NEURO CRITICAL
CARE SOCIETY

ORIGINAL WORK

Exploration of Multiparameter Hematoma 3D Image Analysis for Predicting Outcome After Intracerebral Hemorrhage

Pascal Salazar[1], Mario Di Napoli[2], Mostafa Jafari[3], Alibay Jafarli[3], Wendy Ziai[4], Alexander Petersen[5], Stephan A. Mayer[6], Eric M. Bershad[7], Rahul Damani[7] and Afshin A. Divani[3,8*]

Abstract

Background: Rapid diagnosis and proper management of intracerebral hemorrhage (ICH) play a crucial role in the outcome. Prediction of the outcome with a high degree of accuracy based on admission data including imaging information can potentially influence clinical decision-making practice.

Methods: We conducted a retrospective multicenter study of consecutive ICH patients admitted between 2012–2017. Medical history, admission data, and initial head computed tomography (CT) scan were collected. CT scans were semiautomatically segmented for hematoma volume, hematoma density histograms, and sphericity index (SI). Discharge unfavorable outcomes were defined as death or severe disability (modified Rankin Scores 4–6). We compared (1) hematoma volume alone; (2) multiparameter imaging data including hematoma volume, location, density heterogeneity, SI, and midline shift; and (3) multiparameter imaging data with clinical information available on admission for ICH outcome prediction. Multivariate analysis and predictive modeling were used to determine the significance of hematoma characteristics on the outcome.

Results: We included 430 subjects in this analysis. Models using automated hematoma segmentation showed incremental predictive accuracies for in-hospital mortality using hematoma volume only: area under the curve (AUC): 0.85 [0.76–0.93], multiparameter imaging data (hematoma volume, location, CT density, SI, and midline shift): AUC: 0.91 [0.86–0.97], and multiparameter imaging data plus clinical information on admission (Glasgow Coma Scale (GCS) score and age): AUC: 0.94 [0.89–0.99]. Similarly, severe disability predictive accuracy varied from AUC: 0.84 [0.76–0.93] for volume-only model to AUC: 0.88 [0.80–0.95] for imaging data models and AUC: 0.92 [0.86–0.98] for imaging plus clinical predictors.

Conclusions: Multiparameter models combining imaging and admission clinical data show high accuracy for predicting discharge unfavorable outcome after ICH.

Keywords: Intracerebral hemorrhage, Hematoma volume, Hematoma shape, Computed tomography, CT density, Radiologic predictors, Outcomes

Introduction

Intracerebral hemorrhage (ICH) represents the second most common stroke subtype and remains a significant cause of morbidity and mortality [1–3]. Rapid

*Correspondence: adivani@gmail.com
[3] Department of Neurology, University of Minnesota, MMC 295, 420 Delaware Street S.E., Minneapolis, MN 55455, USA
Full list of author information is available at the end of the article

diagnosis and attentive management of ICH patients are crucial, since early deterioration is common in the first few hours after ICH onset [4]. Computed tomography (CT) scan is the most common imaging modality to detect acute ICH. The initial hematoma volume combined with other clinical and radiological factors is considered strong predictors of clinical outcomes and mortality after ICH [5–9]. However, hematoma volume

Springer

measurement requires manual interpretation, which carries an inherent subjectivity [10, 11]. Hematoma 3D volumetry may be more accurate and more reproducible compared to the classic ABC/2 method in assessing hematoma volume [12–15]. Furthermore, the required 3D hematoma segmentation would allow automated densitometric and geometric analysis, without the need for operator input. A resulting comprehensive hematoma analysis may complement traditional head CT analysis, since CT density patterns and geometric characteristics of the hematoma may contribute to the prediction of ICH outcome [16–20].

The primary objective of this study was to evaluate the performance of predictive models for ICH outcome using a prototype of hematoma segmentation software that can provide 3D volumetry as well as hematoma density and shape quantification. These models were developed with the hypothesis that user-independent quantitative hematoma characteristics from automated 3D segmentation of ICH (i.e., volume, shape irregularity, and density profiles) combined with other imaging or clinical variables at admission will allow accurate prediction of both mortality and poor outcome at discharge. Our secondary objective was to introduce quantitative methods for hematoma CT density and shape analysis. These methods range from a simple metric to quantify hematoma shape irregularity using sphericity index (SI) to a data-driven approach for understanding the variability of density profiles among hematomas using functional principal component (FPC) analysis. Both objectives should support our long-term target of designing automated hematoma assessment tool for the bedside applications.

Materials and Methods

The Institutional Review Boards at Hennepin Healthcare System and Fairview Health Services in Minneapolis, Minnesota, approved the study prior to initiation. Consecutive spontaneous ICH admissions between January 2012 and December 2017 from the two healthcare systems were included. The data collection protocol for the study has been described previously [21, 22]. ICH subjects were identified using International Classification of Diseases (ICD) coding system with the primary or secondary diagnosis codes of 431, 432, and 432.9 for the 9th Revision (ICD-9) and I61.0–I61.9 and I62.9 diagnosis codes for the ICD-10 (when applicable). Cases were cross-matched with a prospectively collected stroke database to account for missing cases and coding misclassification. We excluded secondary hemorrhage due to other etiologies such as hemorrhagic transformation, venous infarction, trauma, tumor (primary or cerebral metastasis), ruptured vascular malformation, subarachnoid, subdural, or epidural hemorrhage.

Demographic data, medical history, time of symptom onset, vital signs, laboratory tests, initial CT scan, in-hospital treatment, and length of stay were recorded in a standardized data collection form. Baseline neurological and functional deficits were documented using the National Institutes of Health Stroke Scale, Glasgow coma scale (GCS), ICH-score [23], and modified Rankin Scale (mRS), when available. Outcomes at discharge were defined as death or composite unfavorable outcome (severe disability or death, mRS 4–6).

Post-processing of Hematoma in Initial CT Scan

Initial CT scans were post-processed using a research version of Vitrea software (v7, Vital Images, Minnetonka, MN, USA). After skull removal, semiautomatic hematoma segmentation was applied. For each case, the user defined a bounding box where the hematoma segmentation was applied using 1 click extend action on a single axial or coronal view. Then, the algorithm segmented the hematoma based on 3D histogram analysis performed on CT densities and various morphologic operations. Hematoma volume as well as hematoma density (both mean and SD values) was, then, automatically computed. Hematoma density 3D histograms data were obtained and exported for further analysis (Fig. 1a).

