Kumar *et al.*
BMC Medical Informatics and Decision Making (2022) 22:307
https://doi.org/10.1186/s12911-022-02046-7

BMC Medical Informatics and Decision Making

RESEARCH Open Access

Choice of intraoperative ultrasound adjuncts for brain tumor surgery

Manoj Kumar[1*], Santosh Noronha[2], Narayan Rangaraj[1], Aliasgar Moiyadi[3,4], Prakash Shetty[3,4] and Vikas Kumar Singh[3,4]

Abstract

Background: Gliomas are among the most typical brain tumors tackled by neurosurgeons. During navigation for surgery of glioma brain tumors, preoperatively acquired static images may not be accurate due to shifts. Surgeons use intraoperative imaging technologies (2-Dimensional and navigated 3-Dimensional ultrasound) to assess and guide resections. This paper aims to precisely capture the importance of preoperative parameters to decide which type of ultrasound to be used for a particular surgery.

Methods: This paper proposes two bagging algorithms considering base classifier logistic regression and random forest. These algorithms are trained on different subsets of the original data set. The goodness of fit of Logistic regression-based bagging algorithms is established using hypothesis testing. Furthermore, the performance measures for random-forest-based bagging algorithms used are AUC under ROC and AUC under the precision-recall curve. We also present a composite model without compromising the explainability of the models.

Results: These models were trained on the data of 350 patients who have undergone brain surgery from 2015 to 2020. The hypothesis test shows that a single parameter is sufficient instead of all three dimensions related to the tumor ($p < 0.05$). We observed that the choice of intraoperative ultrasound depends on the surgeon making a choice, and years of experience of the surgeon could be a surrogate for this dependence.

Conclusion: This study suggests that neurosurgeons may not need to focus on a large set of preoperative parameters in order to decide on ultrasound. Moreover, it personalizes the use of a particular ultrasound option in surgery. This approach could potentially lead to better resource management and help healthcare institutions improve their decisions to make the surgery more effective.

Keywords: Brain cancer surgery, Medical decision making, Logistic regression, Random forest classifier, Intraoperative adjuncts, Bootstrap sampling

Background

Gliomas are among the commonest brain tumors encountered by neurosurgeons. Surgery is an integral component of its treatment, and the extent of resection is a crucial prognostic factor. The advancements in basic sciences and the availability of sophisticated technological surgical aids have led to the rise of innovative surgical strategies meant to profoundly impact the outcome of patients diagnosed with these aggressive tumors, which can show very different radiological patterns depending on their WHO grade and therefore pose different challenges in terms of surgical excision [1]. Due to the ill-defined nature of these tumors, surgeons increasingly rely on technological adjuncts to identify and remove maximum tumor safely. Navigation or frame-less

*Correspondence: manojiitbpowai@iitb.ac.in

[1] Department of Industrial Engineering and Operations Research, Indian Institute of Technology Bombay, Mumbai, India
Full list of author information is available at the end of the article

stereotaxy is the standard tool based on preoperatively acquired MR images which acts like a GPS providing a road map for the surgical procedure. However, due to the changes in anatomical structures during surgery, its accuracy is compromised, necessitating updated intraoperative imaging.

Tumor surgery has two main stages - lesion localization, and control of resection during surgery [2]. State-of-the-art intraoperative micro-surgical techniques are supplemented by techniques such as Magnetic Resonance Imaging (MRI), CT, and ultrasonography to improve the real-time updates [3].

Although intraoperative MRI (iMRI) would be ideal, it is not widely available and is costly. Intraoperative ultrasound (US) has emerged as a very popular solution [4], both 2D US (standard default US modality) as well as advanced navigated 3D US (3DUS). 3DUS provides navigated multiplanar images often with fusion imaging with preoperative MR and is believed to provide better orientation and image interpretation, thereby making it a viable option.

However, the navigated 3D ultrasound costs more as one scan takes 3–5 min (1–2 min for 2D), and the setup cost is higher than 2D ultrasound. Thus there is a trade-off between these imaging technologies, and one of the objectives in this paper is to analyze how neurosurgeons choose to deploy these two modalities during routine practice.

Different studies have shown that the resolution of ultrasound images deteriorates as the surgery progresses [2, 5]. Thus, navigated 3D ultrasound might not be beneficial in later stages.

In this study, we attempt to understand the preoperative factors that affect the choice of ultrasound.

Objective of the study

This study proposes a data-driven optimal decision policy based on patients and tumor characteristics. We have investigated the following research questions:

1. What attributes of the patient and the tumor affect the choice of ultrasound?
2. Does the experience of the surgeon affect the decision?

While assessing the benefits of certain technical adjuncts in healthcare, it is important to understand the patterns of use during routine care, which may differ from those under controlled trial conditions. Routine practices reflect day-to-day factors which are often difficult to pinpoint in preliminary observation. These factors need to be better understood to make conscious and well-informed decisions regarding the deployment of such health care technologies. This is more important if the eventual outcomes are affected by this choice or if there is a significant cost-consequence of these choices. In health care situations, it is often very difficult to test the effect of different states of the same factor (different types of techniques/adjuncts) due to practical and logistical difficulties. Using large databases and employing rigorous data science methods may be the best option.

The major conclusion is that contrast enhancement pattern, prior treatment, and surgeon's experience variables are statistically significant in most models. The patient's age is the only demographic factor that is statistically significant.

