



OPEN

## Combining multi-site magnetic resonance imaging with machine learning predicts survival in pediatric brain tumors

James T. Grist<sup>1</sup>, Stephanie Withey<sup>1,2,3</sup>, Christopher Bennett<sup>1</sup>, Heather E. L. Rose<sup>1,2</sup>, Lesley MacPherson<sup>4</sup>, Adam Oates<sup>4</sup>, Stephen Powell<sup>1</sup>, Jan Novak<sup>2,5,6</sup>, Laurence Abernethy<sup>7</sup>, Barry Pizer<sup>8</sup>, Simon Bailey<sup>9</sup>, Steven C. Clifford<sup>10</sup>, Dipayan Mitra<sup>11</sup>, Theodoros N. Arvanitis<sup>1,2,12</sup>, Dorothee P. Auer<sup>13,14</sup>, Shivaram Avula<sup>7</sup>, Richard Grundy<sup>15</sup> & Andrew C. Peet<sup>1,2</sup>✉

Brain tumors represent the highest cause of mortality in the pediatric oncological population. Diagnosis is commonly performed with magnetic resonance imaging. Survival biomarkers are challenging to identify due to the relatively low numbers of individual tumor types. 69 children with biopsy-confirmed brain tumors were recruited into this study. All participants had perfusion and diffusion weighted imaging performed at diagnosis. Imaging data were processed using conventional methods, and a Bayesian survival analysis performed. Unsupervised and supervised machine learning were performed with the survival features, to determine novel sub-groups related to survival. Sub-group analysis was undertaken to understand differences in imaging features. Survival analysis showed that a combination of diffusion and perfusion imaging were able to determine two novel sub-groups of brain tumors with different survival characteristics ( $p < 0.01$ ), which were subsequently classified with high accuracy (98%) by a neural network. Analysis of high-grade tumors showed a marked difference in survival ( $p = 0.029$ ) between the two clusters with high risk and low risk imaging features. This study has developed a novel model of survival for pediatric brain tumors. Tumor perfusion plays a key role in determining survival and should be considered as a high priority for future imaging protocols.

Brain tumors represent one of the most common causes of pediatric and adult oncological mortality. Particular challenges are faced in clinical pediatric oncology research due to the highly heterogeneous nature of pediatric tumors, combined with the relative rarity of the disease in the general population<sup>1</sup>. Despite this, multi-center studies have allowed impressive advances to be made in the understanding of the major types of children's brain tumors and these are starting to change clinical practice<sup>2,3</sup>. The majority of studies have relied on analysis of tumor tissue; however, medical imaging is becoming increasingly able to probe tissue properties and has the advantage that measurements are made directly in vivo. This is particularly important for probing the tissue microenvironment since quantities such as perfusion cannot be readily determined in tissue samples. Imaging

<sup>1</sup>Institute of Cancer and Genomic Sciences, School of Medical and Dental Sciences, University of Birmingham, Birmingham, UK. <sup>2</sup>Oncology, Birmingham Women's and Children's NHS Foundation Trust, Birmingham, UK. <sup>3</sup>RRPPS, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK. <sup>4</sup>Radiology, Birmingham Women's and Children's NHS Foundation Trust, Birmingham, UK. <sup>5</sup>Psychology, College of Health and Life Sciences Aston University, Birmingham, UK. <sup>6</sup>Aston Neuroscience Institute, Aston University, Birmingham, UK. <sup>7</sup>Radiology, Alder Hey Children's NHS Foundation Trust, Liverpool, UK. <sup>8</sup>Oncology, Alder Hey Children's NHS Foundation Trust, Liverpool, UK. <sup>9</sup>Sir James Spence Institute of Child Health, Royal Victoria Infirmary, Newcastle upon Tyne, UK. <sup>10</sup>Wolfson Childhood Cancer Research Centre, Newcastle University Centre for Cancer, University of Newcastle, Newcastle upon Tyne, UK. <sup>11</sup>Neuroradiology, Royal Victoria Infirmary, Newcastle Upon Tyne, UK. <sup>12</sup>Institute of Digital Healthcare, WMG, University of Warwick, Coventry, UK. <sup>13</sup>Sir Peter Mansfield Imaging Centre, University of Nottingham Biomedical Research Centre, Nottingham, UK. <sup>14</sup>NIHR Nottingham Biomedical Research Centre, Nottingham, UK. <sup>15</sup>The Children's Brain Tumor Research Centre, University of Nottingham, Nottingham, UK. ✉email: a.peet@bham.ac.uk

therefore has the potential to provide new biomarkers of prognosis which can be obtained early and throughout the patient journey.

Recently, an increased understanding of pediatric brain tumor biology has enabled more accurate prognostication for individual patients. The findings have largely been based on molecular genetic markers identified in tissue. For example, in medulloblastoma, biological subgrouping has shown that WNT subgroup tumors have an excellent prognosis whereas group 3 tumors and subsets of SHH tumors have an inferior outcome<sup>4</sup>. However, in even rarer tumors, such as atypical rhabdoid tumors (ATRT), or midline gliomas, where biopsy derived tissue is challenging to acquire, it is more difficult to perform molecular subtyping studies<sup>5</sup>. Therefore, clinical studies have been more difficult to perform in meaningful numbers for many such rare tumor types, and the small biopsies taken may not provide a representative view of the tumor, particularly its microenvironment.

Medical imaging is an important diagnostic aid for brain tumors, since it is non-invasive and can include the whole tumor and surrounding tissue. It is also capable of probing the tumor microenvironment *in vivo*, improving our understanding of the *in vivo* neovascularization and cellularity of the tumor, as well as surrounding cerebral tissue through perfusion and diffusion imaging, respectively<sup>6,7</sup>. However, as mentioned above for biological studies, recruiting large numbers of patients for imaging studies is challenging, and often requires large multi-center trials to glean meaningful results. In spite of this, these non-invasive modalities represent highly attractive methods to derive crucial information surrounding the diagnosis and progression of tumors.

