PLOS ONE

RESEARCH ARTICLE

Magnetic resonance radiomics features and prognosticators in different molecular subtypes of pediatric Medulloblastoma

Feng-Chi Chang[1☯], Tai-Tong Wong[2,3☯], Kuo-Sheng Wu[2], Chia-Feng Lu[4], Ting-Wei Weng[5], Muh-Lii Liang[6], Chih-Chun Wu[1], Wan Yuo Guo[1], Cheng-Yu Chen[5,7], Kevin Li-Chun Hsieh[5,7]*

1 Department of Radiology, School of Medicine, Taipei Veterans General Hospital and National Yang Ming Chiao Tung University, Taipei, Taiwan, **2** Department of Neurosurgery, Taipei Medical University Hospital, Taipei Medical University, Taipei, Taiwan, **3** Department of Neurosurgery, Neurological Institute, School of Medicine, Taipei Veterans General Hospital and National Yang Ming Chiao Tung University, Taipei, Taiwan, **4** Department of Biomedical Imaging and Radiological Science, National Yang Ming Chiao Tung University, Taipei, Taiwan, **5** Department of Medical Imaging, Taipei Medical University Hospital, Taipei, Taiwan, **6** Department of Neurosurgery, Mackay Memorial Hospital, Taipei, Taiwan, **7** Department of Radiology, School of Medicine, College of Medicine, Taipei Medical University, Taipei, Taiwan

☯ These authors contributed equally to this work.
* kevinh9396@gmail.com

Citation: Chang F-C, Wong T-T, Wu K-S, Lu C-F, Weng T-W, Liang M-L, et al. (2021) Magnetic resonance radiomics features and prognosticators in different molecular subtypes of pediatric Medulloblastoma. PLoS ONE 16(7): e0255500. https://doi.org/10.1371/journal.pone.0255500

Editor: Joseph Najbauer, University of Pécs Medical School, HUNGARY

Received: February 10, 2021

Accepted: July 17, 2021

Published: July 29, 2021

Data Availability Statement: All relevant data are within the manuscript and its Supporting information files.

Funding: K.H received funding from Ministry of Science and Technology, Taiwan (MOST 106-2314-B-038-033) and the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan (DP2-110-21121-03-C-02-03). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

Purpose

Medulloblastoma (MB) is a highly malignant pediatric brain tumor. In the latest classification, medulloblastoma is divided into four distinct groups: wingless (WNT), sonic hedgehog (SHH), Group 3, and Group 4. We analyzed the magnetic resonance imaging radiomics features to find the imaging surrogates of the 4 molecular subgroups of MB.

Material and methods

Frozen tissue, imaging data, and clinical data of 38 patients with medulloblastoma were included from Taipei Medical University Hospital and Taipei Veterans General Hospital. Molecular clustering was performed based on the gene expression level of 22 subgroup-specific signature genes. A total 253 magnetic resonance imaging radiomic features were generated from each subject for comparison between different molecular subgroups.

Results

Our cohort consisted of 7 (18.4%) patients with WNT medulloblastoma, 12 (31.6%) with SHH tumor, 8 (21.1%) with Group 3 tumor, and 11 (28.9%) with Group 4 tumor. 8 radiomics gray-level co-occurrence matrix texture (GLCM) features were significantly different between 4 molecular subgroups of MB. In addition, for tumors with higher values in a gray-level run length matrix feature—Short Run Low Gray-Level Emphasis, patients have shorter survival times than patients with low values of this feature ($p = 0.04$). The receiver operating characteristic analysis revealed optimal performance of the preliminary prediction model

Competing interests: The authors have declared that no competing interests exist.

based on GLCM features for predicting WNT, Group 3, and Group 4 MB (area under the curve = 0.82, 0.72, and 0.78, respectively).

Conclusion

The preliminary result revealed that 8 contrast-enhanced T1-weighted imaging texture features were significantly different between 4 molecular subgroups of MB. Together with the prediction models, the radiomics features may provide suggestions for stratifying patients with MB into different risk groups.

Introduction

Medulloblastoma (MB) is the most common primary malignant brain tumor in children and is currently treated on the basis of pathological and clinicoradiological risk stratification [1]. Despite progress in disease management, the overall survival remains dismal with 5 year survival rate around 70~75% [2]. Current therapeutic options have debilitating effects on developing children, highlighting the need for molecularly targeted treatments with reduced toxicity. Recent developments in genome-wide sequencing techniques have improved understanding regarding MB. Many target genes with dysregulated expression and mutation have been identified as possible drivers of MB. Research conducted over the past years has identified four distinct molecular variants, namely wingless (WNT), sonic hedgehog (SHH), Group 3, and Group 4, each with different demographic, and genomic characteristics [3,4]. Improved knowledge of signaling pathways and important oncogenic drivers in these subgroups has elucidated the reasons for suboptimal clinical outcomes and therapeutic resistance. These findings have ushered in a new era of therapeutic optimism, with latest clinical trials now pursued on the basis of the molecular classification of MB. However, no consensus has been reached on which genomic analysis protocol should be applied to identify MB molecular subtypes. Moreover, the heterogeneous spatial distribution of genomic features makes unguided surgical biopsies prone to sampling errors, which may result in misclassification [5,6].

With the development of diagnostic imaging technologies, magnetic resonance imaging (MRI) is now a standard examination procedure employed for pediatric brain tumors. This is because MRI provides a wide range of physiologically meaningful contrasts for distinguishing different components through imaging, thereby improving depiction of heterogeneous patterns of tissue composition within tumors. Several reports have proved that MRI phenotype can be applied to predict the genetic profiling of brain tumors [7]. A study by Perreault et al. indicated that MRI substitute can be a favorable reference for MB molecular subgroups [8]. On the basis of objective radiomics analysis, imaging data can be transformed into a high-dimensional space by using an automatically high-throughput feature extraction algorithm [9]. Quantitative image-based features, including intensity, shape- and size-based, and textural features, can be obtained using this process. Some of these features have been proven to be associated with underlying gene-expression patterns [10].

In this study, multiparametric radiomics MRI analysis was conducted to reveal MB features. Since different MB subgroups have distinct patient demographics, clinical management and disease outcomes [11]. This non-invasive computer-aided quantitative diagnosis procedure may be included as part of the diagnostic workup.

Materials and methods

Data collection of patients with MB

This retrospective study was approved by the institutional review boards of Taipei Medical University Hospital and Taipei Veterans General Hospital. The original material and data were collected in compliance with all applicable laws, regulations, and policies for the protection of human participants. Informed consent was obtained from a parent and/or legal guardian for inclusion in the study. The clinical information, frozen tumor specimens, and preoperative magnetic resonance (MR) images of 52 patients with MB were obtained. 14 patients with incomplete clinical data, poor tumor tissue quality, and incomplete or poor-quality MR images were excluded from the research. Only 38 cases were finally included in our study.