To assess the hematoma shape irregularity (Fig. 1b), we used the following sphericity definition as a shape irregularity index:
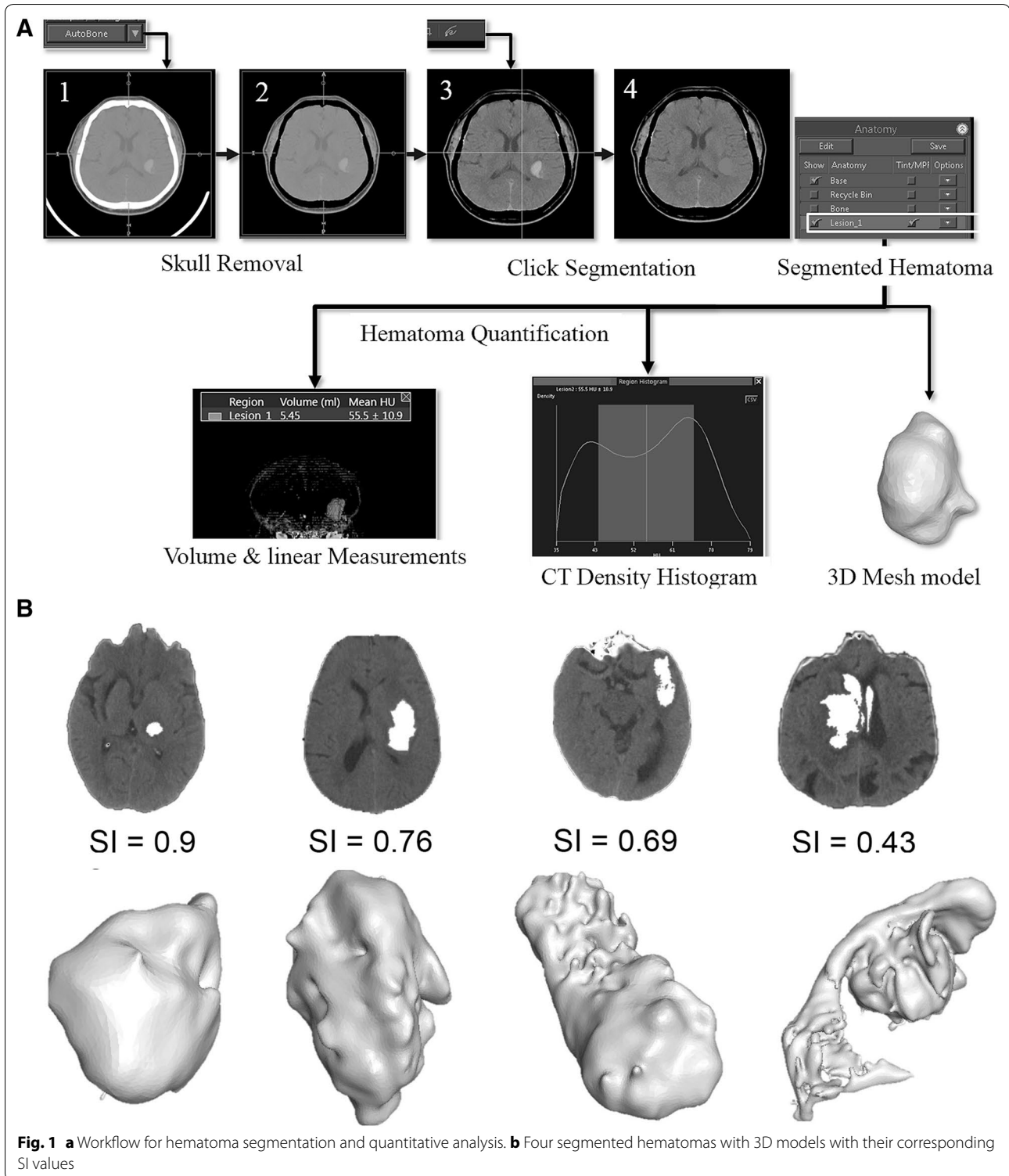
$$\text{Sphericity Index (SI)} = \pi^{1/3}(6V)^{2/3}/A$$

where "V" represents the volume and "A" represents the surface area of the hematoma. The proposed SI quantifies the irregularity of the hematoma shape taking the sphere (SI = 1.0) [24, 25]. Practically, SI for brain hematoma ranges from mostly regular ellipsoids (SI: 0.90–0.80) to a highly irregular shape (SI: 0.2–0.4) in a continuous scale.

To assess intra- and inter-rater variabilities, two trained readers repeated the hematoma segmentation of 50 hematomas twice using the hematoma probe prototype. For ground truth measurements, all the hematoma volumes measurements were repeated using a semiautomated planimetry measurement using the Medical Image Processing, Analysis, and Visualization (MIPAV, National Institute of Health) software [12, 22].

The midline shift (MLS) was determined as a binary variable (yes/no) and an alternative 4-class ordered MLS (0–2 mm [no MLS], 2–5 mm, 5–10 mm, and ≥ 10 mm). The MLS was measured as the perpendicular distance between septum pellucidum and a midline connecting free edges of the anterior and posterior falx [26].

Using hematoma CT density curves, we extracted mean CT (HU), standard deviation CT (HU), and 3

Fig. 1 a Workflow for hematoma segmentation and quantitative analysis. **b** Four segmented hematomas with 3D models with their corresponding SI values

continuous variables (FPC1, FPC2, and FPC3) describing the variation in these curves using a functional data analysis described later.

Statistical Methods

The statistical analysis was performed using the R-statistical programming environment (version 3.2.0, R Core Team, Vienna, Austria). Characteristics of each outcome

(in-hospital mortality and severe disability, mRS 4–6) were computed using median, interquartile range (IQR), Mann–Whitney tests for continuous variables, and Fisher exact tests for binary variables. The agreement between automated volume measurement and reference volume and the intra-rater agreement were assessed using intra-class correlation (ICC).

The quantitative (continuous) and qualitative (binary) predictors of patient outcome have been jointly analyzed using a mixed-type data approach [27] where generalized association values (0–100%) were computed between quantitative and qualitative predictors and summarized in a heatmap plot (see Figure 1s in Supplemental Materials). Confounder plots were generated to visualize the association of any predictors with the patient outcome, with the hematoma volume taken as a reference major predictor (see Figures 2s and 3s in Supplemental Materials).