Diffusion imaging is available on every major commercial MRI scanner and is routinely used to assess brain tumors<sup>8</sup>. Apparent diffusion coefficient (ADC) maps represent the speed of water motion in the tissue and this correlated cellularity. Perfusion imaging is often acquired either with dynamic susceptibility contrast (DSC) or arterial spin labelling (ASL) techniques<sup>9,10</sup>. DSC imaging is undertaken through the introduction of an exogenous contrast agent containing gadolinium, and the passage of this bolus through the cerebral vasculature is rapidly imaged and post-processed to form quantitative cerebral blood volume and flow maps<sup>11</sup>. ASL is an approach for measuring cerebral perfusion, harnessing the use of radiofrequency tagging of blood in the supplying vessels to create blood flow-based image contrast<sup>12</sup>. This technique is increasing in popularity due to the removal of the need for a gadolinium containing contrast agent bolus, which may be retained within the brain<sup>13</sup>.

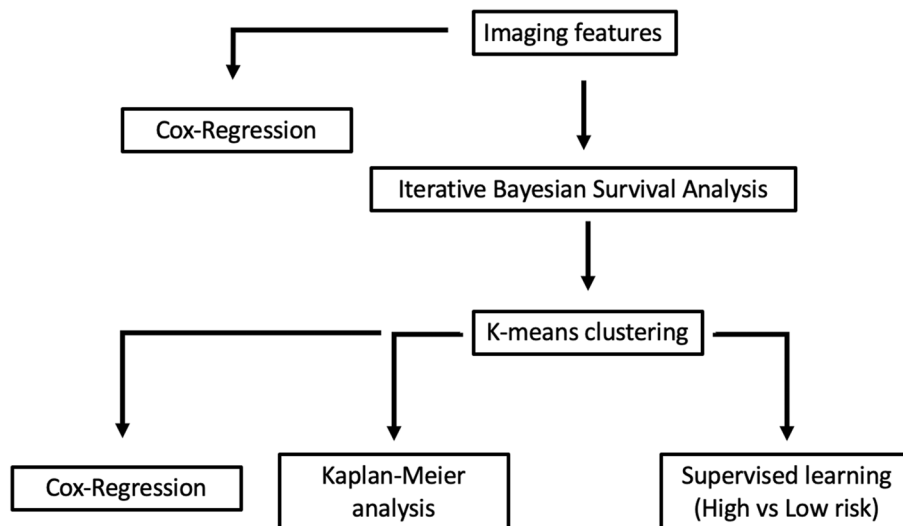
Studies have shown that diffusion and perfusion imaging are able to discriminate between pediatric tumor types *in vivo*, with high cellularity and perfusion in high grade tumors, and vice versa for low-grade<sup>14,15</sup>. This data, in turn, has informed survival analysis models using traditional methods such as Cox-regression to derive significant covariates from imaging data<sup>16</sup>. In particular, ADC mean, elevated cerebral blood flow, and image derived texture parameters have been found to be significant factors in long-term pediatric brain tumor survival<sup>7,17,18</sup>. However, studies which combine these imaging techniques are lacking. Moreover, the use of machine learning has not been implemented in survival analysis of these data types, despite being so successful in improving diagnosis from molecular and imaging data.

In this study we have therefore taken a novel approach to the understanding of risk and survival in pediatric brain tumors through combining diffusion and perfusion MRI and combining this with both supervised and unsupervised machine learning to determine key imaging derived risk factors and novel prognostic groups.

## Methods

**Patient recruitment and imaging.** 69 participants with suspected brain tumors (medulloblastoma (N = 17), pilocytic astrocytoma (N = 22), ependymoma (considered high grade, N = 10), other tumors (N = 20) are found in Supplementary Document 1. They were recruited from four clinical sites in the United Kingdom (Study approved by regional ethics committee, ethics reference: 04/MRE04/41, Birmingham Children's Hospital, Newcastle Royal Victoria Infirmary, Queen's Medical Centre, Alder Hey Children's Hospital, Liverpool). Parents/guardians of participants gave informed consent for their child to be recruited to this study. All methods were performed in accordance with the relevant guidelines and regulations. Recruitment took place between 2009 and 2017. Participants underwent MRI, protocol discussed below, before treatment. Diagnosis was confirmed on tissue samples obtained by surgery performed for clinical reasons. In summary there were 37 complete resections, 17 incomplete resections, 1 open biopsy and 10 stereotactic biopsies with information missing on 4 patients. Further treatment was undertaken according to the clinical needs of the patient as determined by the local tumor board supported by national guidelines and external opinions where appropriate. Radiotherapy tended to be used in the more aggressive tumors and older children, chemotherapy was widely used and tended to be more intensive in younger children with aggressive tumors. Since the study included many different tumor types, stages and patient ages, the treatments varied widely and was tailored to the patient's need. The results of the study were not used in any way to determine or alter treatment. The median age of the cohort at diagnosis was 8 years (range 16 days to 17.6 years). The distribution in Chang stage adapted for use across tumor types was 39 M0, 5 M1, 9 M2, 4 M3, 12 unknown. Diagnoses were made according to the WHO Classification in use at the time of diagnosis and the cohort composition mirrors that found nationally. Tumors were assigned to high-(3&4) and low-grade (1&2) groups. The median follow-up time for the cohort was 4.4 years. Full cohort details can be found in supplementary document 1.

The imaging protocol for all participants was performed either at 3 or 1.5 T and included standard anatomical imaging (T<sub>1</sub>-weighted pre- and post-contrast and T<sub>2</sub>-weighted) as well as diffusion and dynamic susceptibility contrast imaging covering the tumor volume (imaging sequence details found in Supplementary Table 1). Additional clinical data (age at diagnosis and gender) were also collected for analysis. During the data acquisition period a national protocol was in place for both conventional MRI and the advanced imaging analyzed here, some variability in acquisition sequence and parameters was allowed within the protocol<sup>19</sup>.



**Figure 1.** Data processing pipeline used in this study.

**Image post-processing and analysis.** Apparent diffusion coefficient (ADC) maps were calculated from diffusion weighted imaging (DWI) using a linear fit between the two b-value images in Matlab (The Mathworks, MA, 2018a), voxels were fit after the exclusion of noise voxels (any voxel less than a signal to noise ratio of 5 was excluded). DSC time-course data were processed using conventional methods to provide uncorrected cerebral blood volume (uCBV) maps, with a leakage correction undertaken to produce corrected cerebral blood volume (cCBV) and K2 maps<sup>20</sup>.