RNA-seq and clustering analysis

For RNA-Seq analysis, total RNA was purified from MB tissues by using the RNeasy Mini Kit (Qiagen). The quality and quantity of RNA were assessed through spectrophotometry and capillary gel electrophoresis, respectively. Next, the stranded mRNA libraries were generated by using the TruSeq Stranded mRNA Library Prep Kit (Illumina). From each sample, 60 ng/μL of poly(A)-selected RNA was run in two lanes of a MiSeq sequencing instrument (Illumina) for 100 cycles of multiplexed paired-end reads. We also performed read pseudoalignment and quantification using Kallisto [12] with 200 bootstraps/paired-end reads aligned against the human genome assembly, GRCh38 transcriptome definition. We extracted the gene-expression table by using the tximport package in RStudio with R, and fragments per kilobase of transcript per million reads data sets were trimmed to obtain the protein coding genes, which was followed by quantile normalization and log2 transformation of each data set [13]. For clustering, we performed unsupervised clustering analysis with the consensus clustering default parameters by using the t-distributed stochastic neighbor embedding (t-SNE) and nonnegative matrix factorization (NMF) [14,15] packages, and validated by 22 subgroup-specific signature genes expression levels (WNT: WIF1, TNC, GAD1, DKK2, and EMX2; SHH: PDLIM3, EYA1, HHIP, ATOH1, and SFRP1; Group 3: IMPG2, GABRA5, NRL, MAB21L2, NPR3, and EYS [EGFL11]; and Group 4: KCNA1, EOMES, KHDRBS2, RBM24, UNC5D, and OAS1) [16]. RNA-Seq results were sent to Taylor's laboratory in the Hospital for Sick Children, Toronto in helping with counterpart clustering.

MR image analysis

The MR sequences used for analysis included the axial T1 weighted image (T1WI), axial T2 weighted image (T2WI), axial T2 fluid attenuation inversion recovery (FLAIR) image, contrast-enhanced axial T1WI (CET1), and diffusion weighted image. In total, 14 patients were excluded; 11 lacked preoperative MRI, and 3 had poor MRI quality. Thus, 38 patients were enrolled in our study for imaging analysis. Quantitative imaging feature analysis was performed as follows.

Image postprocessing and MR radiomics. To prevent variations in textural analytical features resulting from different scanning protocols [17], we applied several postprocessing steps on the MR images to decrease the discrepancy between imaging parameters used in different hospitals. All images were adjusted to voxel size $0.75 \times 0.75 \times 3.00$ mm^3 without gaps between consecutive slices for each MRI modality. The T2W, T2 FLAIR images, and apparent diffusion coefficient (ADC) maps derived from the diffusion-weighted image were then registered to the CET1s by using a six-parameter rigid body transformation and mutual

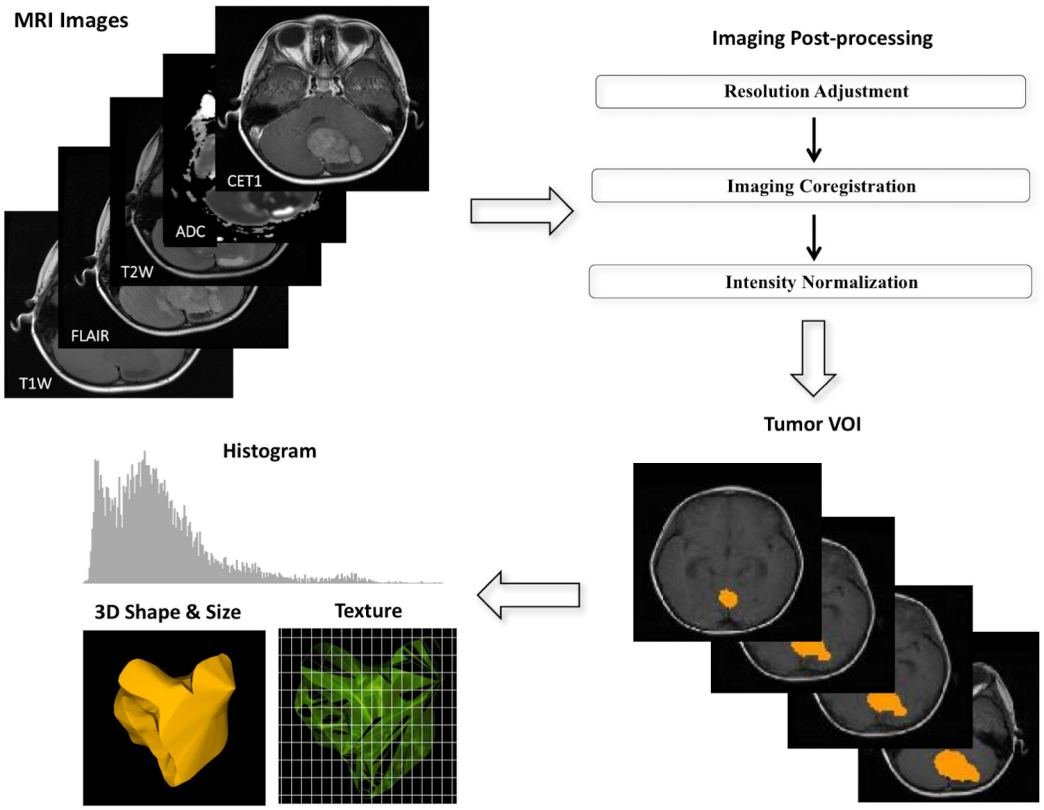

Fig 1. Diagram showing imaging process in radiomics. (a) Multiple postprocessing steps were applied to magnetic resonance images, including image coregistration, adjustment of image resolution, and intensity normalization between patients. (b) The volume of interest covering the total tumor was manually drawn by a neuroradiologist. (c) Three categories of radiomics, namely histogram, three-dimensional geography, and textural analysis, were employed on the processed magnetic resonance images to yield radiomic features.

https://doi.org/10.1371/journal.pone.0255500.g001

information algorithm. Image intensity normalization was employed to transform MRI intensity into standardized ranges among all subjects. The volume of interest (gross tumor volume) covering the total tumor volume was delineated by a board-certified neuroradiologist (K.H., with 14 years of experience) who was blinded to the molecular status of all tumors. Pixels included by the defined tumor contour were applied for feature extraction. A diagram illustrating image processing is displayed in Fig 1.

We extracted 57 radiomic image features that describe tumor characteristics based on Aerts's work [18]. These features can be divided into three categories: tumor intensity (16 features), shape and size (8 features), and texture (33 features). In the first category, tumor intensity information are quantified using first-order statistics, obtained from the histogram of entire tumor voxel intensity values. The second category are features based on the three-dimensional geometry of a tumor, including surface area, volume size, and derived features. The third category consists of three-dimensional texture features that are able to quantify the intratumoral heterogeneity within a full tumor volume. The 16 intensity and 33 textural features were extracted from T1WI, T2WI, T2 FLAIR, CET1, and ADC images to yield 245 features. The eight size and shape features were computed based on the three-dimensional morphology of the tumors. In total, a maximum of 253 MR radiomic features could be generated for each subject. All radiomic features were extracted by using the MR Radiomics

Platform (MRP; http://www.ym.edu.tw/~cflu/MRP_MLinglioma.html) which is in line with the regulations of Image Biomarker Standardization Initiative (IBSI) [19]. Detailed formulas of MR radiomic features are provided in Supplementary S1 Table.

Handling missing values. After calculation of radiomic features for all subjects, we further performed k-nearest neighbor (k-NN) imputation to compensate for the partially missing values caused by the lack of image sequences for some of the subjects [20]. The k-NN method is a common strategy in which missing values are imputed using values calculated from k nearest neighbors that reach a certain similarity measure. K = 3 was selected in this study since the smallest subgroup (WNT) in our cohort has only 7 cases [21]. The nearest, neighbors are determined based on Euclidean distance function. The superiority of the k-NN method over mean imputation is that the imputed values are influenced by only the most similar cases rather than by all values. After application of the feature extraction protocol in T1WI, T2WI, T2 FLAIR, CET1, ADC map with k-NN imputation, we extracted 253 radiomics features from each tumor.