Literature survey

Surgical workflow analysis

The use of adjuncts needs to be understood in the context of surgical workflow. This workflow has different components, including low-level tasks, high-level tasks, patient status, and the use of medical devices. The low level activities described with the terms like *cut the skin with a scalpel* or *remove tissues with forceps* and high level tasks such as *skin incision made, skull opened* or *tumor tissue removal*. Considering the above, a study has been proposed to classify these situations based on multi-perceptive analysis [6]. Medical devices are developed stand-alone to provide specific functionality for a certain stage of the surgery. In [7], the authors have presented a model-driven design of surgical workflow to map the information of all these devices.

Surgeons need to make decisions about various tasks during surgical operations, called intraoperative decisions. Different situations and strategies in general are discussed in [8].

Glioma surgery

An automatic estimation method for brain tumor resection was developed in [9] based on the anatomical information received by the surgical navigation system using a Bayesian technique. The surgical navigation systems' stand-alone use fails to improve the outcome of brain tumor surgeries.

In the literature, many studies highlight the impact of intraoperative ultrasound for controlling the extent of resection of tumor tissue, for example [4, 10, 11]. A study [12] has been conducted to understand the applications and interaction between different modes of intraoperative imaging under the subjective basis of 11 surgical case studies. It highlights that iMRI is always the surgeon's choice, while it is evident from the study that the beneficial imaging modality is linear array intraoperative ultrasound.

Preoperative MRI features such as Eloquent Location, Sub-cortical Depth, Lobar vs non Lobar Glioma Location are considered to develop a predictive grading scheme model for surgical outcome in patients with glioblastoma multiforme [13]. Some of their limitations were highlighted in the study [14], which are selection bias, the premise of the study, finding that high-complexity lesions are significantly less likely to result in complete resection.

The authors in [15] showed that the superimposition of navigable 3D ultrasound with preoperative MRI provides a better orientation of the cross-sectional anatomy. Another study [2] showed that navigated 3D ultrasound without the preoperative images eliminates the registration inaccuracy inherent to image-to-patient registration algorithms. Another study has compared image-guided surgery with surgery being performed by either not using any image guidance or using two different forms of image guidance [3].

Statistical analysis and machine learning in healthcare

In medical decision-making, different statistical techniques have been widely used to improve the understanding of medical practitioners such as Logistic regression [16], principal component analysis [17, 18], and bootstrap sampling [16].

Many researchers have used machine learning algorithms in a variety of applications of healthcare such as diagnosis of a disease, prediction of survivability of a cancer patient, graft survival among kidney transplant recipients [19–24].

Convolutional neural networks have been proposed to diagnose gastric endoscopy-based gastric cancer, and they performed better than human pathologists [25].

A random forest and Cox proportional-hazard model has been developed to assess the association between contrast enhancement pattern of IHD mutant and diffuse glioma tumor with survival [26].

While several studies have been done to understand the parts of brain cancer workflow and the adjuncts used during surgery, benefits discussed in these studies are stand-alone and do not include the decision regarding the imaging modality to use in a particular patient case. Most importantly, none address the factors that influence the choice of using a particular technical modality. Our study integrating the above aspects would be novel and relevant for the field of brain cancer surgery. This could better inform neurosurgeons on selecting the most suitable modality a priori and potentially dictate decision making when identifying and inducting appropriate adjuncts in setting up a service.

Methods

Problem and data description

Intraoperative imaging technologies play a vital role in brain cancer surgery. Some of the possible technologies are intraoperative 2-Dimensional ultrasound (2DUS), navigated 3-Dimensional ultrasound (3DUS), and Magnetic Resonance Imaging (MRI). We try to identify factors that govern the choice of using the different types of intraoperative ultrasound based on the demographic factors of the patient, surgeon's experience, and tumor characteristics. These factors are known a priori and can be built into a decision-making algorithm during the preoperative stage allowing optimal allocation and utilization of resources as well as serving as a recommendation in different types of scenarios.

We also explore whether the surgeon's personal choice affects intraoperative 2DUS versus 3DUS decisions.

The data used in this analysis is secondary data collected from the electronic records of a tertiary care referral neurosurgical oncology centre. All patients undergoing resection for gliomas where intraoperative ultrasound was utilized and had preoperative MRI available for review during the time period 2015–2020 were analyzed. The use of anonymised retrospective data was approved for this study.

Clinical and radiological features based on preoperative routine MRI were extracted. The attributes of interest included patient's age, gender, prior treatment status (yes/no), eloquent location (yes/no), depth of tumor (surfacing/sub-cortical/deep), histology (high grade/low grade), glioma location (lobar/no-lobar), delineation (good/moderate/poor), contrast enhancement pattern (negligible/mixed/predominant), tumor dimensions in three orthogonal planes (height, length, width), and surgeon experience. Additionally, a variable spherical diameter was computed using the volume equivalent spherical diameter using MRI height, length, and width of the tumor.

We have included 350 procedures, out of which 2D ultrasound was used for 143 surgeries. Out of these three values were missing for contrast enhancement patterns, these were imputed using mode value.

In this data set, four surgeons have performed all the surgeries. The number of surgeries accomplished by a surgeon is taken as the surgeon's experience. The average (SD) age of patients is 41.23 (14.71) years. Appropriate correlation methods measuring the association between the variables were applied [27] and are shown in the attached Additional file 1.

The correlation among the tumor's length, height, and width is significant, and all other variables showed negligible correlation.