T<sub>2</sub>-weighted imaging and ADC maps were registered to the first DSC volume with SPM12 (UCL), using a non-linear spatial transform with mutual information. Regions of interest (ROI) segmenting the tumor volume were drawn on the T<sub>2</sub>-weighted imaging<sup>21</sup> after viewing the whole available image set, large cysts and peritumoral oedema were excluded.

Image analysis was performed in Matlab (2018b, The Mathworks, MA), with the image mean, standard deviation, skewness, and kurtosis were calculated on a volume by volume basis for ADC and UCBV/CCBV/K2 maps for regions of interest and the whole brain as previously described<sup>21</sup>. Briefly, a binary whole brain mask calculated and eroded by 3 voxels to remove any signals outside of the head. The whole brain mask included the tumor. Tumor volume (cm<sup>3</sup>) was calculated from the T<sub>2</sub> ROI masks drawn by S.W. (A clinical physicist with more than 15 years of experience in neuroimaging and trained by a consultant pediatric radiologist with 17 years experience L.M., Grist et al.<sup>21</sup>), a random sample of 17 cases had ROIs reviewed by A.O. a consultant pediatric radiologist with 7 years of experience. The list of imaging features are found in the supplementary materials. Regions of interest were also drawn in normal appearing deep grey and white matter for each participant to calculate average diffusion and perfusion measures in normal appearing tissue by J.G. (A physicist with more than 5 years of experience in neuroimaging). Each ROI contained at least 50 voxels. Medulloblastoma Chang stage was derived from radiological reports and lumbar puncture fluid information, other tumors were given a stage using an equivalent definition.

**Histological and genetic analysis.** Histological (including MiB1, Ki67, Glial fibrillary acidic protein (GFAP), INI-1, Isocitrate dehydrogenase-1 (IDH-1), Neuron specific enolase (NSE), S-100, BAF47, BRAF fusions, P53) and genetic data (MYC status and medulloblastoma sub-type), where available, were collected from local sites and are found in Supplementary Document 1.

Medulloblastomas were analyzed for histological type, subgroup, and MYC and MYCN amplification status were determined by protocols established at Newcastle University<sup>22–24</sup>. Medulloblastoma histology was centrally reviewed at the Royal Victoria Infirmary. Data are summarized in Supplementary Document 1.

**Statistical analysis.** All statistical analyses were performed in R (3.6.1) with significance defined at  $p < 0.05$ , and Bonferroni correction for multiple comparisons used where appropriate.

The data processing pipeline used in this study is summarized in Fig. 1.

**Univariate statistical analysis.** Data normality was assessed using a Shapiro–Wilk test. Subsequently, differences in clinical and imaging features between high- and low-grade tumors were assessed using unpaired t-tests or Mann–Whitney *U* tests, where appropriate. Area under the ROC curve (AUC) values were calculated for each imaging feature for high/low grade discrimination. Differences in high-/low- risk (defined below) participants were assessed using unpaired two-tailed t-tests or Mann–Whitney *U* test, depending on data normality. A chi squared test was performed to assess for differences in cohort characteristics including surgical resection rates between high and low risk clusters (described below).

After unsupervised clustering (described below), further Mann–Whitney *U* tests were performed to assess for differences in imaging features between low grade tumors in low and high-risk categories, and between alive high-grade tumors in low- and high-risk categories.

**Survival and correlation analysis.** Univariate Cox-regression was performed with each individual imaging feature, clinical data, and tumor grade used to assess survival hazard coefficients. Tumor grade and type were not used in the analysis detailed below.

Iterative Bayesian survival analysis was undertaken using the iterative BMA<sub>surv</sub> package in R using fivefold stratified cross validation to determine the posterior probabilities and coefficients of the top 5 imaging features that best describe the survival data<sup>25</sup>. Iterative analysis including up to 15 data features in combination at any one time.

**Unsupervised and supervised machine learning.** K means clustering was performed with the imaging features from Bayesian survival analysis, with the optimal number of clusters determined from the largest average silhouette width. Groups were clustered into high and low risk groups, and subsequently used for further Kaplan–Meier analysis to assess for differences in survival between clusters.

Supervised machine learning using the aforementioned Bayesian features was used to predict high/low risk groupings using the Orange toolbox (Orange) in Python (3.6), with Random Forest, a single layer Neural Network, and a support vector machine used. Validation of classifiers was performed using tenfold stratified cross-validation.

Clinical and imaging data were subset into Whole Brain (WB), and Region of Interest (ROI) features, and tumor volume and used for supervised learning. Principal component analysis was used to reduce data dimensionality with 95% of data variance or N-1 (where N is the size of the smallest group) used. The top 5 Bayesian features were also used as input into the classifiers, with no further principal component analysis performed. Classifier performance was determined from the classifier accuracy (% correctly classified cases) and F-statistic.

## Results

A total of 69 patients were analyzed in this study with 33 imaging features, including tumor volume, derived per patient. Example tumor anatomical, diffusion, and perfusion imaging can be seen in Fig. 2. The survival curve for the whole cohort is seen in Fig. 3A, showing 75% overall survival at 5 years from diagnosis.

**Diffusion and perfusion imaging can detect differences between tumor grade.** Univariate statistical analysis showed significant differences in both whole brain and ROI imaging features between all high and low-grade tumors (feature with highest AUC = ADC mean (0.82) range: 0.63–0.82) full results detailed in Supplementary Table 2. Comparative normal appearing grey and white matter imaging results are detailed in Supplementary Tables 3.