Comparison of radiomics features in 4 molecular groups. The extracted 253 MR radiomic image features were compared to find the difference between the 4 molecular subtypes of medulloblastoma. Dunn's multiple comparisons test was applied and the significance level (alpha) is set at 0.05.

Analyzing the impact of imaging biomarkers on disease survival. We applied univariate and multivariate Cox proportional hazards modeling to investigate the relationships between quantitative radiomic features and overall survival. The Wald test was used to determine the significance of Cox models, and the most significant features were selected to define the radiomic risk subgroups. The log-rank (Mantel–Cox) test was used to compare the survival curves of different subgroups, and hazard ratios were used to report the direction of the survival effect. The significance level (alpha) is set at 0.05.

Establishing a prediction model of 4 molecular subtypes of MB. We tried to establish a prediction model to estimate the molecular subtypes of each MB based on different radiomics imaging features. Different feature selection algorithms including minimum redundancy maximum relevance, sequential backward elimination and sequential forward selection were applied to obtain the best future combination [22,23]. A support vector machine (SVM) implemented with nested leave one out cross validation was applied to find the best model [24,25]. The performance of models was scored using several metrics, namely sensitivity, specificity, and accuracy (ACC). However, a meaningful prediction model was hard to be created based on such a small cohort. The detailed methods and preliminary results were presented in supplement S1 Text, S2 and S3 Tables.

Results

Clinical characteristics and molecular subtypes of the study cohort

In total, 52 children with MBs were collected from the archives. We clustered these cases into 4 molecular subgroups by running unsupervised clustering analysis with NMF and t-SNE according to significantly different expressed genes (Fig 2). Furthermore, we validated the cluster using the 22 subgroup-specific signature gene list [16]. However, 14 patients were excluded because complete preoperative MR images were unavailable. Thus, 38 cases were included in our research, and the mean age of this cohort at diagnosis was 7.8 years. The final subgroups of WNT, SHH, Group 3, and Group 4 consisted of 7 (18.4%), 12 (31.6%), 8 (21.1%), and 11 (28.9%) patients, respectively. Clinical profiles of the included patients are listed in Table 1.

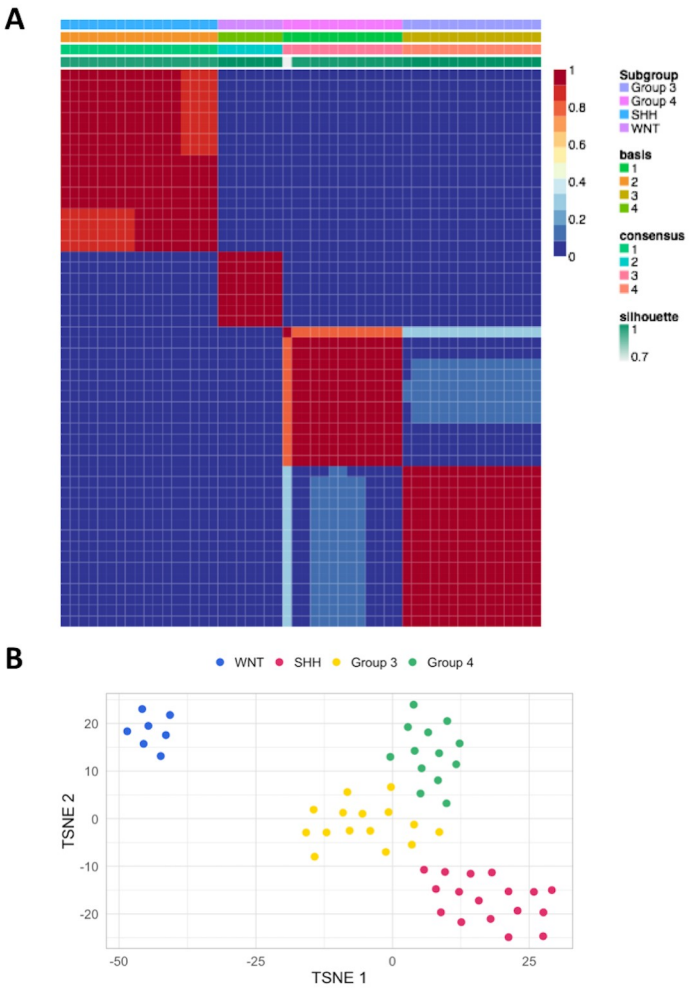

Fig 2. Pediatric MBs were divided into four molecular subgroups. (A) unsupervised NMF and (B) t-SNE analyses of the RNA-Seq gene-expression table. The 52 patients with MB were classified into WNT (7), SHH (17), Group 3 (14), and Group 4 (14). Abbreviations: NMF: Nonnegative matrix factorization; t-SNE: t-distributed stochastic neighbor embedding; WNT: Wingless; SHH: Sonic hedgehog; MB: Medulloblastoma.

https://doi.org/10.1371/journal.pone.0255500.g002

Difference of radiomics features in 4 groups. We have compared the 253 radiomic image features between 4 molecular subtypes of medulloblastoma. 6 features were significant different between 3 groups of medulloblastoma and 2 features were significant different between 4 groups of medulloblastoma. In 4 features (Cluster Tendency, Contrast, Difference entropy, Dissimilarity, Fig 3A–3D), WNT and G3 groups tended to have the higher values than SHH and G4 groups. In the rest 4 features (Entropy, Inverse Difference Normalized (IDN), Inverse Difference Moment Normalized (IDMN), Cluster Prominence, Fig 3E–3H) SHH and G4 groups tended to have the higher values than WNT and G3 groups. All of them are textural features illustrating local patterns in tumors.

The impact of imaging biomarkers on disease survival. In univariate Cox Proportional-Hazards regression model, 6 radiomic features were proved to be significant correlated to the period of patient survival: Mean ($p = 0.04$), Median ($p = 0.04$), Skewness ($p = 0.02$), Energy ($p = 0.01$), Short Run Low Gray-Level Emphasis (SRLGLE) ($p = 0.01$), and Long Run Low Gray-Level Emphasis (LRLGLE) ($p = 0.04$). However, only Short Run Low Gray-Level

Table 1. Clinical profiles of the study cohort.

	WNT	SHH	Group 3	Group 4
Case number	7	12	8	11
Age at diagnosis	6.2 ± 4.1	5.5 ± 4.5	8.3 ± 5.1	10.0 ± 2.9
Boy–girl ratio	1:6	5:7	5:3	6:5
Clinical risk stratification				
Non-met (M0-1) AR	7	7	4	3
Non-met (M0-1) HR	0	3	3	5
Met (M2-3) HR	0	2	1	3
Treatment strategy				
CMT alone	0	1	0	0
RT alone	0	1	1	1
RT + CMT	7	9	7	10

Abbreviations: WNT: Wingless; SHH: Sonic hedgehog; Met: Metastasis; AR: Average risk; HR: High risk; CMT: Chemotherapy; RT: Radiation therapy.

https://doi.org/10.1371/journal.pone.0255500.t001

Emphasis (SRLGLE) (p = 0.02) were significant correlated to patient overall survival in multivariate Cox Proportional-Hazards regression model.