The univariate analysis of the predictors was performed for both unfavorable outcome at discharge (mRS: 4–6) and in-hospital mortality using a receiver operating characteristic (ROC) area under the ROC curve (AUC) metric for continuous predictors and odds ratio (OR) for binary predictors. Optimal cutoff points were determined using the Youden index method. AUC confidence intervals were computed with DeLong's method, and significant differences between AUCs were tested using DeLong's method for correlated AUCs 2-sided tests. In the multivariate analysis, the linearity assumptions and the presence of interactions for the main predictors in logistic regression models were assessed using nonlinear regression plots, ANOVA, and Wald tests [28], see Figures 5s–6s and Tables 1s and 2s in Supplemental Materials. Multivariable logistic models and one model using a random forest classifier were evaluated for their performance to predict in-hospital mortality or unfavorable outcome using custom programs written in R-language [29]. Training of the classifiers was performed on a random sample of 80% of the subjects with tenfold cross-validation while testing the predictive performances used the remaining 20% independent testing set. Performances were assessed as ROC AUC, sensitivity, specificity, and accuracy. Initial variable selection in model building was based on relative univariate AUC variable performances, variable interaction tests results, and the degree of correlation between model variables using a mixed data analysis as shown in Supplemental Materials. Model selection was based on the AUC criterion. Calibration assessment was performed to evaluate the quality of predicted probabilities of the different models using a bootstrapping method detailed in Supplemental Materials (Figures 7s–16s).

Functional Data Analysis of Hematoma CT Density Curves

A functional data analysis was performed on the hematoma density curves to visualize their main modes of variation in our cohort and to extract data-driven density features for patient outcome classification [30, 31]. Original CT histograms were converted to smooth curves defined between 0 and 100 HU [32]. Three FPCs explaining 87.7% of the variation between density curves were selected following Petersen and Müller's FPC method for frequency distributions [33]. The resulting FPC scores (FPC1, FPC2, and FPC3) are non-correlated new variables related to the variation in the density curves to be added to our set of patient outcome predictors (Fig. 2)—see further details on statistical analysis in Supplemental Materials.

Results

We included 430 subjects in the present analysis. The usability performance of the current automated segmentation was assessed on a sample of 30 randomly selected cases using the number of user clicks to obtain the final segmentation (median: 1, IQR: [1–3]). The median segmentation time including computation and user visual assessment was 13.7s (IQR: 16.9s). The ICC between automated and MIPAV volumes was 0.91 (95% CI 0.81–0.96) using a random sample of 50 cases. The intra-rater and inter-rater agreements of automated segmented hematoma volume were 0.98 (95% CI 0.97–0.99) and 0.96 (95% CI 0.94–0.98), respectively, on the same case sample.

The baseline characteristics are shown in Table 1. Table 2 shows the ROC AUC, sensitivity, specificity, and the cutoff values for the variables used in the predictive models. The imaging findings that are shown in Table 3 were significantly associated with death and poor functional outcomes (mRS 4–6). Table 4 summarizes the predicting models for mortality and composite unfavorable outcome (mRS 4–6) at discharge. Our best predictive logistic regression model for mortality included age, admission GCS, hematoma volume, MLS, the interaction of hematoma volume * MLS, and hemorrhage in the pons. Our best predictive logistic regression model for discharge composite unfavorable outcome included age, admission glucose, admission GCS, hematoma volume, SI, MLS, interaction of hematoma volume * SI, interaction of hematoma volume * GCS, and hemorrhage in the pons as predictors. Beside classification accuracy, our calibration assessment confirmed that all models predicting death and the two most accurate models predicting composite unfavorable outcome provide an acceptable prediction of outcome probabilities with mean absolute errors between 0.007 and 0.022 (see Figure 7s to Figure 16s—Supplemental Materials).
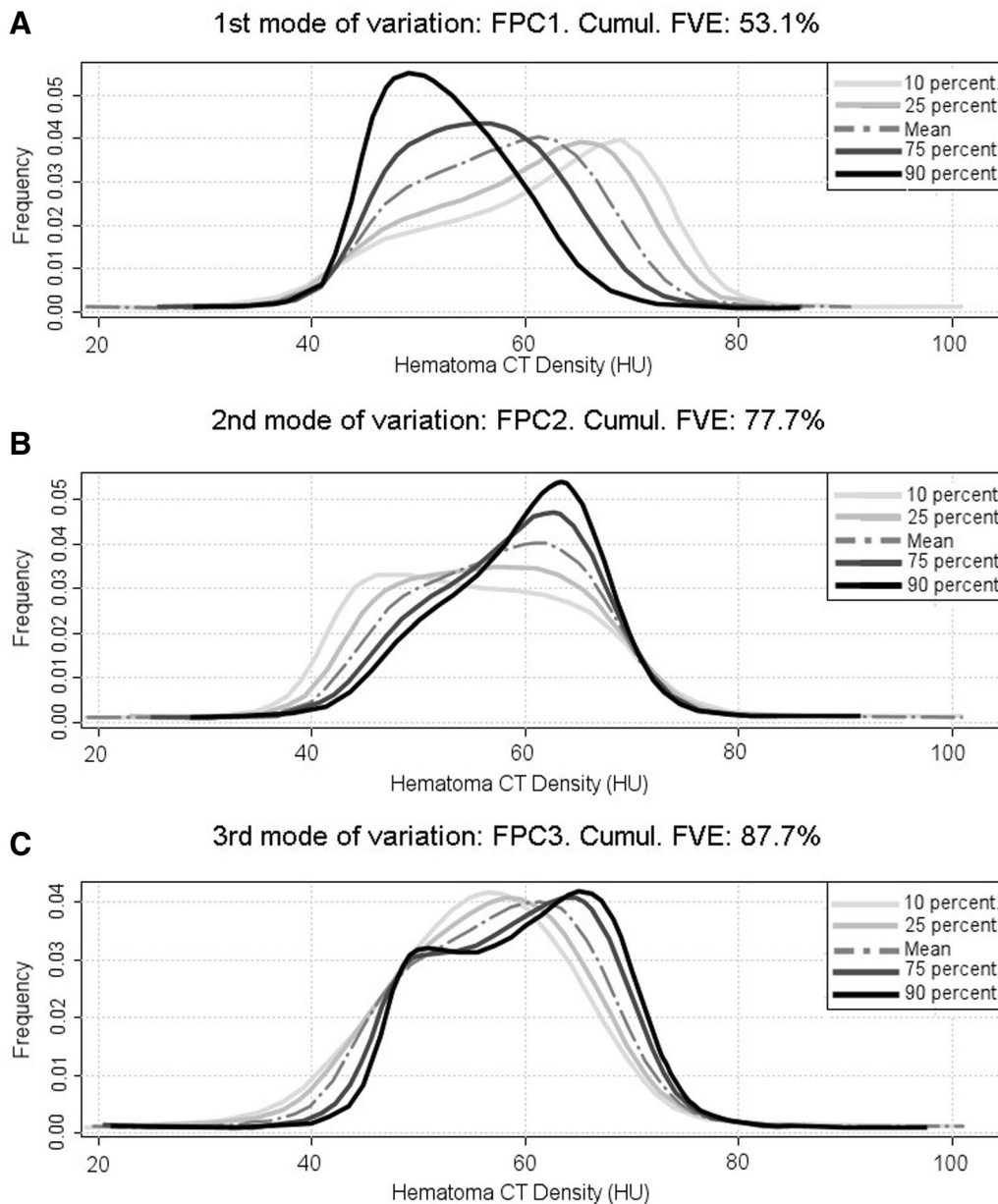
Fig. 2 Functional principal component (FPC) scores for hematoma density profiles. **a** FPC1: First mode of variation in the hematoma density curves starting from highly heterogeneous higher densities curves (10th percentile—light gray curve) moving toward heterogeneous curve profiles (25th percentile, mean, 75th percentile) ending with more homogeneous low densities hematoma (90th percentile—black curve). **b** FPC2: second independent type of variation starting from highly heterogeneous pattern of mixed low (predominant) and mid-densities (10th percentile—light gray) to a moderately high-density profile (90th percentile—black). **c** FPC3: Third independent mode of variation in hematoma density profile shows hematoma density transitioning from homogeneous curves (10th percentile—light gray) centered around 55HU to bimodal high-density dominant curves (around 65–70 HU and lower densities (40–45 HU)—90th percentile—black)