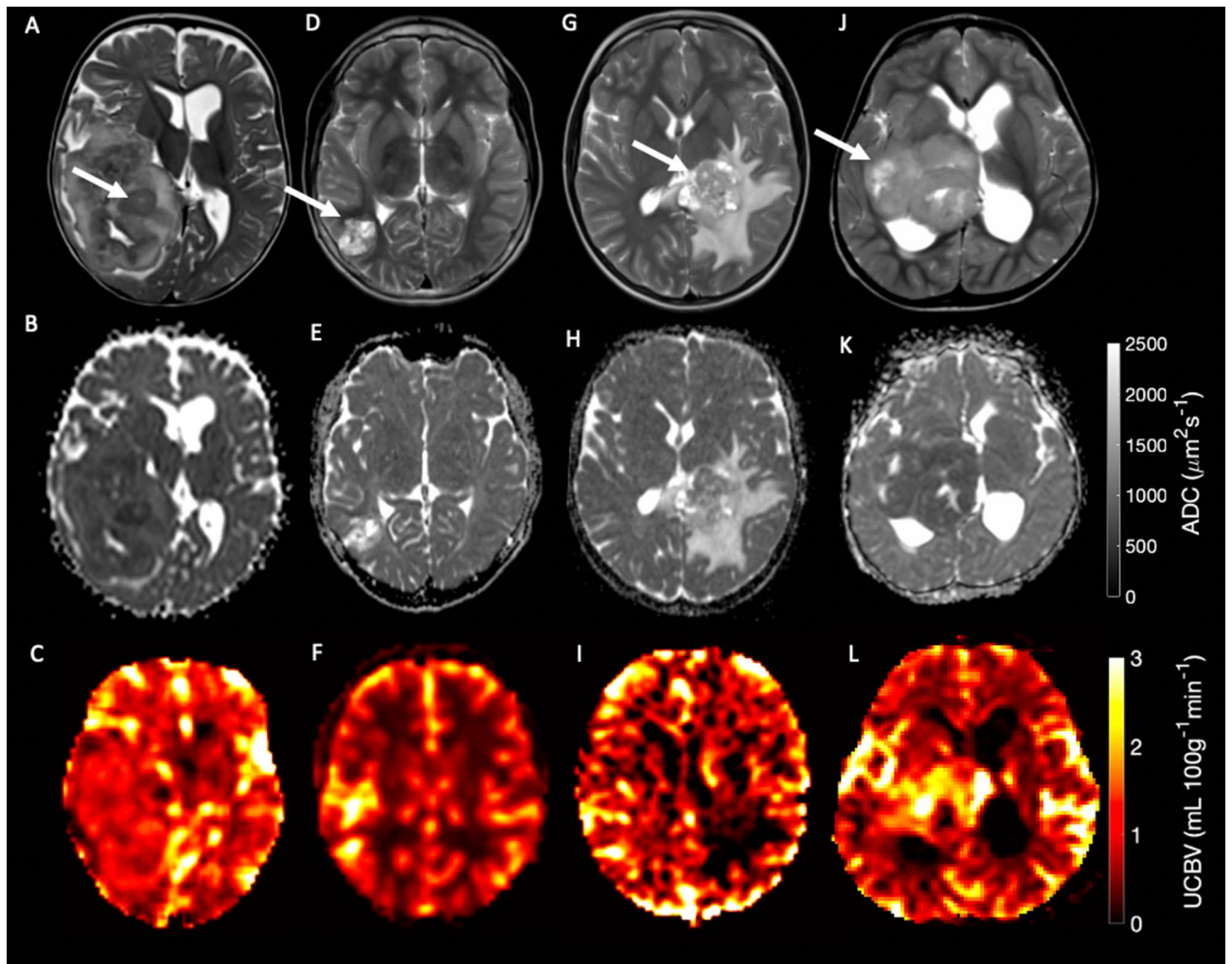
**Perfusion imaging plays a key role in assessing survival in pediatric brain tumors.** Whole cohort univariate cox regression revealed a number of imaging features with significantly elevated hazard ratios (HR), for example Uncorrected CBV ROI mean (HR 3.1, Confidence Intervals (CI) 1.5–6.6,  $p = 0.003$ ), full results detailed in Table 1.

Bayesian analysis revealed the 5 most likely features to predict survival (probability that the feature coefficient is greater than 0, posterior coefficient) to be uCBV ROI mean (96%, 0.85), K2 ROI mean (39%, -0.17), uCBV whole brain mean (40%, 0.3), tumor volume (27%, 0.05), and ADC ROI kurtosis (20%, 0.02). Full results detailed in Table 2.

**Unsupervised clustering detects distinct groups with significantly different survival and imaging characteristics.** Using the Bayesian imaging features, k means clustering revealed two distinct clusters, shown in Fig. 3B, which when combined with Kaplan–Meier analysis revealed a significant difference between a high and low risk population (see Fig. 3C,  $p = 0.0015$ —overall survival for high and low risk = 55% and 90%, respectively). Cox regression revealed an elevated Hazard Ratio (HR 5.6, confidence intervals 1.6–20.1,  $p < 0.001$ ) for the high-risk cluster, relative to the low-risk cluster.

Further univariate analysis of each cluster showed significant differences in a number of imaging features, for example elevated ADC kurtosis in high vs low risk clusters ( $10.1 \pm 5.3$  vs  $4.3 \pm 1.8$ ,  $p < 0.001$ , respectively). A combination of both high- and low-grade tumors were found in both clusters, all other results detailed in Table 3.

The patient characteristics, split by risk cluster, is provided in Supplementary document 1. From this data, there was no significant difference between the low- and high-risk clusters in age at diagnosis (low-risk cluster mean age 7.3 years, high-risk cluster mean age 9 years, Mann–Whitney test  $p > 0.05$ ), extent of surgical resection (complete macroscopic resection 18 cases in low-risk cluster, 19 cases in high-risk cluster, chi squared test,  $p > 0.05$ ) or presence of metastatic disease (low-risk: focal 21 versus metastatic 7; high-risk: 18 focal versus 11 metastatic; chi squared test  $p > 0.05$ ). However, the low-risk cluster contained a greater proportion of low-grade tumors (23 low-grade, 10 high-grade) than the high-risk cluster (7 low-grade; 26 high-grade), chi squared test  $p < 0.05$ . Whilst detailed data was not collected on adjuvant treatment with radiotherapy and chemotherapy, national treatment guidelines would have dictated that the high-risk cluster would in general have received more intense treatment regimens due to their generally higher grade.



**Figure 2.** Example  $T_2$  weighted, diffusion, and perfusion imaging of Ependymoma (A–C respectively), Pilocytic astrocytoma (D–F, respectively), choroid plexus carcinoma (F–H respectively) and a Glioblastoma (I–K, respectively). Tumor regions are highlighted with white arrows.