SRLGLE was selected as the prognosticator. And we found that in tumors with SRLGLE values higher than the third quartile values 0.08, patients have shorter median survival than tumor don't have these features (627 vs 1666 days, Hazard ratio: 4.28; 95% Confidence Interval 1.03 to 17.78, p = 0.04, Log-rank [Mantel-Cox] test, Fig 4).

Preliminary prediction model for molecular subgroups of MB. Our result revealed that features extracted from CET1 GLCM features by using sequential forward selection algorithm delivered the highest performance to differentiate 4 molecular subgroups. Receiver operating characteristic analysis revealed optimal performance for predicting WNT, Group 3, and Group 4 MB (area under the curve [AUC] = 0.82, 0.72, and 0.78, respectively). But the prediction performance for SHH MB is suboptimal (AUC = 0.50). (Fig 5) The overall accuracy of the model to estimate the molecular subtypes of each MB is 71%. The model proposed in this study is still in the preliminary stage because a meaningful prediction model was hard to be created based on such a small cohort. The detailed methods and other preliminary results were presented in supplement S1 Text, S2 and S3 Tables.

Discussion

Our result demonstrated that there is significant difference in MR radiomics features between different MB molecular subtypes. Some features can be applied as prognosticators to predict the outcome of MB.

In 2010, an international panel of experts reached consensus on the four main subgroups of MB [24,26]. This proposal was subsequently adopted in the revised World Health Organization classification of central nervous system tumors published in 2016 [27]. WNT group MBs originate from dorsal brainstem precursors [28] and are featured by activation of the WNT pathway. Mutations in exon 3 of *CTNNB1* and monosomy chromosome 6 are usually detected [29]. SHH group MBs arise from granule cell precursors of the cerebellum and cochlear nucleus [30] and are characterized by activation of the SHH pathway. This group commonly contain mutations in the SHH pathway, including *SMO*, *PTCH* and *SUFU* [31]. Group 3 MBs are featured by recurrent MYC amplifications and almost 20% of patients harboring an MYC amplicon [29]. Overall outcome is poorer compared with the other subgroups. Group 4 is the

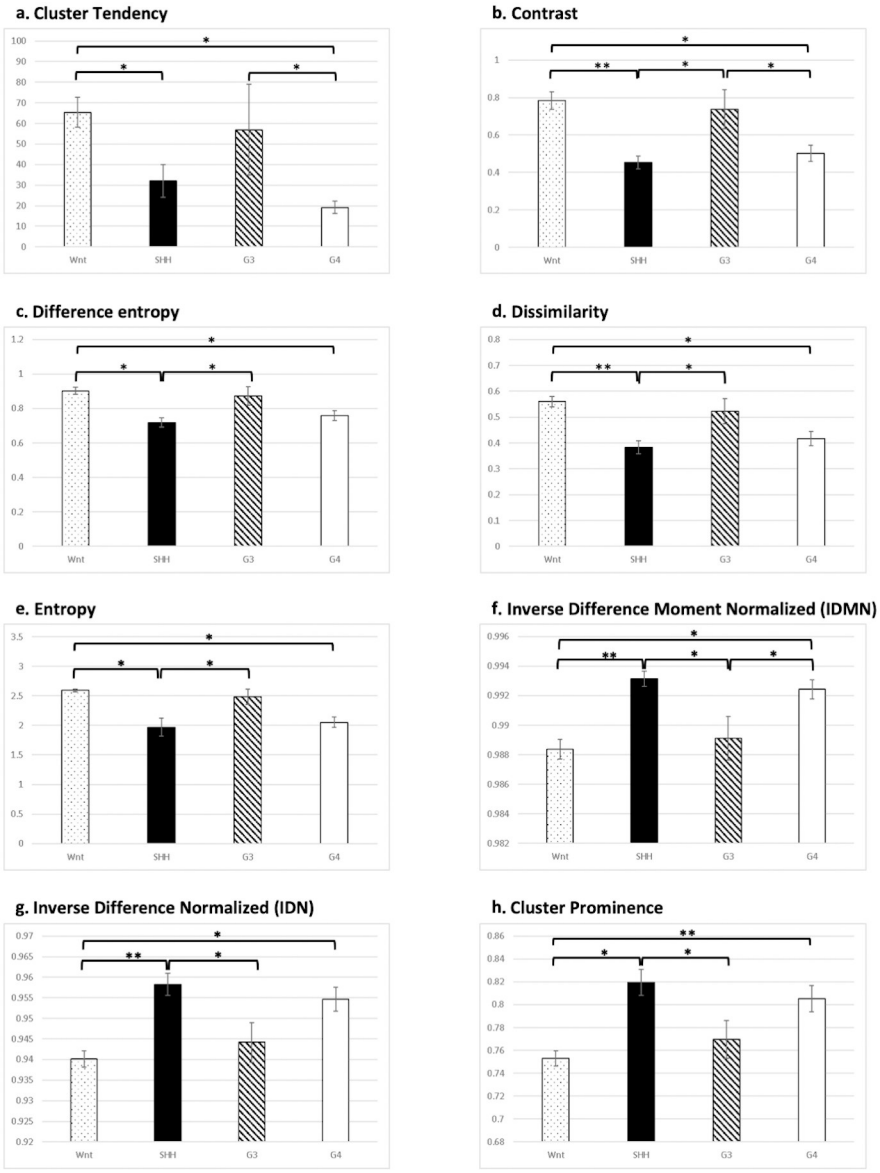

Fig 3. Comparison of the radiomics features which were statistically different between 4 molecular subtypes of medulloblastoma. In first 4 features (A-D) WNT and G3 groups tended to have the higher values than SHH and G4 groups. In the rest 4 features (E-H) SHH and G4 groups tended to have the higher values than WNT and G3 groups. (*: $p<0.05$; **: $p<0.01$).

https://doi.org/10.1371/journal.pone.0255500.g003

most common subgroup, and the most common aberration is isochromosome 17q, *MYCN* amplifications, *SNCAIP* duplications and loss of 11q [32]. Molecular subgroups can be identified using various genomic approaches and platforms, including fluorescence *in situ* hybridization or expression array methods, molecular inversion probe single-nucleotide polymorphism assay, or genome-wide methylation [16]. Two independent validated methods performed in accredited laboratories were recommended to reach the correct subgroup assignment [33]. Clinical trials which tailor therapies for each of the subgroups and assess whether this approach can improve outcomes are underway [34].