Table 1 Description of complete data set

	2D (n = 143) Mean (SD)	3DUS (n = 207)	*p*-value
Age (in years)	40.32 (15.72)	41.86 (13.97)	0.35
Length (in cm)	4.81 (1.59)	4.98 (1.65)	0.21
Width (in cm)	3.91 (1.18)	3.83 (1.14)	0.53
Height (in cm)	4.03 (1.30)	4.15 (1.33)	0.34
Surgeon experience (no. of surgeries)	143.17 (64.65)	184.23 (55.20)	0.00*
Gender			
Male	97	149	0.4
Female	46	58	
Prior treatment			
Yes	45	36	
No	98	171	0.00*
Eloquent location			
Yes	55	98	
No	88	109	0.1
Depth of tumor			
Surfacing	68	113	
Sub-cortical	41	49	
Deep	34	45	0.26
Histology			
Low grade	31	49	
High grade	112	158	0.66
Glioma location			
Lobar	135	203	0.06
No-Lobar	8	4	
Delineation			
Poor	10	23	
Moderate	80	102	
Good	53	82	0.97
Contrast enhancement pattern			
Negligible	33	78	
Mixed	88	80	
Predominant	22	49	0.30

*$p < 0.05$; *p*-value are corresponding to Mann–Whitney test

Statistical analysis

We have performed both parametric (*t*-test) and non-parametric (Mann–Whitney test) tests on the data sets to confirm the normality of the data. We have presented the *p*-values corresponding to the Mann–Whitney test here as we obtained the same result from both methods. The hypothesis tested is that both groups of technologies result in the same mean/proportions for the variables listed. Table 1 depicts the description of the complete data set with *p*-values for the hypotheses designed above.

Surgeon experience and prior treatment status are statistically significant in both groups. The most experienced surgeon has used the navigated 3D ultrasound more often. The average age of the patients, length, and height are more in 3DUS group but not statistically significant.

Table 2 Description of surgeon group 1 data

	2D (n = 58) Mean (SD)	3DUS (n = 156)	*p*-value
Age (in years)	43.79 (16.64)	41.42 (13.77)	0.43
Gender			
Male	39	114	
Female	19	42	0.4
Prior treatment			
Yes	17	27	
No	41	129	0.05
Eloquent location			
Yes	25	79	
No	33	77	0.32
Depth of tumor			
Surfacing	26	88	
Sub-cortical	18	36	0.18
Deep	14	32	
Histology			
Low grade	12	34	
High grade	46	122	0.86
Glioma location			
Lobar	54	153	
No-Lobar	4	3	0.07
Delineation			
Poor	3	16	
Moderate	32	80	0.59
Good	23	60	
Contrast enhancement pattern			
Negligible	12	63	
Mixed	36	54	0.22
Predominant	10	39	

p-values are corresponding to Mann–Whitney test

The data set is stratified into two groups—surgeon group 1, which includes the patients whose surgery was performed by the most experienced surgeon, and surgeon group 2 that consists of the patients whose surgeries were performed by three other surgeons.

Surgeon group 1 has performed 214 surgeries, out of which 58 (27%) surgeries are with 2DUS. Table 2 depicts the description of surgeon group 1 data set. The *t*-test and Mann–Whitney tests showed that none of the attributes are statistically significant.

Surgeon group 2 has performed 136 surgeries, out of which 51 (37.5%) surgeries are with navigated 3D ultrasound. Mann–Whitney test shows that none of the parameters (taken one at a time) are statistically

Table 3 Description of surgeons group 2 data set

	2D (n = 85) Mean (SD)	3DUS (n = 51)	*p*-value
Age (in years)	37.95 (14.70)	43.23 (14.64)	0.07
Gender			
Male	58	35	
Female	27	16	0.96
Prior treatment			
Yes	28	9	
No	57	42	0.05
Eloquent location			
Yes	30	19	
No	55	32	0.82
Depth of tumor			
Surfacing	42	25	
Sub-cortical	23	13	0.89
Deep	20	13	
Histology			
Low grade	19	15	
High grade	66	36	0.36
Glioma location			
Lobar	81	50	
No-lobar	4	1	0.41
Delineation			
Poor	7	7	
Moderate	48	22	0.69
Good	30	22	
Contrast enhancement pattern			
Negligible	21	15	
Mixed	52	26	0.99
Predominant	12	10	

p-values are corresponding to Mann–Whitney test

significant except prior treatment, which is at borderline, as shown in Table 3.

Thus, the major difference between surgeon groups 1 and 2 is that the former had more navigated 3D procedures, and the latter had more 2D ultrasound ones.

Methodology

We have designed two bootstrap cum aggregation (bagging) algorithms with logistic regression, and random forest as base (weak) classifiers [28, 29]. The non-parametric bootstrap sampling technique [16] was used for generating different learning set from the data set. The bagging algorithms are an aggregation of weak classifiers trained on bootstrap samples. We aggregated the final prediction by averaging the predicted probabilities of each of the weak classifiers.

Data analysis

Both bagging algorithms have been trained on the complete data set, on various subsets, and after dimensional reduction of the data set. We have also combined some levels of ordinal features and trained the logistic regression and random forest classifier to develop a composite model. All these data sets were standardized to standard normal distribution beforehand. All the data sets are divided randomly in the training set (80%) and testing test (20%), and models were trained on bootstraps samples drawn from training data sets.

Actual data set analysis

In this section, we have discussed the models trained on the actual data set and their corresponding results.