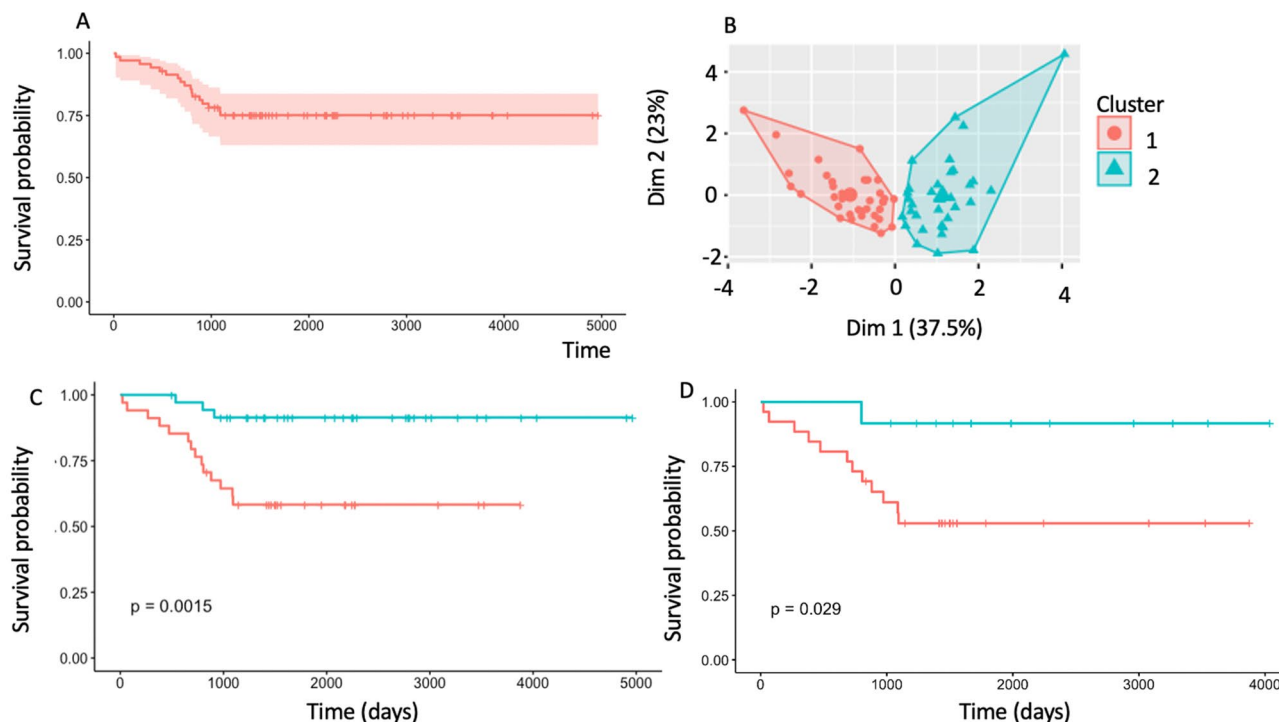
**Supervised machine learning can be used to distinguish between high/low risk clusters.** Supervised machine learning using imaging features showed that the Bayesian features combined with a single layer neural network, after stratified tenfold cross validation, provided the most accurate classification of high- and low-risk patients (accuracy = 98%, F-statistic = 0.98). Classifier accuracy ranged from 90 (logistic regression) to 98% (single layer neural network).

**There is a distinct difference in survival between high- and low-risk high-grade tumors.** Kaplan–Meier analysis of high-grade tumors in the high and low risk clusters revealed a significant difference in survival ( $p < 0.05$ ) with a hazard ratio of 7 (0.9–53 lower and upper bounds, respectively). The Kaplan–Meier curves for high grade tumors in both clusters can be seen in Fig. 3D. There was no detectable difference in survival between the high- and low- risk groups within the low-grade tumors ( $p > 0.05$ ). Imaging of example cases by risk and grade given in Fig. 4.

Qualitative sub-group analysis of histology and genetics between low- and high-risk medulloblastomas revealed no significant differences between MYC amplification or groupings. The high-risk cluster exhibited a trend toward having a larger number of high Chang stage Medulloblastomas (M3 = 6, M2 = 4, M1 = 3) in comparison to the low-risk cluster (M2 = 1, M1 = 1, M0 = 2)—data shown in supplementary document 1.

## Discussion

This study has shown the power of combining diffusion and perfusion imaging with machine learning to predict survival risk in a mixed cohort of pediatric brain tumors. A handful of studies have previously looked at assessing survival with one of the aforementioned imaging techniques<sup>14,16,27</sup>; however, here we have shown the utility of combined diffusion-perfusion measures to provide advanced modelling of survival. The univariate results assessing low/high grade suggested a number of key diffusion and perfusion features for the discrimination between



**Figure 3.** (A) Overall survival curve for the cohort, (B) K Means clustering survival results showing two distinct clusters, (C) Kaplan–Meier curve for the two clusters showing a significant difference in survival, 1 = High risk, 2 = Low risk, (D) Kaplan–Meier curves for high-grade low-risk (green) and high-risk (red) patients showing a significant difference in survival from imaging at diagnosis.

Feature	Beta	Hazard ratio	95% confidence interval	Significance
CBV ROI uncorrected mean	1.13	3.1	1.5–6.6	p = 0.003
CBV uncorrected standard deviation	– 1.12	0.33	0.11–0.99	p = 0.05
K2 ROI mean	– 2.02	0.13	0.03–0.63	p = 0.011
CBV uncorrected whole brain mean	1.12	3.02	1.06–8.91	p = 0.04

**Table 1.** Cox regression results.

Feature	Probability (%)	Posterior coefficient
Tumor volume	27	0.05
CBV ROI uncorrected mean	96	0.85
K2 ROI mean	39	– 0.17
ADC ROI Kurtosis	20	0.02
CBV uncorrected WB mean	40	0.3