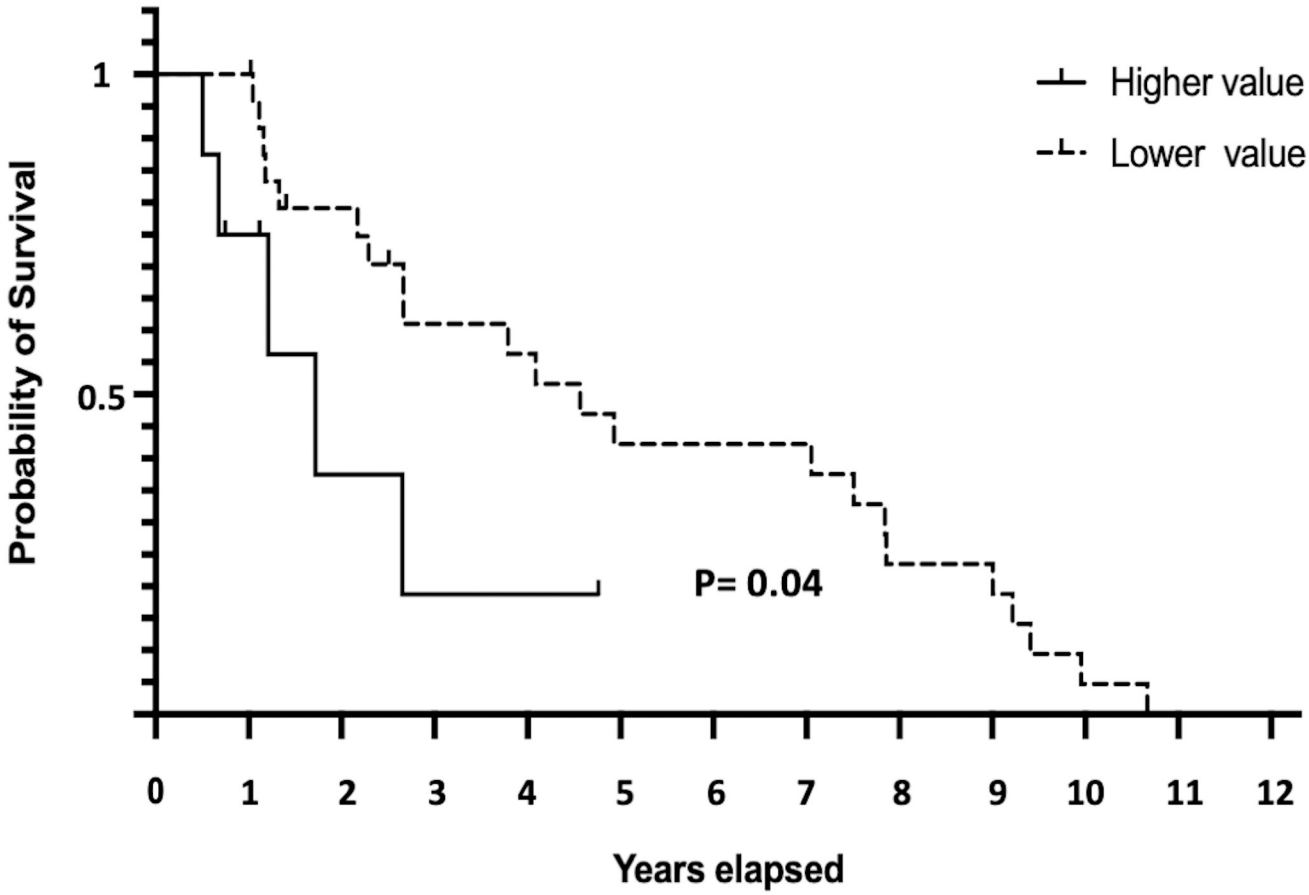

Fig 4. Comparison of survival in MB patients with different radiomics values. For tumors with Energy value > 0.23 or SRLGLE value > 0.08, patients have shorter survival time than tumor don't have these features (p = 0.04 in Log-rank [Mantel-Cox] test).

https://doi.org/10.1371/journal.pone.0255500.g004

Linking MRI features with genomic data can help identify not only unique genetic information but also MR phenotypes, which may play a role in genetic dosage alteration [8]. We proposed several quantitative radiomic features in this study for predicting molecular subgroups. These features were categorized according to intensity, morphology, and texture, each providing different perspectives for distinguishing MB molecular subgroups. According to our results, texture features proved the most distinct in differentiating MBs at the molecular level. In previous reports [10,35], researchers likewise observed that texture features in MR images were the most illustrative phenotype for differentiating molecular subtypes of brain glioma. Texture analysis is a defining feature of radiomics that describes the patterns and spatial variations of voxel intensities, calculated from gray-level co-occurrence and gray-level run-length texture matrices, respectively. The classifying features in our cohort were Cluster tendency, Cluster prominence, Entropy, Difference entropy, Dissimilarity, IDN, IDMN, and Contrast. Cluster tendency measures groupings of voxels with similar gray-level values. As a measure of heterogeneity, it places higher weights on neighboring intensity level pairs that deviate more from the mean [18]. Cluster prominence is a measurement of the skewness and asymmetry of gray-level co-occurrence. Higher values imply more asymmetry of the mean value, whereas lower values imply a peak near the mean with less variation [36]. Entropy is a statistical measure of randomness that has been proven to capture intratumoral heterogeneity [37] and has

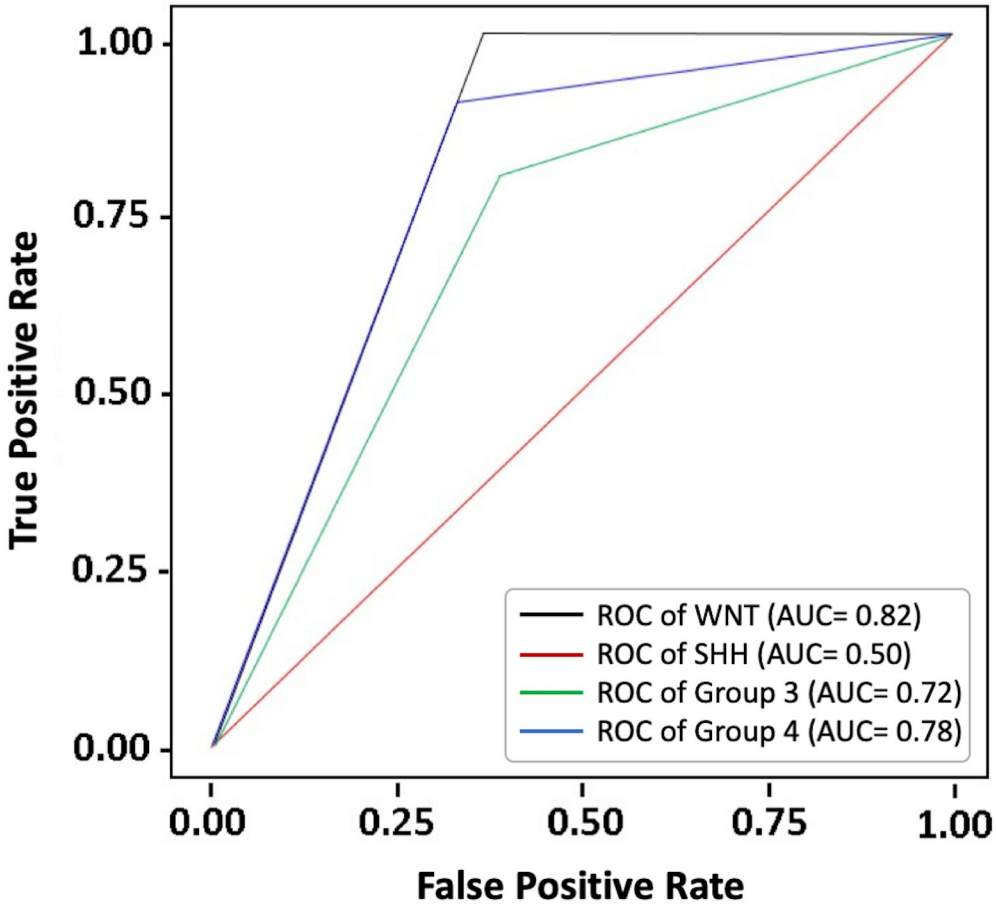

Fig 5. Receiver operating characteristic analysis of the preliminary prediction model. The prediction performance for predicting WNT, SHH, Group 3, and Group 4 MB was illustrated (area under the curve = 0.82, 0.50, 0.72, and 0.78, respectively).

https://doi.org/10.1371/journal.pone.0255500.g005

also been associated with tumor staging [38], outcome [39], and expression of molecular pathways. Dissimilarity illustrates local intensity variation, which is defined as the mean absolute difference between the neighboring pairs; a higher value represent greater disparity in intensities among neighboring voxels [40]. IDN (also called similarity) and IDMN (also called homogeneity) are both measures of the local homogeneity [40]. Our study revealed that WNT-MBs have a similar texture value to Group 3 MBs and SHH-MBs have similar texture features to Group 4 MBs. Further studies on the similarities and dissimilarities of these molecular subgroups are warranted.