Complete data set

The complete data set was used to generate 11000 bootstrap samples and logistic regression (referred as LR-full model) and random forest-based bagging (referred as RF-full model) were trained.

Odds ratio (OR) of patient's gender, surgeon's experience, eloquent location, length, height, and glioma location are greater than one, and for all other variables, they are less than one. OR > 1 indicates the likelihood of navigated 3D is higher as compared to 2D ultrasound, and OR < 1 indicates a decrease in the likelihood of navigated 3D ultrasound.

Complete data set with spherical diameter

Brain tumor shape was approximated using spherical harmonics, which is defined by the orthogonal basis of functions over unit sphere for image-guided surgery [30]. In this data set, instead of height, length, and width, we have introduced a new parameter 'spherical diameter' which is computed from equivalent spherical volume ($\sqrt[3]{\frac{6}{\pi} \times \text{length} \times \text{height} \times \text{width}}$). The logistic regression (LR-Spherical model) and random forest (RF-Spherical model) based bagging algorithm were trained on this dataset.

The spherical diameter variable is constructed because there is a high correlation between the tumor's length, height, and width (tumors are relatively unlikely to grow along only one dimension). A χ^2-test shows that the hypothesis that the spherical diameter is sufficient to capture the information of the three-dimension-related variables cannot be ruled out ($p = 0.763 > 0.05$). This allows us to make the models more compact.

The surgeon experience variable is a most important feature in RF-full and RF-Spherical shown in Fig. 1 and has a larger coefficient in LR-full and LR-Spherical model as well. Also, its coefficient is statistically significant in both the models defined earlier. We have performed the

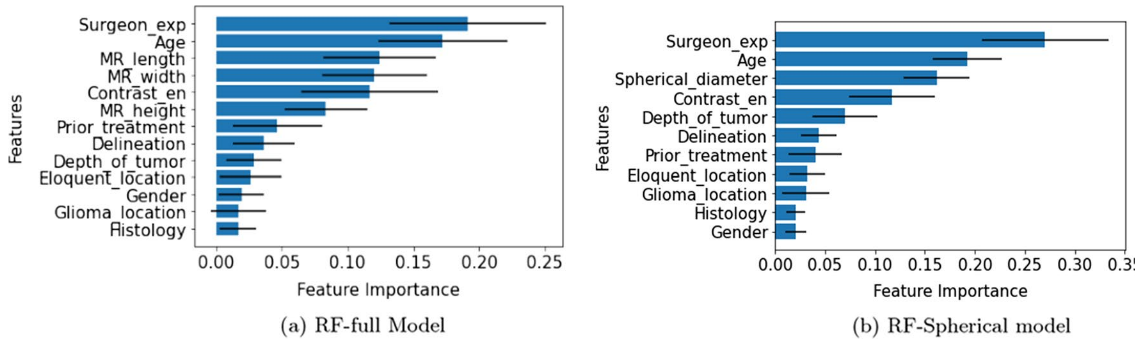

Fig. 1 Feature Importance for random forest model: the bars show the 95% confidence interval centered on the mean value. Surgeon_exp represents surgeon experience and contrast_en represents contrast enhancement pattern

analysis after excluding it to claim that this is indeed an important factor.

Data set with surgeon experience removed

We have removed the surgeon's experience from the complete data set and randomly divided the data set into training (80%) and testing set (20%). The bootstrap samples were drawn from the training set. The random forest (RF-Surgeon's Experience Removed) and logistic regression-based bagging algorithm (LR-Surgeon's Experience Removed) were trained on each of these bootstrap samples.

We have observed that the performance of both bagging algorithms worsened after dropping the surgeon experience feature.

The χ^2-test showed that LR-Spherical model and LR-Surgeon's Experience Removed are statistically different ($p = 0.005 < 0.05$). Therefore, removing the surgeon's experience from the model increases the deviance of the model and thus degrades the performance of the model. Hence surgeon's experience is an important factor in the choice of intraoperative ultrasound.

Surgeon based stratification

The complete data set is stratified into two groups based on the surgeons who have executed the surgeries. The details are discussed in the Tables 2 and 3.

Surgeon group 1 has 27% 2DUS samples, and surgeon group 2 has 37.5% 3DUS samples. Class imbalance is when one class has more elements than another in the data set, which biases predictive models towards the majority class. To prevent this, we have used the Synthetic Minority Oversampling Technique (SMOTE) [31]. SMOTE over-samples the minority class using k nearest neighbors technique.

In surgeon group 1, we have considered the value of k as six. The balanced data set of surgeon group 1 was divided into training and test set with an 80:20 ratio. The logistic regression (LR-Surgeon 1 group) and random forest-based bagging algorithm (RF-Surgeon 1 group) were trained on 11500 bootstrap samples from the training set.

In surgeon group 2, we have over-sampled the navigated 3D ultrasound class using k is equal to 3 and divided the data into 80:20 ratio for training and testing set. The logistic regression (LR-Surgeon 2 group) and random forest-based bagging algorithm (RF-Surgeon 2 group) were trained on 11500 bootstrap samples of size equal to the training set.

The prior treatment is statistically significant in LR-full ($p = 0.008$), LR-Surgeon's Experience Removed model ($p = 0.023$), LR-Surgeon 1 group model ($p = 0.007$), and at border line in LR-Spherical model. The width is statistically significant in LR-full model ($p = 0.04$). The LR-Surgeon 1 showed that patient's age is statistically significant as $p = 0.011$, whereas histology ($p = 0.018$) in LR-Surgeon 2 model.