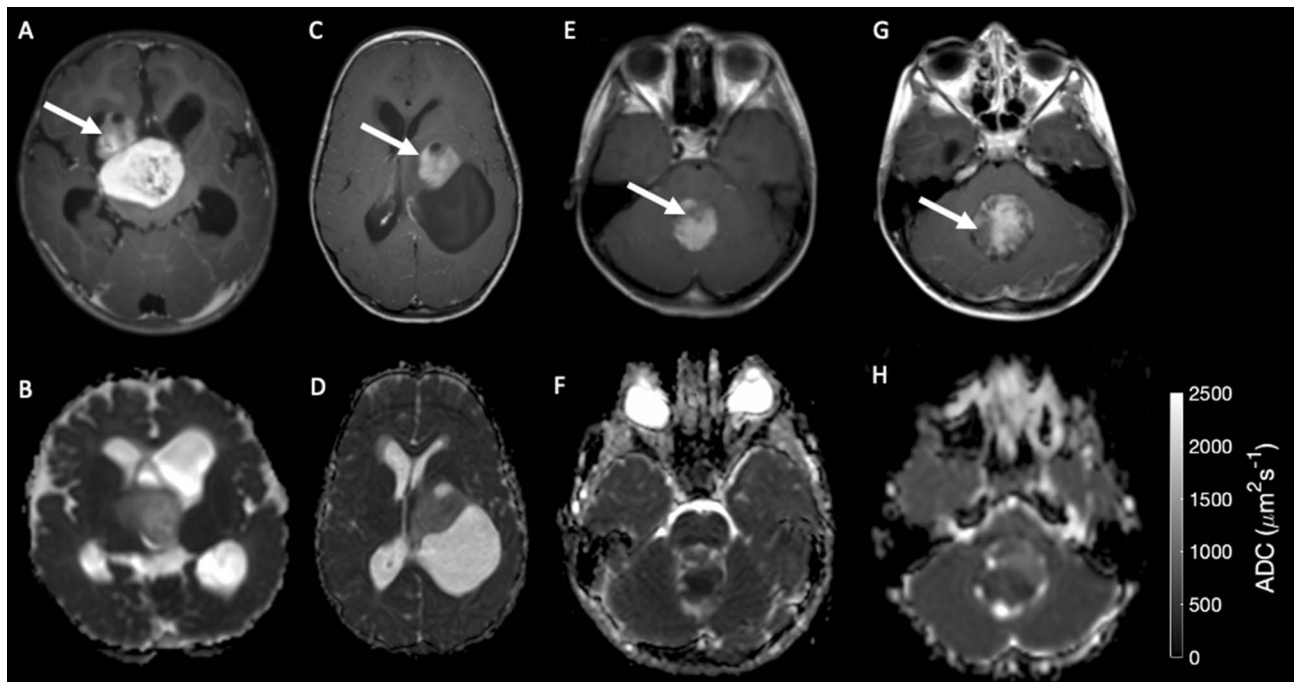
**Table 2.** Bayesian survival results.

groups, however most had a poor AUC. Therefore, this represented an ideal situation for the use of machine learning to combine these features to provide highly accurate classifiers to solve this challenge.

Interestingly, the majority of parameters predicting survival were from the perfusion imaging which is not currently part of routine clinical practice in many centers. DWI has become a standard method for investigating childhood brain tumors and low ADC is seen as being a marker for higher cellularity and grade which would be associated with poorer survival. The current study substantiates this but shows that DSC-MRI may be an even better modality for predicting survival. The importance of the vessel leakiness parameter K in survival prediction also implies that DSC-MRI may have advantages in survival prediction beyond that available from methods which do not include the injection of contrast agent such as ASL. Furthermore, for medulloblastomas, clustering demonstrated a reasonable separation of high (M1 to M3) from low (M0) Chang stage<sup>28</sup> tumors

Feature	Low risk	High risk	Significance
Male: female	16:19	19:15	N/A
Low: high grade	23:12	7:27	N/A
Censored: events	32:3	20:14	N/A
Tumor volume (cm <sup>3</sup> )	2.3 + 2.8	5.6 ± 7.0	p = 0.015
ROI ADC Kurtosis	4.3 + 1.8	10.1 + 5.3	p < 0.001
ROI ADC Skewness	0.1 ± 1.0	2.1 ± 1.0	p < 0.001
ROI K2 mean (min <sup>-1</sup> )	0.0018 ± 0.0027	- 0.005 ± 0.002	p < 0.001
ROI CBV uncorrected standard deviation (mL 100 g <sup>-1</sup> min <sup>-1</sup> )	1.44 + 0.74	0.88 ± 0.42	p < 0.001
K2 whole brain standard deviation (min <sup>-1</sup> )	0.03 ± 0.02	0.019 + 0.008	p = 0.007
CBV corrected whole brain mean (mL 100 g <sup>-1</sup> min <sup>-1</sup> )	1.14 + 0.27	1.29 ± 0.26	p < 0.02

**Table 3.** Low and high-risk cluster group features.



**Figure 4.** Example high and low risk, high and low-grade tumors. (A T1 post contrast & B ADC map) high risk and (C T1 post contrast & D ADC map) low risk Pilocytic astrocytoma, respectively showing elevated ADC skew and kurtosis in the tumor region. (E, F) high risk and (G, H) low risk medulloblastomas, respectively, showing increased ADC kurtosis. Tumor regions are highlighted with white arrows.

suggesting that these imaging features identify some properties in the primary tumor which are associated with metastatic potential.

There are a number of clinical risk factors which are commonly used to stratify treatment and it is important to consider the new imaging risk classification with these. The size of the cohort and relatively low number of events precluded a formal survival analysis with multiple risk factors. However, there was no significant difference in age, presence of metastatic disease or complete resection rate between the high- and low-risk imaging groups implying that the new imaging risk stratification will add value to these well-known risk factors. There was a significantly greater proportion of high-grade tumors in the high-risk imaging group as we would expect but the split is far from complete and a separate analysis of high-grade tumors showed a difference in survival between the high- and low-risk imaging groups showing added value beyond grade. Survival will be affected by treatment received and we did not systematically acquire information on radiotherapy and chemotherapy. However, on average the higher-risk imaging group will have received more intense treatment due to its greater proportion of high-grade tumors and so whilst adjuvant treatment is a confounder in the survival analysis, it will have acted to reduce the effect size thereby further increasing confidence in the robustness of the imaging risk classification.