In our machine learning protocol, features generated from multiparametric MR sequences including CET1, T1W, T2W, T2 FLAIR, and ADC map were input into the algorithms to determine the optimal features for molecular subgroup prediction. The results revealed that the combination of textural features generated from CET1 achieved the highest ACC. We therefore only compared features generated from CET1 images. In CET1 images, several pathological features including tumor necrosis and cysts can be clearly identified. Furthermore, the intensity of contrast enhancement is linked to angiogenesis module activity within the tumor [41,42]. Therefore, it is reasonable that CET1 features may depict the molecular profiles and predict outcome of MB because they provide substantial information regarding tumor

characteristics. Iv et al's report also investigated the prediction model for medulloblastoma [43]. However, their model didn't take all routine MR sequences into consideration. And even with substantial case number, the prediction accuracy for some molecular subgroup is still suboptimal.

Several radiomic features have been identified as potential prognosticators in our study. Many prognostic factors of medulloblastoma patients has been proposed, including presence of metastases at diagnosis, age < 3 years, extend of resection, residual disease $\geq 1{\cdot}5\ cm^2$, molecular subgroups, craniospinal radiation, and geographic location of therapy [44–47]. Imaging is a routine examination for pediatric brain tumors, with no additional costs associated. By applying radiomic image features, we can quantify all tumor characteristics without risking a biopsy. Therefore, radiomics analysis may be another helpful tool for stratifying patients into different risk groups.

Limitations

Because the MB occurrence rate is low, this preliminary study was limited by its small sample size. Therefore the accuracy of the established prediction model based on the machine learning is not optimal (S3 Table). Further studies with a larger sample size are warranted. The trained prediction model and developed MRP are available on our website (www.ym.edu.tw/~cflu/MRP_MLinglioma.html), and the complete content will be available after paper publication to encourage researchers worldwide to test these models and refine them accordingly. Another limitation is that we only used images generated from original MR images. Further investigations on other image postprocessing techniques, such as wavelets or ranklets, are necessary.

In conclusion, 8 CET1 texture features were found to be significant different between 4 molecular subgroups of MB. Another CET1 features were found to be a good prognosticator of MB. These features may provide suggestions for further stratifying patients with MB into different risk subgroups.

Supporting information

S1 Fig. Diagram showing nested leave-one-out cross-validation procedure. The data are repeatedly split in testing and decoding sets. The decoding set itself is split in multiple training and validation sets with the same decoding set, forming an inner cross-validation loop used to set the regularization hyperparameter, while the external loop varying the testing set is used to measure the performance of prediction. The process was repeated according to the case number in both inner and external loops. (* 4 cases were excluded during the validation because of their missing values).
(PDF)

S1 Table. Employed 57 radiomic features.
(PDF)

S2 Table. Prediction results obtained in different molecular subgroups using the proposed model based on the sequential forward selection algorithm. The details of the prediction results in different molecular subgroups were demonstrated by using the proposed model based on the selected CET1 features. Overall prediction accuracy was highest in WNT.
(PDF)

S3 Table. Accuracy of the prediction model based on different combinations of imaging features extracted from different MR parameters and feature selection algorithms. The Highest accuracy was obtained with sequential forward selection algorithm using CET1 images. (Abbreviations: mRMR, minimum redundancy maximum relevance; SBE, sequential

backward elimination; SFS, sequential forward selection).
(PDF)

S1 Text. Establishing a prediction model. Taking the RNA-Seq-based molecular subtypes as the gold standard, a prediction model was established to estimate the molecular subtypes of each MB based on the most significant radiomics imaging features defined by the feature selection procedure. The generalizability of the selected features was validated through leave-one-out cross-validation (LOOCV).
(PDF)

Acknowledgments

We thank Dr. Michael Taylor and the laboratory staff for helping validate the subgroup classification based on counterpart clustering of the RNA-Seq data. The authors also like to thank Dr. Po-Chih Kuo for his astute statistical consultation.

Author Contributions

Conceptualization: Tai-Tong Wong, Kevin Li-Chun Hsieh.

Data curation: Feng-Chi Chang, Tai-Tong Wong, Kuo-Sheng Wu, Chia-Feng Lu, Chih-Chun Wu, Wan Yuo Guo, Kevin Li-Chun Hsieh.

Formal analysis: Ting-Wei Weng, Kevin Li-Chun Hsieh.

Funding acquisition: Tai-Tong Wong, Muh-Lii Liang, Kevin Li-Chun Hsieh.

Investigation: Feng-Chi Chang, Tai-Tong Wong, Ting-Wei Weng, Kevin Li-Chun Hsieh.

Methodology: Kuo-Sheng Wu, Ting-Wei Weng, Kevin Li-Chun Hsieh.

Project administration: Tai-Tong Wong, Ting-Wei Weng, Kevin Li-Chun Hsieh.

Resources: Chia-Feng Lu, Cheng-Yu Chen, Kevin Li-Chun Hsieh.

Software: Kuo-Sheng Wu, Chia-Feng Lu, Ting-Wei Weng, Kevin Li-Chun Hsieh.

Supervision: Tai-Tong Wong, Cheng-Yu Chen, Kevin Li-Chun Hsieh.

Validation: Tai-Tong Wong, Kevin Li-Chun Hsieh.

Visualization: Ting-Wei Weng, Kevin Li-Chun Hsieh.

Writing – original draft: Muh-Lii Liang, Kevin Li-Chun Hsieh.

Writing – review & editing: Feng-Chi Chang, Tai-Tong Wong, Muh-Lii Liang, Kevin Li-Chun Hsieh.

References

1. Polkinghorn WR, Tarbell NJ. Medulloblastoma: Tumorigenesis, current clinical paradigm, and efforts to improve risk stratification. Nature Clinical Practice Oncology. 2007; 4(5):295–304. https://doi.org/10.1038/ncponc0794 PMID: 17464337
2. Wu K-S, Ho DM-T, Jou S-T, Yu AL, Tran HM, Liang M-L, et al. Molecular-clinical correlation in pediatric medulloblastoma: A cohort series study of 52 cases in Taiwan. Cancers. 2020; 12(3):653. https://doi.org/10.3390/cancers12030653 PMID: 32168907
3. Northcott PA, Korshunov A, Witt H, Hielscher T, Eberhart CG, Mack S, et al. Medulloblastoma comprises four distinct molecular variants. J Clin Oncol [Internet]. 2011; 29(11):1408–14. Available from: https://www.ncbi.nlm.nih.gov/pubmed/20823417. https://doi.org/10.1200/JCO.2009.27.4324

4. Roussel MF, Stripay JL. Epigenetic Drivers in Pediatric Medulloblastoma. Cerebellum [Internet]. 2017; 17(1):28–36. Available from: https://www.ncbi.nlm.nih.gov/pubmed/29178021.

