of oedema segmentation on Analyze and Osirix. The solid line represents mean volume difference between raters and the dotted lines are the limits of 95% agreement. The outliers all had significant white matter changes.

**Supplementary table** Technical code of the workflow for volume calculation
Table 1. Interrater, intrarater and intersoftware intraclass correlation coefficients (ICC)

<table>
<thead>
<tr>
<th></th>
<th>Analyze, n=100</th>
<th>Osirix, n=100</th>
<th>n=100</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Interrater ICC (95% CI)</td>
<td>Intrarater ICC (95% CI)</td>
<td>Interrater ICC (95% CI)</td>
</tr>
<tr>
<td>All ICH</td>
<td>0.994 (0.992-0.996)</td>
<td>0.998 (0.996-0.999)</td>
<td>0.997 (0.996-0.998)</td>
</tr>
<tr>
<td>All Oedema</td>
<td>0.952 (0.933-0.966)</td>
<td>0.983 (0.971-0.990)</td>
<td>0.944 (0.922-0.961)</td>
</tr>
</tbody>
</table>
Table 2. Mean volume output of intracerebral haemorrhage and oedema.

<table>
<thead>
<tr>
<th>Rater</th>
<th>Analyze, n=100</th>
<th>Osirix, n=100</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ICH, mL +/- SD</td>
<td>Oedema, mL +/- SD</td>
</tr>
<tr>
<td>Rater 1</td>
<td>23.4 mL +/- 28.7</td>
<td>21.1 mL +/- 27.3</td>
</tr>
<tr>
<td>Rater 2</td>
<td>22.8 mL +/- 28.5</td>
<td>20.6 mL +/- 21.4</td>
</tr>
<tr>
<td>Rater 3</td>
<td>23.9 mL +/- 31.8</td>
<td>21.0 mL +/- 27.9</td>
</tr>
<tr>
<td>p value*</td>
<td>0.97</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Abbreviations: ICH, intracerebral haemorrhage; mL, millilitre; SD, standard deviation.
* One way ANOVA
Table 3. Reliability statistics and volume difference between in-house technique and automated output from Analyze and Osirix.

<table>
<thead>
<tr>
<th></th>
<th>Automated output, mL +/- SD</th>
<th>Analyze In-house output, mL +/- SD</th>
<th>Difference, mL (%)</th>
<th>P*</th>
<th>ICC (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Full sample, (n=100)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICH</td>
<td>20.8 +/- 28.5</td>
<td>23.4 +/- 29.6</td>
<td>-2.6 (-11%)</td>
<td>&lt;0.001</td>
<td>0.986 (0.983-0.989)</td>
</tr>
<tr>
<td>Oedema</td>
<td>18.4 +/- 23.7</td>
<td>20.9 +/- 25.6</td>
<td>-2.5 (-12%)</td>
<td>&lt;0.001</td>
<td>0.980 (0.975-0.984)</td>
</tr>
</tbody>
</table>

| **Variable thickness, (n=50)** |                             |                                   |                   |    |              |
| ICH            | 15.4 +/- 17.7               | 21.8 +/- 25.1                    | -6.4 (-29%)       | <0.001 | 0.968 (0.956-0.977) |
| Oedema         | 14.6 +/- 17.3               | 20.6 +/- 24.5                    | -6.0 (-29%)       | <0.001 | 0.961 (0.947-0.972) |

| **Uniform thickness, (n=50)** |                             |                                   |                   |    |              |
| ICH            | 26.3 +/- 35.5               | 25.0 +/- 33.5                    | 1.3 (+5%)         | <0.001 | 0.999 (0.999-0.999) |
| Oedema         | 22.1 +/- 27.3               | 21.2 +/- 26.9                    | 1.0 (+5%)         | <0.001 | 0.999 (0.998-0.999) |

| Osirix         |                             |                                   |                   |    |              |
| **Full sample, (n=98)** |                             |                                   |                   |    |              |
| ICH            | 18.4 +/- 22.3               | 22.4 +/- 27.1                    | -4.0 (-18%)       | <0.001 | 0.983 (0.979-0.987) |
| Oedema         | 16.5 +/- 20.8               | 22.0 +/- 24.5                    | -5.5 (-25%)       | <0.001 | 0.921 (0.901-0.937) |

| **Variable thickness, (n=49)** |                             |                                   |                   |    |              |
| ICH            | 13.6 +/- 16.2               | 21.1 +/- 23.6                    | -7.5 (-36%)       | <0.001 | 0.961 (0.946-0.972) |
| Oedema         | 13.5 +/- 17.0               | 22.9 +/- 25.1                    | -9.4 (-41%)       | <0.001 | 0.957 (0.940-0.969) |

| **Uniform thickness, (n=49)** |                             |                                   |                   |    |              |
| ICH            | 23.3 +/- 31.2               | 23.6 +/- 30.2                    | -0.3 (-1%)        | 0.045 | 0.999 (0.999-0.999) |
| Oedema         | 19.4 +/- 23.8               | 21.1 +/- 23.9                    | -1.7 (-8%)        | <0.001 | 0.949 (0.930-0.963) |