The unsupervised machine learning identified two groups of tumors which did not correspond to any obvious non-imaging tumor characteristics. The credibility of these groups as being distinct entities was substantiated by the high accuracy (up to 98% on cross validation) with which the tumors could be assigned to the correct group

by a supervised learner. A number of patients in the high-risk cluster were still alive at the study end although some of these, including those from known poor prognostic groups had short follow-up times. Further analysis showed that a number of surviving high-risk low-grade tumors had imaging features similar to high grade tumors (such as elevated ADC kurtosis and CBV) and were significantly different to low-risk low grade tumors. It will be interesting to ascertain the clinical course of these tumors over longer periods of follow-up. Interestingly, it is noted that survival risk status was not always associated with tumor grade (for example grade 3 vs 4) or histology, and this may point to further underlying tumor heterogeneity within current groupings.

A particular strength of this work is that imaging features with clinical data provide a non-invasive tool that can assess risk early in the patient journey. Indeed, the use of a supervised classifier to predict risk category allows for the prospective integration of this model into a clinical decision support system—whereby radiological analysis of a small number of imaging features can rapidly identify patients that should be considered for inclusion into clinical trials for prospective evaluation and subsequent stratification. The use of in vivo imaging also has the advantage that it provides information that cannot be found from analysis of resected tissue, perfusion in particular is inherently an in vivo property.

A further strength of this study is the use of multi-site, multi-scanner data—providing reassurance that the results are robust to the natural variability that occurs in protocols and scanners within clinical practice. Using multiple centers also provided a more statistically powerful study from which clinically relevant results could be obtained.

The imaging modalities used in this study are widely available and so data acquisition should be readily achieved in routine clinical practice. The image processing and classification should be made available by integration into a clinical decision support tool which are increasingly being developed<sup>29</sup>. Indeed, the results shown above show that it is possible to stratify patients into high and low risk groups with a trained supervised neural network, therefore enabling further real-time decisions to be made with regards to appropriate clinical management and inclusion into research trials for novel therapies to aid those with the current worst prognosis.

With the current uncertainty surrounding the use of Gadolinium in clinical practice, and the inability to be used in patients with impaired renal function, future work will include the addition of ASL<sup>12</sup>, a technique to estimate perfusion without the introduction of exogenous contrast agents, as data from this technique has been shown to correlate well with DSC cerebral blood volume<sup>30,31</sup>.

The main limitation of this study is that it is based on a relatively small heterogeneous cohort treated in a diverse manner. The results should be verified in a larger prospective study but it would also be interesting to apply the same methodology to specific tumor cohorts treated in a designated manner as part of a clinical treatment trial. The multi-center nature of the study is a strength but presents particular challenges notably with regards to variations in scanner protocol which may introduce variability in results. To some extent this was mitigated by calculating the ADC values directly from the raw DWI data for all cases and performing a leakage correction on the DSC time courses to allow for differences in Gd contrast injection protocols. Regions of interest were drawn by hand which could lead to variability but a previous study in children's brain tumors by our consortium has shown good inter-rater consistency for all the ADC histogram metrics<sup>32</sup>. Finally, it is noted that there are a number of children alive at study end with high-risk tumors and currently limited follow-up. For example, a Choroid Plexus Carcinoma with a current follow-up of 1 year and a national average 5-year survival rate of 26%<sup>26</sup> and a medulloblastoma with less than 3-year follow-up and M3 Chang stage.

In conclusion, this work has demonstrated a highly novel clinical application of advanced survival modeling and machine learning to non-invasively stratify patients according to risk. This provides a non-invasive, multi-modal MRI approach to determining the malignant nature of a tumor and its potential for poor prognosis. Both diffusion and perfusion were found to be important in determining risk, with perfusion contributing to a greater extent emphasizing the importance of acquiring perfusion imaging. This work represents an important step forward in the use of machine learning to predict survival and paves the way for further clinical studies focusing on the successful identification and treatment of high-risk children with brain tumors.

Received: 4 January 2021; Accepted: 27 July 2021

Published online: 23 September 2021

## References

1. NCRA. Childhood Cancer Statistics, England Annual report 2018. *Public Heal. Engl.* (2018).
2. Vicente, J. *et al.* Accurate classification of childhood brain tumours by in vivo 1H MRS-A multi-centre study. *Eur. J. Cancer* **49**, 658–667 (2013).
3. Walker, D. A. *et al.* A multi-disciplinary consensus statement concerning surgical approaches to low-grade, high-grade astrocytomas and diffuse intrinsic pontine gliomas in childhood (CPN Paris 2011) using the Delphi method. *Neuro. Oncol.* **15**, 462–468 (2013).
4. El Doussouki, M., Gajjar, A. & Chamdine, O. Molecular genetics of medulloblastoma in children: Diagnostic, therapeutic and prognostic implications. *Future Neurol.* **14**, 20–33 (2019).
5. Johnson, K. J. *et al.* Childhood brain tumor epidemiology: A brain tumor epidemiology consortium review. *Cancer Epidemiol. Biomarkers* **23**, 2716–2736 (2008).
6. Goo, H. W. & Ra, Y. S. Advanced MRI for pediatric brain tumors with emphasis on clinical benefits. *Korean J. Radiol.* **18**, 194–207 (2017).
7. Li, S. P. & Padhani, A. R. Tumor response assessments with diffusion and perfusion MRI. *J. Magn. Reson. Imaging* **35**, 745–763 (2012).
8. White, N. S. *et al.* Diffusion-weighted imaging in cancer: Physical foundations and applications of Restriction Spectrum Imaging. *Cancer Res.* **74**, 4638–4652 (2015).
9. Detre, J. A., Rao, H., Wang, D. J. J., Chen, Y. F. & Wang, Z. Applications of arterial spin labeled MRI in the brain. *J. Magn. Reson. Imaging* **31**, 1713–1723 (2013).