5. Cavalli FMG, Remke M, Rampasek L, Peacock J, Shih DJH, Luu B, et al. Intertumoral Heterogeneity within Medulloblastoma Subgroups. Cancer Cell [Internet]. 2017; 31(6):737–754.e6. Available from: https://www.ncbi.nlm.nih.gov/pubmed/28609654. https://doi.org/10.1016/j.ccell.2017.05.005

6. Albright AL, Packer RJ, Zimmerman R, Rorke LB, Boyett J, Hammond GD. Magnetic resonance scans should replace biopsies for the diagnosis of diffuse brain stem gliomas: a report from the Children's Cancer Group. Neurosurgery [Internet]. 1993; 33(6):1026–30. Available from: https://www.ncbi.nlm.nih.gov/pubmed/8133987. https://doi.org/10.1227/00006123-199312000-00010

7. Wefers AK, Warmuth-Metz M, Poschl J, von Bueren AO, Monoranu CM, Seelos K, et al. Subgroup-specific localization of human medulloblastoma based on pre-operative MRI. Acta Neuropathol [Internet]. 2014; 127(6):931–3. Available from: https://www.ncbi.nlm.nih.gov/pubmed/24699697. https://doi.org/10.1007/s00401-014-1271-5

8. Perreault S, Ramaswamy V, Achrol AS, Chao K, Liu TT, Shih D, et al. MRI surrogates for molecular subgroups of medulloblastoma. American Journal of Neuroradiology. 2014; 35(7):1263–9. https://doi.org/10.3174/ajnr.A3990 PMID: 24831600

9. Kumar V, Gu Y, Basu S, Berglund A, Eschrich SA, Schabath MB, et al. Radiomics: The process and the challenges. Magnetic Resonance Imaging. 2012; 30(9):1234–48. https://doi.org/10.1016/j.mri.2012.06.010 PMID: 22898692

10. Lu C-F, Hsu F-T, Hsieh KL-C, Kao Y-CJ, Cheng S-J, Hsu JB-K, et al. Machine learning–based radiomics for molecular subtyping of gliomas. Clinical Cancer Research. 2018; 24(18):4429–36. https://doi.org/10.1158/1078-0432.CCR-17-3445 PMID: 29789422

11. Ellison DW. Childhood medulloblastoma: Novel approaches to the classification of a heterogeneous disease. Acta Neuropathologica. 2010; 120(3):305–16. https://doi.org/10.1007/s00401-010-0726-6 PMID: 20652577

12. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. Nat Biotechnol [Internet]. 2016; 34(5):525–7. Available from: https://www.ncbi.nlm.nih.gov/pubmed/27043002. https://doi.org/10.1038/nbt.3519

13. Soneson C, Love MI, Robinson MD. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Res [Internet]. 2015; 30(4):1521. Available from: https://www.ncbi.nlm.nih.gov/pubmed/26925227. https://doi.org/10.12688/f1000research.7563.2

14. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol [Internet]. 2014; 15(12):550. Available from: https://www.ncbi.nlm.nih.gov/pubmed/25516281. https://doi.org/10.1186/s13059-014-0550-8

15. Gaujoux R, Seoighe C. A flexible R package for nonnegative matrix factorization. BMC Bioinformatics [Internet]. 2010; 11:367. Available from: https://www.ncbi.nlm.nih.gov/pubmed/20598126. https://doi.org/10.1186/1471-2105-11-367

16. Northcott PA, Shih DJ, Remke M, Cho YJ, Kool M, Hawkins C, et al. Rapid, reliable, and reproducible molecular sub-grouping of clinical medulloblastoma samples. Acta Neuropathol [Internet]. 2012; 123(4):615–26. Available from: https://www.ncbi.nlm.nih.gov/pubmed/22057785. https://doi.org/10.1007/s00401-011-0899-7

17. Buch K, Li B, Qureshi MM, Kuno H, Anderson SW, Sakai O. Quantitative assessment of variation in CT parameters on texture features: Pilot study using a nonanatomic phantom. American Journal of Neuroradiology. 2017; 38(5):981–5. https://doi.org/10.3174/ajnr.A5139 PMID: 28341714

18. Aerts HJWL, Velazquez ER, Leijenaar RTH, Parmar C, Grossmann P, Cavalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nature Communications. 2014; 3(5):4006. https://doi.org/10.1038/ncomms5006 PMID: 24892406

19. Zwanenburg A, Vallières M, Abdalah MA, Aerts HJWL, Andrearczyk V, Apte A, et al. The image biomarker standardization initiative: Standardized quantitative radiomics for high-throughput image-based phenotyping. Radiology [Internet]. 2020 May 1 [cited 2021 May 21]; 295(2):328–38. Available from: https://doi.org/10.1148/radiol.2020191145. PMID: 32154773

20. Yin P, Shi JQ. Simulation-based sensitivity analysis for non-ignorably missing data. Stat Methods Med Res [Internet]. 2019; 28(1):289–308. Available from: https://www.ncbi.nlm.nih.gov/pubmed/28747095. https://doi.org/10.1177/0962280217722382

21. Batista GEAPA, Monard MC. A Study of K-Nearest Neighbour as an Imputation Method. Computer Science. 2002.

22. Radovic M, Ghalwash M, Filipovic N, Obradovic Z. Minimum redundancy maximum relevance feature selection approach for temporal gene expression data. BMC Bioinformatics [Internet]. 2017/01/05. 2017; 18(1):9. Available from: https://www.ncbi.nlm.nih.gov/pubmed/28049413. https://doi.org/10.1186/s12859-016-1423-9

23. Chandrashekar G, Sahin F. A survey on feature selection methods. Computers & Electrical Engineering. 2014; 40(1):16–28.

24. Huang S, Cai N, Pacheco PP, Narandes S, Wang Y, Xu W. Applications of Support Vector Machine (SVM) Learning in Cancer Genomics. Cancer Genomics Proteomics [Internet]. 2017/12/25. 2018; 15 (1):41–51. Available from: https://www.ncbi.nlm.nih.gov/pubmed/29275361. https://doi.org/10.21873/cgp.20063

25. Varoquaux G, Raamana PR, Engemann DA, Hoyos-Idrobo A, Schwartz Y, Thirion B. Assessing and tuning brain decoders: Cross-validation, caveats, and guidelines. NeuroImage. 2017 Jan 15; 145:166–79. https://doi.org/10.1016/j.neuroimage.2016.10.038 PMID: 27989847

26. Taylor MD, Northcott PA, Korshunov A, Remke M, Cho YJ, Clifford SC, et al. Molecular subgroups of medulloblastoma: the current consensus. Acta Neuropathol [Internet]. 2012; 123(4):465–72. Available from: https://www.ncbi.nlm.nih.gov/pubmed/22134537. https://doi.org/10.1007/s00401-011-0922-z

27. Louis DN, Perry A, Reifenberger G, von Deimling A, Figarella-Branger D, Cavenee WK, et al. The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. Acta Neuropathol [Internet]. 2016; 131(6):803–20. Available from: https://www.ncbi.nlm.nih.gov/pubmed/27157931. https://doi.org/10.1007/s00401-016-1545-1

28. Gibson P, Tong Y, Robinson G, Thompson MC, Currle DS, Eden C, et al. Subtypes of medulloblastoma have distinct developmental origins. Nature [Internet]. 2010; 468(7327):1095–9. Available from: https://www.ncbi.nlm.nih.gov/pubmed/21150899. https://doi.org/10.1038/nature09587