Table 4 shows the likelihood of choice of ultrasound-based on the odds ratios corresponding to all logistic regression models discussed so far.

Analysis after redesigning of some parameters

The above models include the patient as well as tumor characteristics. It is very unlikely that patient's age and gender would influence the choice of the US being used. Therefore, we will consider only the tumor characteristics visible to surgeons before starting the surgery in further models, but after redefining some of them, such as contrast enhancement pattern, delineation, and location of tumor.

The contrast enhancement pattern is redefined as follows: 'predominant' + 'mixed' is taken as enhancing, and 'negligible' is taken as 'non-enhancing'.

Delineation is redefined as a dichotomous variable in two distinct ways:

Table 4 Likelihood of ultrasound of different logistic regression based bagging models

Name	LR Full	LR-Spherical	LR-Surgeon's Exp	LR-Surgeon 1	LR-Surgeon 2
Age	2D	3DUS	3DUS	2D	3DUS
Gender	3DUS	3DUS	3DUS	2D	2D
Surgeon's Exp	3DUS	3DUS	–	–	–
Prior treatment	2D	2D	2D	2D	2D
Contrast enhancement pattern	2D	2D	2D	2D	2D
Delineation	2D	2D	2D	2D	2D
Eloquent location	3DUS	3DUS	3DUS	3DUS	2D
Histology	2D	2D	2D	3DUS	2D
Depth of tumor	2D	2D	2D	2D	2D
Length	3DUS	–	–	–	–
Width	2D	–	–	–	–
Height	3DUS	–	–	–	–
Glioma location	3DUS	3DUS	3DUS	3DUS	-
Spherical diameter	–	3DUS	2D	3DUS	2D

2D represents the 2D ultrasound whereas 3DUS represents the navigated 3D ultrasound. – represents the exclusion of that variable in the model

1. Moderate grouped with poor delineation, and good delineation kept separate (denoted as PMD)
2. Moderate grouped with good, and poor delineation kept separate (denoted as GMD)

This is because the definition of moderate may be subjective, whereas poor and good delineations are more easily and reproducibly defined. The location of tumor was also redefined by coupling the depth of the tumor and height of the tumor as follows:

1. We club 'surfacing' and 'sub-cortical' tumors as 'surfacing' and assign all of them as fixed 'surface depth' value of 0.5 cm as these were defined as less than 1 cm. The 'surface depth' of deep tumors is considered as 1 cm.
2. Then we use the height value of each tumor, take its midpoint and add it to 'surface depth' (which is 0.5 or 1) to get the epicenter depth of the tumor.
3. Then we have defined the new variable location of tumor as 'superficial' if epicenter depth is less than 3 cm and otherwise 'deep' tumor.

A threshold of 3 cm is reasonable as total depth of the brain practically is observed to be 5–6 cm. The statistical analysis of all these variables showed that the prior treatment and contrast enhancement pattern are statistically significant (see the attached Additional file 1).

.

Based on this, we have constructed the following data set after suitable changes.

A. *Complete data set with PMD* The logistic regression and random forest models trained on this data set are referred as LR-PMD Spherical and RF-PMD Spherical.
B. *Surgeon Stratified data sets with PMD* The logistic regression and random forest models trained on these data sets are referred as LR-PMD Spherical Surgeon 1, LR-PMD Spherical Surgeon 2, RF-PMD Spherical Surgeon 1, and RF-PMD Spherical Surgeon 2.
C. *Complete data set with GMD* The logistic regression and random forest models trained on this data set are referred as LR-GMD Spherical and RF-GMD Spherical.
D. *Surgeon stratified data sets with GMD* The logistic regression and random forest models trained on these data sets are referred as LR-GMD Spherical Surgeon 1, LR-GMD Spherical Surgeon 2, RF-GMD Spherical Surgeon 1, and RF-GMD Spherical Surgeon 2.

In (A) and (B), we have only included the tumor characteristics with delineation defined as PMD, whereas in (C) and (D), delineation is defined as GMD. We have trained different logistic regression and random forest models on these data sets. Table 5 shows the likelihood of choice of intraoperative ultrasound in different models.

We have compared all logistic regression models using the chi-square test as depicted in Table 6. This table summarizes the important models that lead to our conclusions.

Our composite models with redesigned variables showed performance comparable to the LR-spherical model. Hence, a decision may be taken with fewer parameters instead of all patient and tumor characteristics. This can also be concluded from the surgeon group 1 and surgeon group 2 data.

Table 7 shows the performance of all the models. RF-Surgeon 1 model resulted in all performance measures (accuracy, AUC ROC score, and AUC PR being more than 80%. All except LR-Surgeon Experience removed, and RF-Surgeon Experience removed models resulted in the AUC ROC score of more than 70%. The AUC ROC score 0.7–0.8 is considered acceptable, 0.8–0.9 is considered excellent, and more than 0.9 is considered outstanding [32]. Hence all our models expect LR-Surgeon's Experience removed and RF-Surgeon's Experience removed are acceptable.