Discussion

Among the admission radiologic imaging variables, ICH volume has long been known as a strong predictor of mortality [34]. Our univariate analysis confirmed the value of ICH volume for the prediction of in-hospital death (AUC: 0.82 [0.83–0.89]) and composite unfavorable

outcome (AUC: 0.83 [0.79–0.86]). The optimal cutoff point for the hematoma volumes was > 41.6 ml (Table 2). Other published studies have reported hematoma volume cutoff values of > 32 ml (supratentorial) and > 21 ml (infratentorial) for predicting 30-day mortality [35], 40 ml cutoff volume predicting a poor outcome in patients with

Table 1 Baseline continuous characteristics of ICH per in-hospital mortality and disability outcome

Variables	In-hospital death			Composite unfavorable outcome		
	Discharged alive (n = 298)	Died in hospital (n = 132)	p value	mRS 0–3 (n = 155)	mRS 4–6 (n = 275)	p value
Demographics						
Age (year)	67 [57–79]	73 [58.5–83]	0.037	64 [56–75]	71 [59–81]	< 0.001
Weight (lb)	169 [141–202]	169 [140–194.5]	0.873	172 [143.25–207]	165 [141–192.8]	0.140
Sex (M/F)	169/129 (56.71%)	73/59 (55.30%)	0.830	96/59 (61.94%)	146/129 (53.09%)	0.085
Baseline clinical characteristics						
GCS	14 [11–15]	3 [3–7.5]	< 0.001	15 [14, 15]	8 [3–13]	< 0.001
DBP (mmHg)	94 [80–109]	96.5 [78–111]	0.759	93 [78.25–107]	96 [80–111]	0.189
SBP (mmHg)	168.5 [143–192]	161 [142–197]	0.831	164 [140–190]	172 [146.25–197]	0.040
Creatinine (mg/dl)	0.89 [0.72–1.1]	0.94 [0.75–1.195]	0.039	0.88 [0.70–1.10]	0.9 [0.72–1.15]	0.417
Glucose (mg/dl)	124 [105–151]	158.5 [134–204]	< 0.001	118 [103–137.75]	147 [116.25–181]	< 0.001
INR	1.07 [1–1.2]	1.1 [1–1.45]	0.017	1.08 [1–1.19]	1.1 [1–1.3]	0.037
Warfarin intake	246/44	247/28	0.077	136/18	194/54	0.012
DNR (1/0)	8/290 (2.68%)	16/116 (12.12%)	< 0.001	2/153 (1.31%)	22/253 (8.0%)	0.003
Baseline imaging characteristics						
Onset to CT (h)	1.5 [1–3]	1.5 [1–2.5]	0.347	2 [1–3.5]	1.5 [1–2.5]	0.008
Volume (ml)	9.41 [3.01–26.46]	61.74 [26.83–91.78]	< 0.001	6.08 [2.08–12.29]	34.4 [13.93–70.58]	< 0.001
SI	0.63 [0.49–0.73]	0.45 [0.36–0.54]	< 0.001	0.67 [0.55–0.74]	0.51 [0.41–0.64]	< 0.001
Mean CT (HU)	56.45 [52.8–60.1]	57.15 [53.2–60.8]	0.227	56.1 [52.83–59.8]	56.9 [53.1–61.03]	0.070
SD CT (HU)	8.1 [6.9–9.2]	8.95 [7.55–10.35]	< 0.001	8.1 [6.4–7.3]	8.1 [6.9–9.1]	0.022
FPC1	0.0084 [− 0.118 to 0.175]	− 0.1016 [−0.225–0.0592]	< 0.001	0.0267 [− 0.076 to 0.193]	− 0.0718 [− 0.187 to 0.114]	< 0.001
FPC2	0.0021 [− 0.110 to 0.102]	0.0207 [− 0.081 to 0.015]	0.376	− 0.045 [− 0.145 to 0.082]	0.025 [− 0.073 to 0.11]	0.001
FPC3	− 0.015 [−0.092–0.059]	0.027 [− 0.034 to 0.074]	< 0.001	− 0.031 [0.10–0.059]	0.016 [0.067–0.072]	0.005
Lobar	127/171 (42.62%)	63/69 (47.73%)	0.340	63/92 (40.65%)	127/148 (46.18%)	0.310
Basal ganglia	86/212 (28.86%)	39/93 (29.55%)	0.910	44/111 (80.00%)	81/194 (29.45%)	0.830
Midbrain	5/293 (1.68%)	12/120 (9.09%)	< 0.001	3/152 (1.94%)	14/261 (5.09%)	0.130
Thalamus	57/241 (19.13%)	19/113 (14.39%)	0.270	31/124 (20.00%)	45/230 (16.36%)	0.360
Cerebellum	26/27 (28.72%)	13/119 (9.85%)	0.720	14/141 (9.03%)	25/250 (9.09%)	1.000
Pons	10/288 (3.36%)	18/114 (13.64%)	< 0.001	4/151 (2.58%)	24/251 (8.73%)	0.014
IVH	107/191 (35.91%)	98/34 (74.24%)	< 0.001	45/110 (29.03%)	160/115 (58.18%)	< 0.001
MLS	64/234 (21.48%)	80/52 (60.61%)	< 0.001	21/134 (13.55%)	123/152 (44.73%)	< 0.001
Hydrocephalus	20/278 (6.71%)	45/87 (34.09%)	< 0.001	7/148 (4.52%)	58/217 (21.09%)	< 0.001
Herniation	15/283 (5.03%)	35/97 (26.52%)	< 0.001	3/15 2 (1.94%)	47/228 (17.09%)	< 0.001