Abbreviations: ICC, intraclass correlation coefficients; ICH, intracerebral haemorrhage; mL, millilitre; SD, standard deviation.
* Paired T test.
** Automated output in Osirix failed in two patients (one with variable slice thickness, one with uniform slice thickness) with region of interest limited to one image slice therefore 98 patients were used in this analysis.
Automated output in variable slice thickness

**Ascending:** slice 9 thickness = slice 8 midpoint to slice 9 midpoint

**Halfpoints:** slice 9 thickness = (slice 8/9 halfpoint) + (slice 9/10 halfpoint)

**Automated output in variable slice thickness**
Mean (mL)

Rater 2 - Rater 3

60
40
20
0
-20
-40
-60

ICH segmented on Analyze

Inter-rater ICC 0.998
Rater 1 intrarater ICC 0.999

Rater 1 - Rater 2

60
40
20
0
-20
-40
-60

ICH segmented on Analyze

Inter-rater ICC 0.988
Rater 2 intrarater ICC 0.999

Rater 1 - Rater 3

60
40
20
0
-20
-40
-60

ICH segmented on Analyze

Inter-rater ICC 0.989
Rater 3 intrarater ICC 0.995

Rater 2 - Rater 3

60
40
20
0
-20
-40
-60

ICH segmented on Osirix

Inter-rater ICC 0.992
Rater 3 intrarater ICC 0.992

Rater 1 - Rater 2

60
40
20
0
-20
-40
-60

ICH segmented on Osirix

Inter-rater ICC 0.996
Rater 1 intrarater ICC 0.992

Rater 2 - Rater 3

60
40
20
0
-20
-40
-60

ICH segmented on Osirix

Inter-rater ICC 0.995
Rater 2 intrarater ICC 0.999

Rater 2 - Rater 3

60
40
20
0
-20
-40
-60

ICH segmented on Osirix

Inter-rater ICC 0.992
Rater 3 intrarater ICC 0.992
ICH volume difference in uniform slice thickness

ICH volume difference in variable slice thickness

Graph/
SCATTERPLOT
(BIVAR)
= Mean_PHO_Osirix_inhouse
WITH Diff_PHO_Osiri
x_automated_inhouse
/MISSING = LISTWISE.
EXAMINE VARIABLES = Diff_PHO_Osirix_automated_inhouse Diff_ICH_Osirix_automated_inhouse / PLOT BOXPLOT / COMPARE GROUPS / STATISTICS DESCRIPTIVES / CINTERVAL 95 / MISSING LISTWISE / NOTOTAL.
Graph:

**Oedema segmented on Osirix**

- **Interrater ICC 0.918**
- **Rater 1 intrarater ICC 0.992**

- **Interrater ICC 0.895**
- **Rater 2 intrarater ICC 0.988**

- **Interrater ICC 0.954**
- **Rater 3 intrarater ICC 0.971**

**Oedema segmented on Analyze**

- **Interrater ICC 0.934**
- **Rater 1 intrarater ICC 0.992**

- **Interrater ICC 0.898**
- **Rater 2 intrarater ICC 0.913**

- **Interrater ICC 0.936**
- **Rater 3 intrarater ICC 0.971**

- **Interrater ICC 0.918**
- **Rater 3 intrarater ICC 0.905**

**Mean (mL)**

0 30 60 90 120

**Rater 1 - Rater 2**

0 20 40 60

**Rater 1 - Rater 3**

0 20 40 60

**Rater 2 - Rater 3**

0 20 40 60

**Inter rater ICC**

0.934

**Rater 1 intrarater ICC**

0.992

**Rater 2 intrarater ICC**

0.988

**Rater 3 intrarater ICC**

0.971
Zpos is the representation of voxel depth for each slice

Calculation of Zpos using

- DICOM tag ImageOrientation which contains six values represented as Image_Ori
- DICOM tag ImagePositionPatient which contains three values represented as Image_Pos

Function Calculate_Zpos (Image_Ori, Image_Pos)

```plaintext
return Zpos}
```

We calculated interslice distance as

Function Calculate_inter_slice_distance (Zpos)

```plaintext
{For slice n = 2: last - 1
 inter_slice_distance[slice] = [ Zpos(slice n) - Zpos(slice n - 1) ] / 2 + [Zpos(slice n+1) - Zpos (slice n)] / 2
 inter_slice_distance[last] = inter_slice_distance[last - 1]
return inter_slice_distance}
```