10. Pedersen, M., van Gelderen, P. & Moonen, C. T. W. *Imaging Techniques for Dynamic Susceptibility Contrast-Enhanced MRI*. 95–108 (2005) [https://doi.org/10.1007/3-540-26420-5\\_7](https://doi.org/10.1007/3-540-26420-5_7).
11. Essock-burns, E. *et al.* Comparison of DSC-MRI post-processing techniques in predicting microvascular histopathology in patients newly diagnosed with GBM. *J. Magn. Reson. Imaging* **38**, 388–400 (2014).
12. van Osch, M. J. P. *et al.* Advances in arterial spin labelling MRI methods for measuring perfusion and collateral flow. *J. Cereb. Blood Flow Metab.* **38**, 1461–1480 (2018).
13. Radbruch, A. *et al.* Gadolinium retention in the dentate nucleus and globus pallidus is dependent on the class of contrast agent. *Radiology* **275**, 150337 (2015).
14. Hales, P. W. *et al.* Arterial spin labelling and diffusion-weighted imaging in paediatric brain tumours. *NeuroImage Clin.* **22**, 101696 (2019).
15. Bull, J. G., Saunders, D. E. & Clark, C. A. Discrimination of paediatric brain tumours using apparent diffusion coefficient histograms. *Eur. Radiol.* **22**, 447–457 (2012).
16. Poussaint, T. Y. *et al.* Apparent diffusion coefficient histogram metrics correlate with survival in diffuse intrinsic pontine glioma: A report from the Pediatric Brain Tumor Consortium. *Neuro. Oncol.* **18**, 725–734 (2016).
17. Morrell, G. R. & Schabel, M. C. An analysis of the accuracy of magnetic resonance flip angle measurement methods. *Phys. Med. Biol.* **55**, 6157–6174 (2010).
18. Zhang, J. Multivariate analysis in pediatric brain tumor. *Int. Radiol.* **2**, 6 (2017).
19. Avula, S. *et al.* European Society for Paediatric Oncology (SIOPE)-Brain Tumor Imaging Group. European Society for Paediatric Oncology (SIOPE) MRI guidelines for imaging patients with central nervous system tumors. *Child. Nerv. Syst.* **37**, 2497–2508 (2021).
20. Shiroishi, M. S. *et al.* Principles of T2\*-weighted dynamic susceptibility contrast MRI technique in brain tumor imaging. *J. Magn. Reson. Imaging* **41**, 296–313 (2015).
21. Grist, J. T. *et al.* Distinguishing between paediatric brain tumour types using multi-parametric magnetic resonance imaging and machine learning: a multi-site study. *NeuroImage Clin.* **25**, 102172. <https://doi.org/10.1016/j.nicl.2020.102172> (2020).
22. Schwalbe, E. C. *et al.* Novel molecular subgroups for clinical classification and outcome prediction in childhood medulloblastoma: a cohort study. *Lancet Oncol.* **18**, 958–971 (2017).
23. Schwalbe, E. C. *et al.* Minimal methylation classifier (MIMIC): A novel method for derivation and rapid diagnostic detection of disease-associated DNA methylation signatures. *Sci. Rep.* **7**, 1–8 (2017).
24. Schwalbe, E. C. *et al.* DNA methylation profiling of medulloblastoma allows robust subclassification and improved outcome prediction using formalin-fixed biopsies. *Acta Neuropathol.* **125**, 359–371 (2013).
25. Annest, A., Bumgarner, R. E., Raftery, A. E. & Yee, K. Y. Iterative bayesian model averaging: A method for the application of survival analysis to high-dimensional microarray data. *BMC Bioinform.* **10**, 1–72 (2009).
26. Berger, C. *et al.* Choroid plexus carcinomas in childhood: Clinical features and prognostic factors. *Neurosurgery* **42**, 470–475 (1998).
27. Qaddoumi, I., Sultan, I. & Gajjar, A. Outcome and prognostic features in pediatric gliomas: A review of 6212 cases from the surveillance, epidemiology, and end results database. *Cancer* **115**, 5761–5770 (2009).
28. Dufour, C. *et al.* Metastatic medulloblastoma in childhood: Chang's classification revisited. *Int. J. Surg. Oncol.* **2012**, 1–6 (2012).
29. Zarinabad, N., Meeus, E. M., Manias, K., Foster, K. & Peet, A. Automated modular magnetic resonance imaging clinical decision support system (MIROR): An application in pediatric cancer diagnosis. *J. Med. Internet Res.* **20**, 1–16 (2018).
30. Novak, J. *et al.* A comparison of pseudo-continuous arterial spin labelling and dynamic susceptibility contrast MRI with and without contrast agent leakage correction in paediatric brain tumours. *Br. J. Radiol.* **92**, 1094 (2019).
31. Choi, J. W. & Moon, W. J. Gadolinium deposition in the brain: Current updates. *Korean J. Radiol.* **20**, 134–147 (2019).
32. Novak, J. *et al.* Classification of paediatric brain tumours by diffusion weighted imaging and machine learning. *Sci. Rep.* **11**, 1–8 (2021).

## Author contributions

Study conception and design: A.P., R.G., B.P., S.B., S.C. Data analysis: J.G., S.W., A.P. Manuscript writing: J.G., A.P. Manuscript reviewing: J.G., S.W., C.B., H.R., L.M., B.P., S.B., S.C., D.M., T.A., D.A., S.A., R.G., A.P., A.O., S.P., J.N., L.A.

## Funding

We would like to acknowledge funding from The CCLG and Little Princess Trust (CCLGA 2017 15) who funded Dr James Grist, Action Medical Research and the Brain Tumor Charity (GN2181), Children with Cancer (15/188), Birmingham Children's Hospital Research Foundation, Help Harry Help Others, Poppyfields, Children's Research Fund, Cancer Research UK and EPSRC Cancer Imaging Programme the Children's Cancer and Leukaemia Group (CCLG) in association with the MRC and Department of Health (England) (C7809/A10342), the Cancer Research UK and NIHR Experimental Cancer Medicine Centre Paediatric Network (C8232/A25261), the Medical Research Council—Health Data Research UK Substantive Site and Help Harry Help Others charity. Professor Peet is funded through an NIHR Research Professorship, NIHR-RP-R2-12-019. Stephen Powell gratefully acknowledges financial support from EPSRC through a studentship from the Physical Sciences for Health Centre for Doctoral Training (EP/L016346/1). Theodoros Arvanitis is partially funded by the MRC (HDR UK). We would also like to acknowledge the MR radiographers at Birmingham Children's Hospital, Alder Hey Children's Hospital, the Royal Victoria Infirmary in Newcastle and Nottingham Children's Hospital for scanning the patients in this study. We would also like to thank Selene Rowe at Nottingham University Hospitals NHS Trust for help with gaining MRI protocol information.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-96189-8>.

**Correspondence** and requests for materials should be addressed to A.C.P.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021