CT computed tomography, *DBP* diastolic blood pressure, *DNR* do not resuscitate, *FPC* functional principal component, *GCS* Glasgow Coma Scale, *ICH* intracerebral hemorrhage, *INR* international normalized ratio, *IVH* intraventricular hemorrhage, *mRS* modified Rankin Scale, *MLS* midline shift, *SBP* systolic blood pressure, *SI* sphericity index

lobar hemorrhage [36], and > 60 ml cutoff point for the highest mortality rate [37]. The difference in cutoff values may be due to different outcome scales, different follow-up times, or different studied cohorts [35–37].

Our multivariate analysis revealed a significant interaction between hematoma volume and MLS ($p = 0.038$) for mortality prediction adjusted for SI, hematoma in the cerebellum and the pons (see Figure 5s and Table 1s in Supplemental Materials). It should be noted that a strong nonlinear effect seen in the OR change for 4-class ordered MLS > 10 mm (see Table 3) and the interaction MLS*volume in our model may explain the lack of significant benefit of using 4-category MLS over the binary MLS.

Further, we also observed a significant interaction between hematoma volume and GCS ($p = 0.004$) for severe disability prediction adjusted for hemorrhage in the pons and age (see Figure 6s and Table 2s in Supplemental Materials). Regarding SI, irregular-shaped hematomas were associated with larger hematoma,

Table 2 In-hospital death and composite unfavorable outcome cutoff points

Variable	In-hospital death				Composite unfavorable outcome (mRS: 4–6)			
	Cutoff point	AUC [95% CI]	Sensitivity	Specificity	Cutoff point	AUC [95% CI]	Sensitivity (%)	Specificity (%)
GCS	<11	0.90 [0.87–0.93]	93.18%	73.49%	≤12	0.88 [0.84–0.91]	73.45	90.97
Age (years)	>72	0.56 [0.51–0.61]	50.76%	63.09%	>74	0.60 [0.55–0.64]	45.45	74.84
Admission glucose (mg/dl)	>133	0.72 [0.67–0.76]	77.27%	60.40%	>133	0.69 [0.65–0.74]	63.64	70.97
Initial hematoma volume (ml)	>41.8	0.86 [0.83–0.89]	68.18%	87.58%	>23.6	0.83 [0.79–0.86]	64.0	90.32
SI	≤0.54	0.78 [0.73–0.81]	75.75%	67.61%	≤0.54	0.71 [0.67–0.75]	58.18	75.48
FPC1	<0.078 [40th percentile]	0.62 [0.58–0.67]	49.45%	75.48%	<0.078 [40th percentile]	0.64 [0.59–0.68]	56.82	67.45
FPC2	NA	NA	NA	NA	>0.032 [40th percentile]	0.59 [0.55–0.64]	67.64	54.19
FPC3	>−0.052 [35th percentile]	0.61 [0.56–0.65]	80.3%	42.3%	>0.080 [26th percentile]	0.58 [0.53–0.63]	80.0	36.77

CI conidence interval, *FPC* functional principal component, *GCS* Glasgow Coma Scale, *mRS* modiied Rankin Scale, *SI* sphericity index

Table 3 Imaging predictors of in-hospital outcomes

Predictors	In–hospital death OR [95% CI]	Composite unfavorable outcome OR [95% CI]
MLS (binary)	5.5 [3.53–8.57]	9.63 [5.41–17.34]
MLS-2 (2–5 mm)	3.12 [1.48–6.57]	1.58 [0.76–3.32]
MLS-3 (5–10 mm)	2.92 [1.52–5.59]	1.95 [1.95–9.56]
MLS-4 (>10 mm)	15.76 [7.95–31.22]	50.14 [6.85–366.92]
IVH	4.97 [3.16–7.82]	7.12 [4.22–11.99]
Hydrocephalus	7.1 [3.97–12.63]	10.81 [4.67–25.02]
Herniation	6.7 [3.52–12.82]	18.1 [5.42–60.45]
Hemorrhage in pons	4.49 [2.013–10.02]	5.91 [1.95–17.93]

CI confidence interval, *IVH* intraventricular hemorrhage, *mRS* modified Rankin Scale, *MLS* midline shift, binary and 4-class, *OR* odds ratio

intraventricular hemorrhage (IVH), death, and severe disability at the time of discharge. Both SI cutoff values for mortality and severe disability were <0.54. SI was a significant predictor of both death ($p < 0.001$) and composite unfavorable outcome ($p = 0.002$) even after adjusting for hematoma volume and MLS. The previous studies defining the irregularity of the hematoma shape by visual inspection have shown a relationship between the irregular shape and poor clinical outcome at 30 days and 3 months post-ICH [19, 38]. The underlying mechanism might be the higher risk of hematoma expansion (HE) in irregular-shaped hematomas, which is a predictor of poor clinical outcomes in patients with ICH [19].