Region of interest files were converted into MINC format (3D format) containing
- Region one representing haemorrhage
- Region two representing oedema
- Region three representing ventricular haemorrhage

for the complete volume

Volume Calculation

```plaintext
{Read MINC volume comprising region of interest from segmented from software
 Read Zpos from the MINC header

 Create a variable volume_region_one
 Create a variable volume_region_two
 Create a variable volume_region_three

 inter_slice_distance = Calculate_inter_slice_distance (Zpos) }

For each slice[first-last] of MINC volume

(Region_One_nonzero_voxels = Nonzero voxel of region one in slice (each)
Region_two_nonzero_voxels = Nonzero voxel of region two in slice (each)
Region_three_nonzero_voxels = Nonzero voxel of region three in slice (each)

Slice Volume of Region one = Region_One_nonzero_voxel * Voxel width from header * Voxel height from header * inter_slice_distance(each)
Slice Volume of Region two = Region_Two_nonzero_voxel * Voxel width from header * Voxel height from header * inter_slice_distance(each)
Slice Volume of Region three = Region_Three_nonzero_voxel * Voxel width from header * Voxel height from header * inter_slice_distance(each)

Volume_Region_One += Slice Volume of Region one
Volume_Region_Two += Slice Volume of Region two
Volume_Region_Three += Slice Volume of Region three
```
<table>
<thead>
<tr>
<th>Slice</th>
<th>Actual thickness</th>
<th>Slice number</th>
<th>Area of S</th>
<th>Volume of S</th>
<th>Volume of B</th>
<th>Area of A</th>
<th>Volume of A</th>
<th>Area of B</th>
<th>Volume of B</th>
<th>Area of A</th>
<th>Volume of A</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>5.05</td>
<td>19</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
</tr>
<tr>
<td>21</td>
<td>5.05</td>
<td>18</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
</tr>
<tr>
<td>22</td>
<td>5.05</td>
<td>17</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
</tr>
<tr>
<td>23</td>
<td>5.05</td>
<td>16</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
</tr>
<tr>
<td>24</td>
<td>5.05</td>
<td>15</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
</tr>
<tr>
<td>25</td>
<td>5.05</td>
<td>14</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
</tr>
<tr>
<td>26</td>
<td>5.05</td>
<td>13</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
</tr>
<tr>
<td>27</td>
<td>5.05</td>
<td>12</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
</tr>
<tr>
<td>28</td>
<td>5.05</td>
<td>11</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
</tr>
<tr>
<td>29</td>
<td>5.05</td>
<td>10</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
</tr>
</tbody>
</table>

DICOM HEADER DATA

**VOXEL DEPTH CALCULATED ON HEADER**

**(Zx)-(Zx-1), no gantry tilt correction**

**(Zx)-(Zx-1), corrected for gantry tilt using cosine of gantry tilt**

**(Zx+1)-(Zx)- (Zx-1))/2, corrected for cosine gantry tilt**

**THICKNESS**

**REFERENCE METHOD FOR THICKNESS**

**VOLUME**

**REFERENCE METHOD FOR VOLUME**

**AREA**

**REFERENCE METHOD FOR AREA**

**VOLUME = E * G**

**VOLUME = F * (Voxel Surface Area)**

**VOLUME = G**

**VOLUME = H**

**VOLUME = I**

**VOLUME = J**

**VOLUME = K**

**VOLUME = L**

**VOLUME = N**

**VOLUME = P**

<table>
<thead>
<tr>
<th>Slice</th>
<th>Actual thickness</th>
<th>Slice number</th>
<th>Area of S</th>
<th>Volume of S</th>
<th>Volume of B</th>
<th>Area of A</th>
<th>Volume of A</th>
<th>Area of B</th>
<th>Volume of B</th>
<th>Area of A</th>
<th>Volume of A</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>5.05</td>
<td>19</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
</tr>
<tr>
<td>21</td>
<td>5.05</td>
<td>18</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
</tr>
<tr>
<td>22</td>
<td>5.05</td>
<td>17</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
</tr>
<tr>
<td>23</td>
<td>5.05</td>
<td>16</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
</tr>
<tr>
<td>24</td>
<td>5.05</td>
<td>15</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
</tr>
<tr>
<td>25</td>
<td>5.05</td>
<td>14</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
</tr>
<tr>
<td>26</td>
<td>5.05</td>
<td>13</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
</tr>
<tr>
<td>27</td>
<td>5.05</td>
<td>12</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
</tr>
<tr>
<td>28</td>
<td>5.05</td>
<td>11</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
</tr>
<tr>
<td>29</td>
<td>5.05</td>
<td>10</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
<td>1.319</td>
</tr>
</tbody>
</table>
Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Wu, TY; Sobowale, O; Hurford, R; Sharma, G; Christensen, S; Yassi, N; Tatlisumak, T; Desmond, PM; Campbell, BCV; Davis, SM; Parry-Jones, AR; Meretoja, A

Title:
Software output from semi-automated planimetry can underestimate intracerebral haemorrhage and peri-haematoma oedema volumes by up to 41 %

Date:
2016-09-01

Citation:
Wu, TY; Sobowale, O; Hurford, R; Sharma, G; Christensen, S; Yassi, N; Tatlisumak, T; Desmond, PM; Campbell, BCV; Davis, SM; Parry-Jones, AR; Meretoja, A, Software output from semi-automated planimetry can underestimate intracerebral haemorrhage and peri-haematoma oedema volumes by up to 41 %, NEURORADIOLOGY, 2016, 58 (9), pp. 867 - 876

Persistent Link:
http://hdl.handle.net/11343/219781

File Description:
Accepted version