29. Kool M, Korshunov A, Remke M, Jones DT, Schlanstein M, Northcott PA, et al. Molecular subgroups of medulloblastoma: an international meta-analysis of transcriptome, genetic aberrations, and clinical data of WNT, SHH, Group 3, and Group 4 medulloblastomas. Acta Neuropathol [Internet]. 2012; 123 (4):473–84. Available from: https://www.ncbi.nlm.nih.gov/pubmed/22358457. https://doi.org/10.1007/s00401-012-0958-8

30. Grammel D, Warmuth-Metz M, von Bueren AO, Kool M, Pietsch T, Kretzschmar HA, et al. Sonic hedgehog-associated medulloblastoma arising from the cochlear nuclei of the brainstem. Acta Neuropathol [Internet]. 2012; 123(4):601–14. Available from: https://www.ncbi.nlm.nih.gov/pubmed/22349907. https://doi.org/10.1007/s00401-012-0961-0

31. Kool M, Jones DT, Jager N, Northcott PA, Pugh TJ, Hovestadt V, et al. Genome sequencing of SHH medulloblastoma predicts genotype-related response to smoothened inhibition. Cancer Cell [Internet]. 2014; 25(3):393–405. Available from: https://www.ncbi.nlm.nih.gov/pubmed/24651015. https://doi.org/10.1016/j.ccr.2014.02.004

32. Northcott PA, Shih DJ, Peacock J, Garzia L, Morrissy AS, Zichner T, et al. Subgroup-specific structural variation across 1,000 medulloblastoma genomes. Nature [Internet]. 2012; 488(7409):49–56. Available from: https://www.ncbi.nlm.nih.gov/pubmed/22832581. https://doi.org/10.1038/nature11327

33. Gottardo NG, Hansford JR, McGlade JP, Alvaro F, Ashley DM, Bailey S, et al. Medulloblastoma Down Under 2013: a report from the third annual meeting of the International Medulloblastoma Working Group. Acta Neuropathol [Internet]. 2014; 127(2):189–201. Available from: https://www.ncbi.nlm.nih.gov/pubmed/24264598. https://doi.org/10.1007/s00401-013-1213-7

34. Robinson GW, Orr BA, Wu G, Gururangan S, Lin T, Qaddoumi I, et al. Vismodegib Exerts Targeted Efficacy Against Recurrent Sonic Hedgehog-Subgroup Medulloblastoma: Results From Phase II Pediatric Brain Tumor Consortium Studies PBTC-025B and PBTC-032. J Clin Oncol [Internet]. 2015; 33 (24):2646–54. Available from: https://www.ncbi.nlm.nih.gov/pubmed/26169613. https://doi.org/10.1200/JCO.2014.60.1591

35. Yang D, Rao G, Martinez J, Veeraraghavan A, Rao A. Evaluation of tumor-derived MRI-texture features for discrimination of molecular subtypes and prediction of 12-month survival status in glioblastoma. Medical Physics. 2015; 42(11):6725–35. https://doi.org/10.1118/1.4934373 PMID: 26520762

36. Unser M. Sum and difference histograms for texture classification. IEEE Trans Pattern Anal Mach Intell [Internet]. 1986; 8(1):118–25. Available from: https://www.ncbi.nlm.nih.gov/pubmed/21869331. https://doi.org/10.1109/tpami.1986.4767760

37. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. Radiology. 2016; 278(2):563–77. https://doi.org/10.1148/radiol.2015151169 PMID: 26579733

38. Ganeshan B, Skogen K, Pressney I, Coutroubis D, Miles K. Tumour heterogeneity in oesophageal cancer assessed by CT texture analysis: Preliminary evidence of an association with tumour metabolism, stage, and survival. Clinical Radiology. 2012; 67(2):157–64. https://doi.org/10.1016/j.crad.2011.08.012 PMID: 21943720

39. Ng F, Kozarski R, Ganeshan B, Goh V. Assessment of tumor heterogeneity by CT texture analysis: Can the largest cross-sectional area be used as an alternative to whole tumor analysis? European Journal of Radiology. 2013; 82(2):342–8. https://doi.org/10.1016/j.ejrad.2012.10.023 PMID: 23194641

40. Van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational radiomics system to decode the radiographic phenotype. Cancer Research. 2017; 77(21):e104–7. https://doi.org/10.1158/0008-5472.CAN-17-0339 PMID: 29092951

41. Diehn M, Nardini C, Wang DS, McGovern S, Jayaraman M, Liang Y, et al. Identification of noninvasive imaging surrogates for brain tumor gene-expression modules. Proc Natl Acad Sci U S A [Internet]. 2008; 105(13):5213–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18362333. https://doi.org/10.1073/pnas.0801279105

42. Pope WB, Chen JH, Dong J, Carlson MR, Perlina A, Cloughesy TF, et al. Relationship between gene expression and enhancement in glioblastoma multiforme: exploratory DNA microarray analysis. Radiology [Internet]. 2008; 249(1):268–77. Available from: http://www.ncbi.nlm.nih.gov/pubmed/18796682. https://doi.org/10.1148/radiol.2491072000

43. Iv M, Zhou M, Shpanskaya K, Perreault S, Wang Z, Tranvinh E, et al. MR imaging-based radiomic signatures of distinct molecular subgroups of medulloblastoma. American Journal of Neuroradiology [Internet]. 2019 Jan 1 [cited 2021 May 21]; 40(1):154–61. Available from: https://pubmed.ncbi.nlm.nih.gov/30523141/. https://doi.org/10.3174/ajnr.A5899 PMID: 30523141

44. Simonsen CZ, Madsen MH, Schmitz ML, Mikkelsen IK, Fisher M, Andersen G. Sensitivity of diffusion- and perfusion-weighted imaging for diagnosing acute ischemic stroke is 97.5%. Stroke [Internet]. 2015; 46(1):98–101. Available from: https://www.ncbi.nlm.nih.gov/pubmed/25388415. https://doi.org/10.1161/STROKEAHA.114.007107

45. Gajjar A, Chintagumpala M, Ashley D, Kellie S, Kun LE, Merchant TE, et al. Risk-adapted craniospinal radiotherapy followed by high-dose chemotherapy and stem-cell rescue in children with newly diagnosed medulloblastoma (St Jude Medulloblastoma-96): long-term results from a prospective, multicentre trial. Lancet Oncology. 2006; 7(10):813–20. https://doi.org/10.1016/S1470-2045(06)70867-1 PMID: 17012043

46. Zeltzer PM, Boyett JM, Finlay JL, Albright AL, Rorke LB, Milstein JM, et al. Metastasis stage, adjuvant treatment, and residual tumor are prognostic factors for medulloblastoma in children: Conclusions from the Children's Cancer Group 921 randomized phase III study. Journal of Clinical Oncology. 1999; 17(3):832–45. https://doi.org/10.1200/JCO.1999.17.3.832 PMID: 10071274

47. Thompson EM, Hielscher T, Bouffet E, Remke M, Luu B, Gururangan S, et al. Prognostic value of medulloblastoma extent of resection after accounting for molecular subgroup: a retrospective integrated clinical and molecular analysis. The Lancet Oncology. 2016; 17(4):484–95. https://doi.org/10.1016/S1470-2045(15)00581-1 PMID: 26976201