Our univariate study showed that heterogeneous hematomas with higher-density profiles were only moderately associated with higher mortality (FPC1, AUC: 0.64) and severe disability (FPC1, AUC: 0.62). FPC1 was also moderately associated with hematoma volume (see

Figures 1s–4s in Supplemental Materials). This correlation partly explains the increased mortality associated with FPC1 higher-density heterogeneous profiles. However, FPC1 remained a significant predictor of mortality using Wald test ($p = 0.0272$) even after adjustment for hematoma volume but only marginally significant ($p = 0.052$) in a model to predict mortality after adjustment for hematoma volume and location. A heterogeneous density hematoma might reflect active hemorrhage, an extended hemorrhagic time course, and multifocality which might be a predictor of HE and worse outcomes [39]. FPC2 predictor was not significantly associated with the patient outcome. It is worth mentioning that Nowinski et al. [40] presented CT density profiles of IVH-ICH acquired at different times during the acute and subacute phases (1–7 days) that are strikingly similar to FPC2 curves shown in Fig. 2, suggesting a possible (yet unconfirmed) common interpretation in terms of hematoma maturation. Considering the FPC3 predictor, heterogeneous bimodal low- and high-density curves were also associated with mortality compared to unimodal medium-density curves. These results are consistent with the previous studies mentioning hematoma density heterogeneity as indicative of poor clinical outcome and mortality [16, 20, 39]. In an exploratory sub-study analysis using the 316 cases of our dataset with follow-up CT scan, FPC3 was also found to be the best significant CT imaging predictor ($p = 0.0138$) of the HE (defined as >33% or >6 ml increase)—see Supplemental Materials. FPC3 was significantly associated with HE even after combining it with the other significant clinical predictors at baseline: INR ($p = 0.0015$), creatinine ($p = 0.0042$), and GCS score ($p = 0.0372$) with a modest overall accuracy

Table 4 Prediction models for in-hospital mortality and severe disability (mRS 4–6) at discharge

Models for in-hospital mortality				
Predictors: Imaging only for in-hospital mortality				
Performance	**Accuracy**	**Sensitivity**	**Specificity**	**AUC**
Model A1: volume only				
Multiple logistic regression	81% [71–89%]	69%	86%	0.85 [0.76–0.93]
Model B1: volume + SI + interaction (volume*SI)				
Multiple logistic regression	81% [71–89%]	73%	85%	0.86 [0.77–0.94]
Model C1: volume + MLS + interaction (volume*MLS) + SI + pons + cerebellum				
Multiple logistic regression	82% [74–90]	81%	85%	0.91 [0.86–0.97]
Random forest classifier	82% [73–90]	69%	90%	0.91 [0.84–0.97]
Predictors: clinical and imaging for in–hospital mortality				
Model D1: GCS + volume + Pons + Age				
Multiple logistic regression	88% [79–94]	81%	92%	0.94 [0.89–0.99]
Model E1: GCS + volume + MLS + interaction (volume*MLS) + pons + age				
Multiple logistic regression	87% [78–93]	81%	90%	0.95 [0.90–0.99]
Models for severe disability				
Predictors: Imaging only for severe disability				
Model A2: volume only				
Multiple logistic regression	74% [64–83]	87%	52%	0.84 [0.76–0.93]
Model B2: volume + SI				
Multiple logistic regression	78% [68–86]	87%	58%	0.85 [0.77–0.93]
Model C2: volume + SI + interaction (volume*SI) + pons				
Multiple logistic regression	79% [69–87]	87%	68%	0.88 [0.80–0.95]
Predictors: clinical and imaging for severe disability				
Model D2: volume + GCS + interaction (volume*GCS) + age + pons				
Multiple logistic regression	85% [76–92]	82%	87%	0.92 [0.86–0.98]
Model E2: volume + MLS + interaction (volume*MLS) + SI + GCS + interaction (GCS*SI) + pons + high glucose + age				
Multiple logistic regression	85% [76–91]	84%	87%	0.92 [0.86–0.98]

FPC functional principal component, *GCS* Glasgow Coma Scale, *MLS* midline shift, *mRS* modiied Rankin Scale, *SI* sphericity index. AUCs significant different from volume-only Model E (DeLong method for correlated AUCs). Model E1: *z*-score: 2.88, $p = 0.0039$. Model D1: *z*-score: 2.69, $p = 0.007$. Model C1: *z*-score: 2.69, $p = 0.007$. Model E2: *z*-score: 1.87, $p = 0.062$. Model D2: *z*-score: 1.98, $p = 0.047$

of AUC: 0.70 [0.65 to 0.75]) for the logistic regression. Moreover, FPC3 values were significantly higher in HE group ($n = 52$) compared to no HE group ($n = 378$), Mann–Whitney test ($p = 0.0009$). See Figures 16s–17s in Supplemental Materials. These findings are consistent with the previous studies showing the role of heterogenous hematoma density in the prediction of HE [16–18, 20].

Nonetheless, further detailed investigation on hematoma density profiles is needed to fully understand the way they affect the outcome. Our study also showed that IVH was significantly associated with mortality and severe disability that is in line with the previous studies [41]. However, IVH did not improve the prediction of the outcome (data are not shown) and was therefore excluded from the final models.

The present study shows high predictive accuracy of the best models for in-hospital mortality using both imaging and clinical variables (Model E1: AUC: 0.95 [0.90–0.99]) and models with radiologic imaging only variables (Model C1: AUC: 0.91 [0.86–0.97]). Similarly, we observed a high prdictive accurarcy of our best models for composite unfavorable outcome using both imaging and clinical variables (Model E2: AUC: 0.92 [0.86–0.98]) and imaging only variables (Model C2: AUC: 0.88 [0.80–0.95]). These results are in agreement with previous predictive models summarized by Gregório et al. [42] meta-analysis which reported a pooled C statistic (AUC) of 0.880 for mortality prediction and 0.872 for functional outcome prediction. Logistic regression models performed as well as more complex nonlinear models such as random forests (see Model C1 in Table 4) in agreement with Gregório et al. [42] results.

Since the logistic regression models also allow fast computation and are easily interpretable, they are the best candidates for future automated classifiers. Our

separation of predictive models between an imaging only predictor group (A1-2, B1-2, and C1-2) and a composite imaging-clinical predictor group (D1-2 and E1-2) anticipates future practical challenges associated with the second group as automatic retrieval of clinical variables in emergency settings is more complex than accessing only the radiologic imaging data since typical CT images contain only basic demographic information (i.e., age and gender). Despite better average performances, best composite models were not significantly more accurate than the best imaging only models. Moreover, the redundancy of predictors in clinical applications may be beneficial when clinical or imaging data are missing or borderline. Nonetheless, knowing the mild restrictive inclusion and exclusion criteria used in this study, performances of both types of models support the idea of feasibility for future automated predictive systems.