Discussion

1. All models discussed in Table 4 except LR-full and LR-Surgeon 1 group model favoring the navigated 3D ultrasound more likely when a patient is older.
2. All the logistic regression models discussed in Tables 4 and 5 agree that choice of 2D ultrasound is more likely when prior treatment is 'yes' or contrast enhancement pattern is 'enhancing'. As discussed with medical practitioners, 2D ultrasound is used either to localize the tumor or for a confirmatory scan whenever any prior treatment is done. Also, 'enhancing' tumors are clearly visible to surgeons; therefore, 2D ultrasound is enough. Wherever the tumor is in eloquent areas or the location is deep, all models recommend using navigated 3D ultrasound as the surgeon's focus is to prevent damage to eloquent areas while achieving maximal possible resec-

Table 5 Likelihood of ultrasound of logistic regression based bagging models

Name	LR-PMD Spherical	LR-PMD Surgeon 1	LR-PMD Surgeon 2	LR-GMD Spherical	LR-GMD Surgeon 1	LR-GMD Surgeon 2
Prior treatment	2D	2D	2D	2D	2D	2D
Contrast enhancement pattern	2D	2D	2D	2D	2D	2D
Delineation	3DUS	2D	3DUS	2D	2D	2D
Eloquent location	3DUS	3DUS	3DUS	3DUS	3DUS	3DUS
Histology	2D	3DUS	2D	2D	3DUS	2D
Location	3DUS	3DUS	3DUS	3DUS	3DUS	3DUS
Spherical diameter	2D	3DUS	2D	2D	3DUS	2D

Table 6 Statistical analysis of different models

	Full data set				
	Model	**Deviance**	**Degree of freedom**	**Model comapred**	***P* value**
1	LR-full model	85.67	56		
2	LR-Spherical model	85.14	58	1–2	0.763
3	LR-Surgeon's Experience removed	92.78	59	1–3	0.068
				2–3	**0.005**
4	LR-GMD Spherical	87.17	62	2–4	0.845
5	LR-PMD Spherical	86.62	62	2–5	0.915
	Surgeon Group 1 data set				
6	LR-Surgeon 1	77.18	52		
7	LR-GMD Spherical Surgeon 1	77.48	55	6–7	0.960
8	LR-PMD Spherical Surgeon 1	76.04	55	6–8	0.767
	Surgeon Group 2 data set				
9	LR-Surgeon 2	40.21	23		
10	LR-GMD Spherical Surgeon 2	39.40	26	9–10	0.847
11	LR-PMD Spherical Surgeon 2	41.06	26	9–11	0.837

Bold values are < significance level 0.05 (i.e. $p < 0.05$)

Table 7 Performance of all models

Model	Accuracy (%)	AUC ROC (%)	AUC PR (%)
LR Full	70	70	75
RF Full	66	72	79
LR Spherical	66	74	82
RF Spherical	69	72	75
LR-Surgeon's Exp removed	56	63	74
RF-Surgeon's Exp removed	57	64	73
LR-Surgeon 1 group	70	76	72
RF-Surgeon 1 group	81	89	97
LR-Surgeon 2 group	71	79	80
RF-Surgeon 2 group	74	80	82
LR-PMD Spherical	64	71	77
RF-PMD Spherical	64	77	83
LR-PMD Spherical Surgeon 1	70	77	75
RF-PMD Spherical Surgeon 1	71	78	76
LR-PMD Spherical Surgeon 2	62	76	77
RF-PMD Spherical Surgeon 2	59	70	71
LR-GMD Spherical	64	70	76
RF-GMD Spherical	64	80	85
LR-GMD Spherical Surgeon 1	65	78	75
RF-GMD Spherical Surgeon 1	71	72	70

AUC ROC denotes area under the receiver operating characteristic curve whereas AUC PR denotes area under the precision-recall curve

tion. In cases, 2D ultrasound may not provide sufficient information about deeply situated tumors, then navigated 3D ultrasound is preferred.

3. The surgeon 1 model elaborated in Table 4 suggests the use of navigated 3D ultrasound for large spherical diameter tumors. In contrast, all other models trained on PMD and GMD data sets suggest the use of 2D ultrasound.
4. Except LR-PMD surgeon 2 and LR-PMD spherical, all other models discussed in Table 5 suggest the use of 2D ultrasound whenever delineation is good.
5. LR-PMD Spherical and LR-GMD Spherical models discussed in Table 5 have the same sign coefficient except for the delineation, which is defined in various ways. The random forest trained on these data sets also exhibits a different order of feature importance. Therefore, how surgeons interpret the moderate delineation is also an essential factor in deciding the type of intraoperative ultrasound.
6. The random forest-based models trained on surgeon's stratified datasets resulted spherical diameter as a most important factor in the models. Also, it is observed that feature importance is distinct for different models, which may be due to the surgeon's personal choice that they would have for 2D or navigated 3D ultrasound.
7. We have also applied Principal Component Analysis (PCA) on complete data set with spherical diameter and found that it does not provide us a model with fewer dimensions that could explain intraoperative ultrasound decisions. However, after redesigning a few ordinal variables without compromising explainability, a more compact model with fewer features was obtained. This perhaps overcomes the limitation of the Principal Component Analysis method.

Contrast enhanced ultrasound (CEUS) is a rapidly evolving US technique which employs nano bubbles and harmonic imaging to produce contrast images. This reflects tissue perfusion and is different from MR contrast enhancement in gliomas which is a function of tissue permeability and extravasation of contrast from a leaky blood brain barrier. However, CEUS is fast emerging as a useful adjunct to standard B Mode US. We did not use CEUS and hence our results may be taken in this context only. we agree that in the future they may need to be revisited as more evidence accumulates. The use of CEUS for brain cancer surgery can be found elsewhere [33].