Limitations of the study include the inability to rule out inherent methodological issues given the retrospective nature of the study. Additionally, only hospital discharge outcomes were considered for the study, and associations with long-term functional outcomes remain to be validated. Other than an exploratory review of HE predictors including FPC3, this study did not assess HE due to a significant number of patients ($n = 114$) with missing follow-up CT scans due to different reasons such as death, care directive measures initiated before the follow-up scan, or the patient who had undergone hematoma evacuation within the first 24 h of admission. Even though one can use different statistical techniques to adjust for such variations, the task is not a straightforward and may lead to unreliable results. Nonetheless, the objective of the present study was to predict the outcome solely based on the admission data. In our model, we didn't compare the 3D hematoma volume measurement method with the ABC/2 method since the planimetry and 3D techniques have been already compared with the ABC/2 method to highlight the shortcomings of the ABC/2 method [12]. The validation of our hematoma segmentation is based on hematoma volume comparison. However, a more accurate voxel-based method such as Dice's similarity coefficient or Jaccard index is needed for validating the performance of the final version of software.

Regarding the proposed quantitative features, the SI mostly represents large-scale shape regularity and thus cannot represent lower-scale heterogeneity (e.g., lobar hematomas). Future studies may consider recent advances applied to tumor shapes using statistical shape analysis [43], 3D spherical harmonics [44], or topological data analysis [45]. Lastly, statistical tests applied on predictive models (either only imaging or composite models) could not establish that models using the new quantitative features are more accurate than their counterpart without these features. Therefore, further validation with large independent datasets will be required.

Conclusion

Automated 3D hematoma segmentation on CT images can offer a comprehensive set of user-independent volumetric, shape-related, and CT densitometric quantitative predictors of the ICH patient outcome. Predictive models combining these imaging predictors with other admission clinical variables demonstrate a high-accuracy prediction of mortality and disability, opening the way to future integrated decision support tools for the clinicians at the time of admission. Future work should determine the best options for such tools including the degree of integration of clinical information from medical history, patient monitoring system, touch screen-based user-initiated segmentation, visualization of outcome probability conditioned to treatment options, and database query of similar past cases.

Author details
[1] Vital Images, Minnetonka, MN, USA. [2] Department of Neurology, San Camillo de' Lellis District General Hospital, Rieti, Italy. [3] Department of Neurology, University of Minnesota, MMC 295, 420 Delaware Street S.E., Minneapolis, MN 55455, USA. [4] Department of Neurology, Neurosurgery, and Anesthesia/Critical Care Medicine, Johns Hopkins, Baltimore, USA. [5] Department of Statistics and Applied Probability, University of California, Santa Barbara, CA, USA. [6] Department of Neurology, Henry Ford Health System, Detroit, MI, USA. [7] Department of Neurology, Baylor College of Medicine, Houston, TX, USA. [8] Department of Neurosurgery, University of Minnesota, Minneapolis, MN, USA.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Published online: 29 July 2019