Contrast enhancement may not be important in decision-making in low grade gliomas [14], which may be because a majority of low grade gliomas are non-enhancing. However, our pool of cases was a mix and since there were many high grade tumors, enhancement did show some significance (as shown in the Additional file 1). It also corroborates the study [34], which showed the importance of intraoperative ultrasonography for resections of non-enhancing tumors. It should be kept in mind that histology is usually not available at the time of surgical planning and the surgeon has to rely on available parameters and information. In such a case, contrast enhancement is especially valuable as a surrogate marker of tumor grade, and for decision making.

Conclusion

In this paper, we have attempted to examine the factors that could have influenced the choice of use of a particular intraoperative imaging adjunct (US) in a large series of patients consecutively treated at a reference neurosurgery centre. Different logistic regression and random forest-based bagging models were fitted over the various data sets generated from a data set of 350 patients and tested on the test data sets. We found that the surgeon experience, prior treatment, and contrast enhancement pattern are statistically significant in almost all logistic regression-based models. The models trained on the surgeon's stratified data sets show that patients' age and histology are also statistically significant.

The random forest-based bagging model also showed that the surgeon experience and patient's age are the two most important factors. The spherical diameter of the tumor is the essential attribute after removing the surgeon experience parameter from the model. The random forest-based model trained on the surgeon's stratified group where only tumor characteristics are considered, depicts the distinct order of feature importance.

Logistic regression-based model highlights that likelihood of ultrasound type depends on how the delineation is considered. Therefore, we can say that different surgeons give different weightage to various features while selecting the intraoperative ultrasound. The models trained on the surgeon's stratified data sets show that the surgeon's personal choice affects the overall decision of intraoperative ultrasound.

We have introduced spherical diameter as a single parameter instead of three MRI measured dimensions of tumor. We found that one parameter, i.e., spherical diameter, is enough to capture the information of all three dimensions of the tumor. Tumor characteristics (delineation/prior treatment/contrast enhancement pattern/eloquent monitoring/histology/location) were found to be adequate to explain the decision irrespective of patient characteristics (age, gender), by and large. Only in one subset of data, age plays some role in the decision-making. The 2D ultrasound was used more likely for previously treated superficial and enhancing tumors situated in non-eloquent areas. The navigated 3D ultrasound was used for non-enhancing tumors situated in eloquent areas and deep inside the brain.

The limitation of our work is that the results reflect associations between tumor factors and the use of a particular US type, but this cannot be interpreted as recommendations for the use of such US type in those subsets. For that, outcome analysis and correlation are important. However, analyzing choice distribution is important to be able to account for surgeon choices in future outcome comparison studies.

Abbreviations
LR: Logistic regression; RF: Random forest; US: Ultrasound; 2D: 2-Dimensional; 3D: 3-Dimensional; CT: Computerized tomography; AUC ROC: Area under receiver operating characteristic curve.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12911-022-02046-7.

Additional file 1. Correlation matrices and details of the redesigned variables.

Acknowledgements
Not applicable.

Author contributions
MK has performed the analysis and wrote the draft version of the manuscript. SN, NR, and AM contributed to the conceptualization of the problem and reviewed the draft version. AM, PS, and VKS collected the data and helped in clinical inputs. All authors read and approved the final manuscript.

Funding
Not applicable.

Availability of data and materials
The data used in the study is not experimental data designed for scientific study per se, but is a retrospective study derived from actual clinical records in practice, with the approval of the institution. The data set is not publicly available as per institutional policy but is available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
Ethics approval has been obtained from the Institutional Ethics committee of Tata Memorial Centre, Mumbai (IEC 3). As per the institutional ethics committee guidelines, consent of participants was not required as this is a retrospective analysis of anonymized data. The study did not require any experiments to be performed and was performed in accordance with ICH-GCP guidelines.

Consent for publication
Consent for publication by participants is not applicable in this case.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Department of Industrial Engineering and Operations Research, Indian Institute of Technology Bombay, Mumbai, India. [2]Department of Chemical Engineering, Indian Institute of Technology Bombay, Mumbai, India. [3]Department of Neurosurgery, Tata Memorial Centre, Mumbai, India. [4]Department of Health Sciences, Homi Bhabha National Institute, Mumbai, India.

Received: 2 February 2022 Accepted: 14 November 2022
Published online: 28 November 2022

References

1. Ganau L, Paris M, Ligarotti G, Ganau M. Management of gliomas: overview of the latest technological advancements and related behavioral drawbacks. Behav Neurol. 2015. https://doi.org/10.1155/2015/862634.
2. Moiyadi AV, Shetty P. Direct navigated 3d ultrasound for resection of brain tumors: a useful tool for intraoperative image guidance. Neurosurg Focus. 2016;40(3):5.
3. Barone DG, Lawrie TA, Hart MG. Image guided surgery for the resection of brain tumours. Cochrane Database Syst Rev. 2014;(1).
4. Moiyadi A, Shetty P. Objective assessment of utility of intraoperative ultrasound in resection of central nervous system tumors: a cost-effective tool for intraoperative navigation in neurosurgery. J Neurosci Rural Pract. 2011;2(01):004–11.
5. Moiyadi AV. Intraoperative ultrasound technology in neuro-oncology practice-current role and future applications. World Neurosurg. 2016;93:81–93.
6. Franke S, Meixensberger J, Neumuth T. Multi-perspective workflow modeling for online surgical situation models. J Biomed Inform. 2015;54:158–66.
7. Liebmann P, Neumuth T. Model driven design of workflow schemata for the operating room of the future. INFORMATIK 2010. Service Science–Neue Perspektiven für die Informatik. Band. 2010;1.
8. Flin R, Youngson G, Yule S. How do surgeons make intraoperative decisions? BMJ Qual Saf. 2007;16(3):235–9.

9. Nakamura R, Aizawa T, Muragaki Y. Automatic surgical workflow estimation method for brain tumor resection using surgical navigation information. J Robot Mechatron. 2012;24(5).
10. Moiyadi A. Impact of navigable ultrasound: how much credit should we give to the "navigable" component? J Neurol Surg Part A Central Eur Neurosurg. 2015;76(02):177–8.
11. Moiyadi AV, Kannan S, Shetty P. Navigated intraoperative ultrasound for resection of gliomas: predictive value, influence on resection and survival. Neurol India. 2015;63(5):727.
12. Lothes TE, Siekmann M, König RW, Wirtz CR, Coburger J. Surgical workflow analysis: ideal application of navigated linear array ultrasound in low-grade glioma surgery. J Neurol Surg Part A Central Eur Neurosurg. 2016;77(06):466–73.
13. Marcus HJ, Williams S, Hughes-Hallett A, Camp SJ, Nandi D, Thorne L. Predicting surgical outcome in patients with glioblastoma multiforme using pre-operative magnetic resonance imaging: development and preliminary validation of a grading system. Neurosurg Rev. 2017;40(4):621–31.
14. Ganau L, Ligarotti GK, Ganau M. Predicting complexity of tumor removal and postoperative outcome in patients with high-grade gliomas. Neurosurg Rev. 2018;41(1):371–3.
15. Moiyadi AV, Shetty PM, Mahajan A, Udare A, Sridhar E. Usefulness of three-dimensional navigable intraoperative ultrasound in resection of brain tumors with a special emphasis on malignant gliomas. Acta Neurochir. 2013;155(12):2217–25.
16. Adjei IA, Karim R. An application of bootstrapping in logistic regression model. Open Access Libr J. 2016;3(9):1–9.
17. Abdi H, Williams LJ. Principal component analysis. Wiley Interdiscip Rev Comput Stat. 2010;2(4):433–59.
18. Wold S, Esbensen K, Geladi P. Principal component analysis. Chimometrics and intelligent laboratory systems. In: IEEE conference on emerging technologies and factory automation EFTA volume, 1987. p. 704–6.
19. Olsen CR, Mentz RJ, Anstrom KJ, Page D, Patel PA. Clinical applications of machine learning in the diagnosis, classification, and prediction of heart failure. Am Heart J. 2020;229:1–17.
20. Khaleel FA, Al-Bakry AM. Diagnosis of diabetes using machine learning algorithms. Mater Today Proc. 2021.
21. Bilgen I, Guvercin G, Rekik I. Machine learning methods for brain network classification: application to autism diagnosis using cortical morphological networks. J Neurosci Methods. 2020;343: 108799.
22. Piri S, Delen D, Liu T, Zolbanin HM. A data analytics approach to building a clinical decision support system for diabetic retinopathy: developing and deploying a model ensemble. Decis Support Syst. 2017;101:12–27.
23. Zolbanin HM, Delen D, Zadeh AH. Predicting overall survivability in comorbidity of cancers: a data mining approach. Decis Support Syst. 2015;74:150–61.
24. Topuz K, Zengul FD, Dag A, Almehmi A, Yildirim MB. Predicting graft survival among kidney transplant recipients: a Bayesian decision support model. Decis Support Syst. 2018;106:97–109.
25. Agarwal S, Yadav AS, Dinesh V, Vatsav KSS, Prakash KSS, Jaiswal S. By artificial intelligence algorithms and machine learning models to diagnosis cancer. Mater Today Proc. 2021.
26. Suchorska B, Schueller U, Biczok A, Lenski M, Albert NL, Giese A, Kreth F-W, Ertl-Wagner B, Tonn J-C, Ingrisch M. Contrast enhancement is a prognostic factor in IDH1/2 mutant, but not in wild-type who grade II/III glioma as confirmed by machine learning. Eur J Cancer. 2019;107:15–27.
27. Khamis H. Measures of association: how to choose? J Diagn Med Sonogr. 2008;24(3):155–62.
28. Breiman L. Bagging predictors. Mach Learn. 1996;24(2):123–40.
29. Li G-Z, Liu T-Y, Cheng VS. Classification of brain glioma by using SVMs bagging with feature selection. In: International workshop on data mining for biomedical applications. Springer; 2006. p. 124–30.
30. Goldberg-Zimring D, Talos I-F, 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(4):459–66.
31. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. J Artif Intell Res. 2002;16:321–57.
32. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. J Thorac Oncol. 2010;5(9):1315–6.
33. Ganau M, Ligarotti GK, Apostolopoulos V. Real-time intraoperative ultrasound in brain surgery: neuronavigation and use of contrast-enhanced image fusion. Quant Imaging Med Surg. 2019;9(3):350.
34. Moiyadi AV, Prakash Shetty RJ. Non-enhancing gliomas: does intraoperative ultrasonography improve resections? Ultrasonography. 2019;38(2):156.

Publisher's Note

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