References

1. Hemphill JC 3rd, Greenberg SM, Anderson CS, Becker K, Bendok BR, Cushman M, et al. Guidelines for the management of spontaneous intracerebral hemorrhage: a guideline for healthcare professionals from the American Heart Association/American Stroke Association. Stroke. 2015;46:2032–60.
2. Feigin VL. Stroke in developing countries: can the epidemic be stopped and outcomes improved? Lancet Neurol. 2007;6:94–7.
3. Feigin VL, Lawes CM, Bennett DA, Anderson CS. Stroke epidemiology: a review of population-based studies of incidence, prevalence, and case-fatality in the late 20th century. Lancet Neurol. 2003;2:43–53.
4. Moon JS, Janjua N, Ahmed S, Kirmani JF, Harris-Lane P, Jacob M, et al. Prehospital neurologic deterioration in patients with intracerebral hemorrhage. Crit Care Med. 2008;36:172–5.
5. Yang W-S, Li Q, Li R, Liu Q-J, Wang X-C, Zhao L-B, et al. Defining the optimal midline shift threshold to predict poor outcome in patients with supratentorial spontaneous intracerebral hemorrhage. Neurocrit Care. 2018;28:314–21.
6. Kiphuth IC, Huttner HB, Breuer L, Schwab S, Kohrmann M. Sonographic monitoring of midline shift predicts outcome after intracerebral hemorrhage. Cerebrovasc Dis. 2012;34:297–304.
7. Fogelholm R, Murros K, Rissanen A, Avikainen S. Admission blood glucose and short term survival in primary intracerebral haemorrhage: a population based study. J Neurol Neurosurg Psychiatry. 2005;76:349–53.
8. Broderick JP, Brott TG, Duldner JE, Tomsick T, Huster G. Volume of intracerebral hemorrhage. A powerful and easy-to-use predictor of 30-day mortality. Stroke. 1993;24:987–93.
9. Al-Shahi Salman R, Frantzias J, Lee RJ, Lyden PD, Battey TWK, Ayres AM, et al. Absolute risk and predictors of the growth of acute spontaneous intracerebral haemorrhage: a systematic review and meta-analysis of individual patient data. Lancet Neurol. 2018;17:885–94.
10. Chalela JA, Kidwell CS, Nentwich LM, Luby M, Butman JA, Demchuk AM, et al. Magnetic resonance imaging and computed tomography in emergency assessment of patients with suspected acute stroke: a prospective comparison. Lancet. 2007;369:293–8.
11. Rodriguez-Luna D, Boyko M, Subramaniam S, Klourfeld E, Jo P, Diederichs BJ, et al. Magnitude of hematoma volume measurement error in intracerebral hemorrhage. Stroke. 2016;47:1124–6.
12. Divani AA, Majidi S, Luo X, Souslian FG, Zhang J, Abosch A, et al. The ABCs of accurate volumetric measurement of cerebral hematoma. Stroke. 2011;42:1569–74.
13. Webb AJ, Ullman NL, Morgan TC, Muschelli J, Kornbluth J, Awad IA, et al. Accuracy of the ABC/2 Score for Intracerebral Hemorrhage: systematic Review and Analysis of MISTIE, CLEAR-IVH, and CLEAR III. Stroke. 2015;46:2470–6.
14. Hussein HM, Tariq NA, Palesch YY, Qureshi AI. Reliability of hematoma volume measurement at local sites in a multicenter acute intracerebral hemorrhage clinical trial. Stroke. 2013;44:237–9.
15. Wang CW, Juan CJ, Liu YJ, Hsu HH, Liu HS, Chen CY, et al. Volume-dependent overestimation of spontaneous intracerebral hematoma volume by the ABC/2 formula. Acta Radiol. 2009;50:306–11.
16. Boulouis G, Morotti A, Brouwers H, et al. Association between hypodensities detected by computed tomography and hematoma expansion in patients with intracerebral hemorrhage. JAMA Neurol. 2016;73:961–8.
17. Barras CD, Tress BM, Christensen S, Collins M, Desmond PM, Skolnick BE, et al. Quantitative CT densitometry for predicting intracerebral hemorrhage growth. AJNR Am J Neuroradiol. 2013;34:1139–44.
18. Boulouis G, Morotti A, Brouwers HB, Charidimou A, Jessel MJ, Auriel E, et al. Noncontrast computed tomography hypodensities predict poor outcome in intracerebral hemorrhage patients. Stroke. 2016;47:2511–6.
19. Delcourt C, Zhang S, Arima H, Sato S, Al-Shahi Salman R, Wang X, et al. Significance of hematoma shape and density in intracerebral hemorrhage: the intensive blood pressure reduction in acute intracerebral hemorrhage trial study. Stroke. 2016;47:1227–32.
20. Zhang D, Chen J, Xue Q. Heterogeneity signs on noncontrast computed tomography predict hematoma expansion after intracerebral hemorrhage: a meta-analysis. Biomed Res. 2018;2018:6038193.
21. Hevesi M, Bershad EM, Jafari M, Mayer SA, Selim M, Suarez JI, et al. Untreated hypertension as predictor of in-hospital mortality in intracerebral hemorrhage: a multi-center study. J Crit Care. 2018;43:235–9.
22. Divani AA, Hevesi M, Pulivarthi S, Luo X, Souslian F, Suarez JI, et al. Predictors of nosocomial pneumonia in intracerebral hemorrhage patients: a multi-center observational study. Neurocrit Care. 2015;22:234–42.
23. Hemphill JC 3rd, Bonovich DC, Besmertis L, Manley GT, Johnston SC. The ICH score: a simple, reliable grading scale for intracerebral hemorrhage. Stroke. 2001;32:891–7.
24. Poplawski NJ, Shirinifard A, Agero U, Gens JS, Swat M, Glazier JA. Front instabilities and invasiveness of simulated 3D avascular tumors. PLoS ONE. 2010;5:e10641.
25. Lee JH, Park CM, Park SJ, Bae JS, Lee SM, Goo JM. Value of computerized 3D shape analysis in differentiating encapsulated from invasive thymomas. PLoS ONE. 2015;10:e0126175.
26. Liao CC, Chen YF, Xiao F. Brain midline shift measurement and its automation: a review of techniques and algorithms. Int J Biomed Imaging. 2018;2018:4303161.
27. Hummel M, Edelmann D, Kopp-Schneider A. Clustering of samples and variables with mixed-type data. PLoS ONE. 2017;12:e0188274.
28. Harrell FE. Regression modeling strategies with applications to linear models, logistic regression and survival analysis. Stat Med 2003; 22:2531.
29. Kuhn M, Johnson K. Applied predictive modeling. New York: Springer; 2013.
30. Tian TS. Functional data analysis in brain imaging studies. Front Psychol. 2010;1:35.
31. Petersen A, Zhao J, Carmichael O, Muller HG. Quantifying individual brain connectivity with functional principal component analysis for networks. Brain Connect. 2016;6:540–7.
32. Ramsay JO, Silverman BW. Functional data analysis (springer series in statistics). 2nd ed. New York: Springer; 2005.
33. Petersen A, Muller H-G. Functional data analysis for density functions by transformation to a Hilbert space. Ann Stat. 2016;44:183–218.
34. Al-Mufti F, Thabet AM, Singh T, El-Ghanem M, Amuluru K, Gandhi CD. Clinical and radiographic predictors of intracerebral hemorrhage outcome. Interv Neurol. 2018;7:118–36.
35. Safatli DA, Gunther A, Schlattmann P, Schwarz F, Kalff R, Ewald C. Predictors of 30-day mortality in patients with spontaneous primary intracerebral hemorrhage. Surg Neurol Int. 2016;7:2152–7806.
36. Flemming KD, Wijdicks EF, Li H. Can we predict poor outcome at presentation in patients with lobar hemorrhage? Cerebrovasc Dis. 2001;11:183–9.
37. Nilsson OG, Lindgren A, Brandt L, Saveland H. Prediction of death in patients with primary intracerebral hemorrhage: a prospective study of a defined population. J Neurosurg. 2002;97:531–6.
38. Wang CW, Liu YJ, Lee YH, Hueng DY, Fan HC, Yang FC, et al. Hematoma shape, hematoma size, Glasgow coma scale score and ICH score: which predicts the 30-day mortality better for intracerebral hematoma? PLoS ONE. 2014;9:e102326.
39. Barras CD, Tress BM, Christensen S, MacGregor L, Collins M, Desmond PM, et al. Density and shape as CT predictors of intracerebral hemorrhage growth. Stroke. 2009;40:1325–31.
40. Nowinski WL, Gomolka RS, Qian G, Gupta V, Ullman NL, Hanley DF. Characterization of intraventricular and intracerebral hematomas in non-contrast CT. Neuroradiol J. 2014;27:299–315.
41. Hanley DF. Intraventricular hemorrhage and ICH outcomes: severity factor and treatment target. Stroke J Cereb Circ. 2009;40:1533–8.

42. Gregorio T, Pipa S, Cavaleiro P, Atanasio G, Albuquerque I, Chaves PC, et al. Prognostic models for intracerebral hemorrhage: systematic review and meta-analysis. BMC Med Res Methodol. 2018;18:145.

43. Bharath K, Kurtek S, Rao A, Baladandayuthapani V. Radiologic image-based statistical shape analysis of brain tumours. J R Stat Soc Ser C Appl Stat. 2018;67:1357–78.

44. Goldberg-Zimring D, Talos IF, Bhagwat JG, Haker SJ, Black PM, Zou KH. Statistical validation of brain tumor shape approximation via spherical harmonics for image-guided neurosurgery. Acad Radiol. 2005;12:459–66.

45. Crawford L, Monod A, Chen A, Mukherjee S, Rabadan R. Functional data analysis using a topological summary statistic: the smooth Euler characteristic transform. 2016. arXiv:161106818.